

RICE UNIVERSITY

**Noise Processes in Atomic-Scale Junctions and Two-Dimensional Topological Insulators**

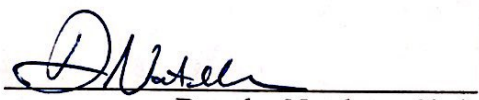
by

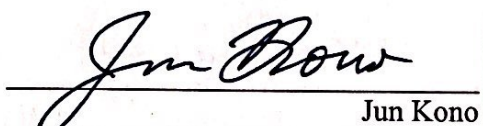
**Loah Ambrose Stevens**

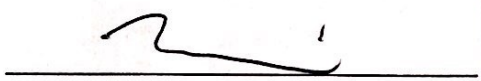
A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

**Doctor of Philosophy**

APPROVED, THESIS COMMITTEE

  
\_\_\_\_\_  
Douglas Natelson, Chair  
Professor of Physics and Astronomy

  
\_\_\_\_\_  
Jun Kono  
Professor of Electrical and Computer  
Engineering

  
\_\_\_\_\_  
Pengcheng Dai  
Professor of Physics and Astronomy

HOUSTON, TEXAS  
August 2019

# ABSTRACT

## Noise Processes in Atomic-Scale Junctions and Two-Dimensional Topological Insulators

by

Loah Ambrose Stevens

Standard transport measurements, focusing on the first moment of the current, are crucial for understanding the behavior of a system as a function of factors such as applied voltage or current, temperature, or external fields. An even more detailed picture may be procured from the second moment of the current, the electronic noise. While the conductance derived from the first moment provides the average state of the system, the electronic noise describes how quantities such as current, voltage, or resistance fluctuate about their average values. These fluctuations respond to a variety of factors within different systems and can thus reveal information not evident by transport measurements alone. In this work, we employ noise, particularly shot noise, to study the behavior of atomic-scale gold junctions and two-dimensional topological insulators.

Chapter 1 provides an introduction to quantum transport and examines how the concept of conductance evolves as the system size decreases from the macroscopic level to the few-channel limit. Chapter 2 describes the basics of electronic noise, specifically Johnson-Nyquist thermal noise,  $1/f$  or flicker noise, and shot noise. This chapter also details the expected behavior of shot noise as relates to bias, temperature, sample size, and interaction effects. Chapter 3 introduces two-dimensional topological insulators (2DTIs), beginning with the quantum Hall effect and building to the theory of quantum spin Hall insulators. Chapter 4 outlines the methods used for

the noise studies of the following chapters. Chapter 5 describes a study of shot noise in STM-style gold junctions. Shot noise was found to obey the finite temperature Landauer-Büttiker model of noise at low applied biases, but at high biases, the derived Fano factors were enhanced, which was attributed to either a bias-dependent channel-mixing mechanism or interactions between the conduction electrons and nonequilibrium phonon populations. Chapter 6 describes two studies of noise processes in InAs/GaSb quantum wells (QWs). In the first, RF noise measurements of silicon-doped InAs/GaSb bar structures revealed that the differential current noise decreases with increasing bias up to some finite bias, above which it increases linearly with increasing bias as expected. The nonmonotonic trend was suppressed by perpendicular magnetic field, increased temperatures, and applied gate voltage, leading to the belief that the trend was caused by contributions by generation-recombination noise. The second study involved both low and high frequency noise measurements in InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb QW Corbino structures and aimed to investigate the noise properties of the 2D bulk and the device contacts. In both frequency ranges, at high temperatures and positive gate voltages, when the 2D bulk is conductive, the measured noise is essentially flat with increasing bias, but as temperature is reduced and the bulk is gapped out, shot noise becomes detectable. The measured noise is much smaller and with broader curvature about zero bias than expected, which can be explained by a model in which the bulk and the contacts contribute to the thermal noise, but only the contacts produce shot noise. This model produces reasonable contact resistances and accurately portrays the zero bias curvature, but at the cost of anomalously large Fano factors. The large Fano factors may be due to either contributions to the shot noise by the bulk that are not included in the model or potentially by a mechanism of positive feedback between the conduction electrons and the buildup of space charge near one of the contacts. Future studies are proposed for using noise to

probe the edge states in InAs/GaSb 2DTIs.

# Acknowledgments

I would like to thank Doug and the rest of the Natelson lab. Thank you to Doug for his guidance over the past six years. Special thank you to Panpan and Liyang for always being willing to help me debug any issues with the instruments or setup and to Charlotte for always providing a sense of camaraderie within the lab. Also thank you to Oana Jurchescu for being an amazing role model and mentor, even when I was no longer her student.

Thank you to my parents, grandparents, and brothers for being endlessly encouraging and supportive through all the hurdles of grad school and for being more assured in my abilities than I am.

Thank you to the Valhalla crew for providing a home on campus and learning opportunities outside of the lab.

Lastly, thank you to Eric for always lending a sympathetic ear, picking me up when lab nights ran late, bringing lunch or dinner to the lab, and for celebrating every little triumph.

# Contents

<b>ACKNOWLEDGMENTS</b> .....	<b>iv</b>
<b>CONTENTS</b> .....	<b>v</b>
<b>LIST OF FIGURES</b> .....	<b>viii</b>
<b>QUANTUM TRANSPORT</b> .....	<b>1</b>
1.1 CONDUCTION AT THE MACROSCALE .....	1
1.2 BULK CONDUCTION IN REAL SYSTEMS .....	3
1.2.1 Bloch States .....	3
1.2.2 Band Structure and Conduction .....	7
1.2.3 The Fermi-Dirac Function .....	10
1.3 LENGTH SCALES AND TRANSPORT: FROM DIFFUSIVE TO BALLISTIC .....	12
1.4 THE LANDAUER-BÜTTIKER MODEL .....	13
1.5 EXPERIMENTAL EVIDENCE OF CONDUCTION QUANTIZATION .....	17
1.6 BACK TO OHM'S LAW .....	22
<b>ELECTRONIC NOISE</b> .....	<b>25</b>
2.1 JOHNSON-NYQUIST NOISE .....	26
2.2 1/F NOISE .....	29
2.3 SHOT NOISE .....	31
2.4 POISSONIAN STATISTICS .....	32
2.5 THE FANO FACTOR .....	34
2.6 SHOT NOISE OF A NON-INTERACTING SINGLE BARRIER .....	35
2.7 SHOT NOISE OF AN N-BARRIER SYSTEM .....	41
2.8 EFFECTS OF INTERACTIONS ON SHOT NOISE .....	44
<b>TOPOLOGICAL INSULATORS</b> .....	<b>48</b>

3.1	TOPOLOGICAL ORDER .....	49
3.2	THE QUANTUM HALL EFFECT .....	50
3.2.1	<i>The Classical Hall Effect and the Discovery of the Quantum Hall Effect</i> .....	50
3.2.2	<i>Quantum Approach to the Quantum Hall Effect</i> .....	56
3.2.3	<i>Edge Modes</i> .....	61
3.2.3	<i>The Role of Disorder and Laughlin's Gedankenexperiment</i> .....	64
3.3	TOPOLOGICAL INVARIANTS .....	69
3.3.1	<i>The TKNN Model of the Quantum Hall Effect</i> .....	69
3.3.2	<i>The Chern Number</i> .....	73
3.4	TOPOLOGICAL INSULATORS: THE QUANTUM SPIN HALL EFFECT .....	74
3.4.1	<i>Time Reversal Symmetry</i> .....	75
3.4.2	<i>The Haldane Model</i> .....	76
3.4.3	<i>Z<sub>2</sub> Topological Invariant</i> .....	79
3.4.4	<i>The Quantum Spin Hall Effect</i> .....	83
3.5	TWO-DIMENSIONAL TOPOLOGICAL INSULATORS .....	86
3.6	SCATTERING MECHANISMS IN TWO-DIMENSIONAL TOPOLOGICAL INSULATORS.....	93
3.7	PREDICTIONS OF NOISE IN TWO-DIMENSIONAL TOPOLOGICAL INSULATORS AND INITIAL EXPERIMENTS .....	98
	<b>EXPERIMENTAL METHODS.....</b>	<b>107</b>
4.1	FABRICATION OF INAs/GA <sub>0.68</sub> IN <sub>0.32</sub> SB QUANTUM WELL STRUCTURES .....	107
4.2	CRYOSTAT SYSTEM .....	108
4.3	CONDUCTANCE MEASUREMENTS .....	109
4.4	RF SHOT NOISE MEASUREMENTS.....	109
4.4.1	<i>Reflection Coefficient Considerations: Equivalent Circuit</i> .....	111
4.5	LOW-FREQUENCY NOISE SPECTROSCOPY .....	114
	<b>CURRENT NOISE ENHANCEMENT VIA NONEQUILIBRIUM PHONON BACKACTION IN ATOMIC-SCALE AU JUNCTIONS.....</b>	<b>117</b>
5.1	INTRODUCTION .....	117

5.2 EXPERIMENTAL SETUP .....	119
5.3 RESULTS AND ANALYSIS .....	121
5.3.1 <i>Low Bias Data</i> .....	121
5.3.2 <i>Bias Dependence</i> .....	126
5.3.3 <i>Candidate Mechanisms for Enhanced Noise</i> .....	128
5.4 CONCLUSIONS AND FURTHER PROGRESS.....	138
<b>NOISE MEASUREMENTS IN INAS/GASB QUANTUM WELL STRUCTURES .....</b>	<b>140</b>
6.1 BAND STRUCTURE OF INAS/GASB QUANTUM WELLS .....	141
6.1.1 <i>Electron-Hole Hybridization</i> .....	144
6.1.2 <i>Tuning the Band Structure in InAs/GaSb Composite Quantum Wells</i> .....	147
6.2 THEORETICAL PREDICTION OF THE QUANTUM SPIN HALL EFFECT IN INVERTED TYPE-II SEMICONDUCTORS	151
6.3 EXPERIMENTAL EVIDENCE OF THE QUANTUM SPIN HALL STATE IN INAS/GASB .....	155
6.4 INITIAL NOISE MEASUREMENTS OF SILICON-DOPED INAS/GASB DEVICES .....	164
6.4.1 <i>Sample Structure and Measurement Setup</i> .....	167
6.4.2 <i>Results</i> .....	169
6.4.3 <i>Discussion</i> .....	173
6.5 NOISE PROCESSES IN INAS/GA <sub>0.68</sub> IN <sub>0.32</sub> SB CORBINO STRUCTURES .....	177
6.5.1 <i>Device Structure and Measurement Setup</i> .....	180
6.5.2 <i>Results and Analysis</i> .....	182
6.5.3 <i>Conclusions</i> .....	192
6.6 FUTURE DIRECTIONS .....	193
<b>APPENDICES .....</b>	<b>197</b>
APPENDIX 1 1838 WAFER COMPOSITION .....	197
APPENDIX 2 DETAILED PROCEDURES FOR FABRICATION OF INAS/GA <sub>0.68</sub> IN <sub>0.32</sub> SB DEVICES.....	198
APPENDIX 3 PROCEDURE FOR COOLING DOWN JANIS CRYOSTAT .....	203
APPENDIX 4 CONVERSION FROM LOCK-IN TO REAL UNITS FOR RF MEASUREMENTS .....	208



# List of Figures

Figure 1.1: a) Top: an example of a periodic potential (in 1D). Bottom: The single-particle states for electrons in a periodic lattice. b) Left: Energy as a function of wavevector for a free electron. Right: In contrast, when the electron is subject to the constraints of a lattice, the allowed energies form specific bands within the Brillouin zone.....	6
Figure 1.2: a) Band structure of a metal at $T=0$ . b) At finite temperature, electrons near the Fermi level can be excited with little energy cost, leading to metallic conduction. ....	9
Figure 1.3: a) Fermi-Dirac distribution for $\mu = 5$ eV. b) The thermal broadening function, a measure of how quickly the occupancy of the single-particle states change with energy.. ..	11
Figure 1.4: a) One-dimensional ballistic conductor, capped by two contacts at chemical potentials $\mu_L$ and $\mu_R$ . b) Dispersion curves of right-moving and left-moving electrons. ....	14
Figure 1.5: Two-terminal conductance of GaAs/AlGaAs 2dEG as the width is tuned by a top gate. ....	18
Figure 1.6: a) An example of an STM-style measurement. b) An example of a MBJ setup.. ....	19
Figure 1.7: Examples of conductance steps seen as the width of metallic atomic-scale junctions are varied. (a) Platinum in vacuum at 1.2K. (b) copper, gold, and sodium at helium temperatures. ....	20
Figure 1.8: Examples of conductance histograms collected for gold, sodium, and copper.. ....	21
Figure 1.9: Possible paths for an electron travelling through a conductor with two barriers with transmission probabilities $T_1$ and $T_2$ .....	23
Figure 2.1: a) Voltage-squared as a function of resistance for several types of conductors. b) Apparent power versus temperature. Johnson found a linear dependence of mean-square voltage fluctuations on both resistance and temperature.....	26
Figure 2.2: Nyquist's system of two resistors used to theoretically derive Johnson-Nyquist noise. ....	27
Figure 2.3: a) J.B. Johnson found excess noise at low frequencies in the thermionic emission of tungsten filaments. b) Voltage noise versus frequency in electromigrated gold junctions. c) $1/f$ noise found in a. Scott Joplin piano rags, and b. classical, c. rock, and d. news and talk radio stations.....	30
Figure 2.4: Examples of Poissonian distributions for three different average number of events..	33
Figure 2.5: a) Conductivity and current noise as gate voltage adjusts the width of a 2DEG. b) Excess noise in 27 gold contacts. Solid line is case of single partially transmitted mode. c) Fano factor of ensembles of silver (left) and gold (right) atomic contacts.....	37
Figure 2.6: Fano factor for current noise measured in gold (filled circles) and aluminum (open circles).. ....	38
Figure 2.7: a) Conductance traces and calculated transmission coefficient ranges for silver, gold, aluminum, and platinum point contacts formed by the mechanical break junction method.	

b) Transmission coefficient estimates for the four most contributing channels in nickel oxide break junctions.....	39
Figure 2.8: a) Shot noise in an Au/hBN/Au tunnel junction as a function of temperature. The curvature about zero-bias sharpens as temperature decreases. b) Zero bias curvature of shot noise can be used for thermometry.....	41
Figure 2.9: Shot noise in the incoherent limit modelled as a chain of N diodes in series, each with its own capacitance.....	43
Figure 2.10: Fano factor as a function of resistor length, broken into regimes in terms of the elastic mean free path, $l$ , the phase coherence length $L_\phi$ , the electron-electron inelastic scattering length, $L_{e-e}$ , and the electron-phonon inelastic scattering length, $L_{e-ph}$ .....	45
Figure 3.1: Example of a typical Hall bar setup.....	52
Figure 3.2: Under very large perpendicular magnetic field, electrons in a Hall bar begin to move in cyclotron orbits with chiral transport along the edges. ....	53
Figure 3.3: a) First evidence of the Quantum Hall effect. b) Another example of oscillatory longitudinal resistance and quantized Hall resistance by Cage et al. ....	55
Figure 3.4: a,b) Magnetic potential and hard wall potential as a function of $y$ for two values of $k_x$ , including the wavefunction and energy level of the ground state for the Hamiltonian from Equation 3.13. c) Ideal Landau levels for a system with boundaries.. ....	58
Figure 3.5: Density of states of ideal Landau levels in case of zero (a) and nonzero (b) electric field.....	59
Figure 3.6: a) Potential similar to Equation 3.8. b) Available states are filled, setting the Fermi level. c) Applying a potential difference across the two sides of the sample. d) Accounting for the tilt of the Landau level under an electric field. e) Arbitrary potential across the interior of the sample. f) Generalizing the model to multiple levels.....	62
Figure 3.7: Laughlin's gedankenexperiment.....	65
Figure 3.8: a,b) A portion of the ribbon from Laughlin's gedankenexperiment. c) The density of states become broader in the presence of disorder, leading to a combination of localized states in the interior and extended states around the edges. ....	68
Figure 3.9: Quantum Hall systems with realistic boundary conditions a) A simple Hall bar with source and drain on left and right and voltage leads on the top and bottom. b) An equivalent geometry where current and voltage leads are replaced by loops of the same material as the Hall bar with flux threaded through each. ....	71
Figure 3.10: Hexagonal lattice of graphene. ....	76
Figure 3.11: Due to the periodic conditions of the Brillouin zone, the unit vector must wrap around the sphere an integer number of times.....	78
Figure 3.12: a) Two options for electronic dispersion between two boundary Kramers degenerate points. b) Another example of the energy dispersions.. ....	81
Figure 3.13: Comparison of the QHE (left) and the QSHE (right). ....	84
Figure 3.14: a) Photons traveling through a lens with antireflective coating. Light reflecting from the top and bottom surfaces interfere destructively. b) Likewise, the possible paths for	

an electron scattering off an impurity interfere destructively such that ballistic transmission is protected.....	85
Figure 3.15: a) HgTe QWs sandwiched between CdTe. b) Relevant bands for HgTe and CdTe. c) Band positions in the quantum well structure. ....	88
Figure 3.16: Energy of the E1 and H1 bands as a function of well thickness. The point at which the bands cross is considered a topological phase transition from the trivial to the nontrivial $Z_2$ topology. ....	89
Figure 3.17: a) Proposed setup for a QSHE experiment b) Predictions for conductance as a function of gate voltage for a thin well (left) and thick well (right).....	90
Figure 3.18: First experimental evidence of the QSHE in HgTe quantum wells. Four-terminal conductance as a function of gate voltage for well with 5.5 nm thickness (black), 7.3 nm thickness with 1 and 0.5 $\mu\text{m}$ length between leads (red and green lines), and 7.3 nm thickness with 20 $\mu\text{m}$ length between leads (blue line). Inset) Comparison of inverted-gap devices with 1 $\mu\text{m}$ length at 30 mK and 1.8K. ....	92
Figure 3.19: Model of backscattering in the QSHE due to the Kondo effect.....	94
Figure 3.20: a) Model of QSHE backscattering due to strong potential scatterers $V_N$ on the edge in the presence of a magnetic field. b) Proposed conductance of edge modes as a function of magnetic field for different values of scattering site potentials.....	95
Figure 3.21: a) Electrons moving in a helical edge tunnel into and out of puddles formed by inhomogeneous charge distribution within the bulk. Electrons may undergo inelastic backscattering within the puddles. b) Proposed edge conduction changes due to inelastic backscattering.....	96
Figure 3.22: a) Proposed experimental setup for a 2DTI with two helical edge modes. b) Charged-pair (left) and spinful-pair (right) tunneling. Both processes lead to opposite cross correlations between spin-up and spin-down tunneling currents. ....	99
Figure 3.23: Contour plots of Fano factor as a function of coupling strength $\phi_L$ and normalized spin flip time $\eta$ for a) one puddle in the absence of energy relaxation, b) continuous distribution of puddles in the absence of energy relaxation, c) one puddle with strong energy relaxation, and d) continuous distribution of puddles with strong energy relaxation. ....	101
Figure 3.24: Shot noise in HgTe QWs. Fano factor is within the range $0.1 < F < 0.3$ , consistent with disordered multi-mode transport. ....	103
Figure 3.25: Conductance (a,c) and $1/f$ noise (b,d) for TI/ $\text{Al}_2\text{O}_3$ /Co tunnel junctions, with bottom TI of $\text{Bi}_2\text{Te}_3$ (a,b) and $\text{Bi}_2\text{Se}_3$ (c,d). Triangles in $1/f$ noise plots indicate points of inflection in the conductance and dashed lines indicate features in the band structure.....	104
Figure 3.26: a) Sketches of backscattering current as a function of time. b) Backscattering Fano factor as a function of $p$ and $q$ for different spin values.....	106
Figure 4.1: Fabrication procedure for $\text{InAs}/\text{Ga}_{0.68}\text{In}_{0.32}\text{Sb}$ Corbino devices. a) Initial wafer structure b) Defining the mesa c) Removing the InAs cap layer d) Defining the contacts e) Depositing Ti/Au for the contacts f) Adding a protective layer of $\text{Al}_2\text{O}_3$ g) Opening windows to the contacts.....	108

Figure 4.2: Setup diagram for RF noise measurements..	110
Figure 4.3: a) Equivalent circuitry for a noise source. b) Reflectance measurement for determining the actual power delivered from the sample to the amplifier chain.	113
Figure 4.4: Low-frequency noise measurement setup.	115
Figure 4.5: Example of a finite bias spectrum fit to Equation 4.1. Spectrum includes contributions from both Johnson-Nyquist and shot noise, which are both white in frequency except for a capacitive roll off toward higher frequencies.	116
Figure 5.1: Excess noise power as a function of scaled bias for low bias (top) and up to high bias (bottom) for three different conductance values.	118
Figure 5.2: Example data and analysis for low bias measurements (160 mV). a) Conductance histogram with typical peaks at integer values of $G_0$ . b) Ensemble averaged Fano factor as a function of conductance with the theoretical minimum Fano factor line shown in black. c) Transmittance histograms of the first six channels for $G = 1.08, 2.18, \text{ and } 3.07G_0$ , with respective Fano factors of 0.07, 0.09, and 0.09, calculated using the inequality in Equation 5.4.	123
Figure 5.3: Ensemble averaged Fano factor as a function of conductance for a sample of increasing biases, minimum Fano factor line shown in black.	127
Figure 5.4: Ensemble averaged Fano factor as a function of bias across the junction for conductances with similar values on the minimum Fano factor line (Insets).	128
Figure 5.5: Conductance histograms for increasing bias applied bias.	132
Figure 5.6: Transmission coefficient histograms for a) $0.8G_0$ , b) $1.08G_0$ , and c) $2.18G_0$ to demonstrate how bias-dependent channel mixing could account for the enhancement of the Fano factor.	133
Figure 5.7: Excess noise as a function of $V^n$ as an attempt to pull out the dominant power law of the bias dependence of the noise.	136
Figure 6.1: Band lineups for the 6.1 Å family of semiconductors.	142
Figure 6.2: a) Composite quantum well of InAs/GaSb with AlSb barriers. b) Schematic band structure diagram of InAs/GaSb.	143
Figure 6.3: Density of states (a) and in-plane dispersion relations (b) of electrons in InAs and holes in GaSb with (solid lines) and without (dashed lines) consideration of the hybridization effect.	147
Figure 6.4: In-plane band structure of a double-gated InAs/GaSb composite quantum well under different external electric fields.	149
Figure 6.5: Phase diagram for InAs/GaSb CQW as a function of front ( $V_f$ ) and back ( $V_b$ ) gates.	154
Figure 6.6: Peaks observed in the longitudinal resistance $R_{xx}$ as a function of $V_{\text{front}}$ in InAs/GaSb CQW with $V_{\text{back}}$ between -10 V and 10 V at $B = 0$ T and $T = 0.3$ K. Inset (a) dispersion of the hybridized bands. Inset (b) The resonance peak heights as a function of the inverse carrier density.	157

Figure 6.7: a) Longitudinal resistance  $R_{xx}$  vs. front gate voltage  $V_{front}$  for various device lengths  $L = 100, 10, 4,$  and  $2 \mu\text{m}$ . b)  $R_{xx}$  vs.  $V_{front}$  for device widths  $W = 0.5, 1, 1.5,$  and  $2 \mu\text{m}$  with  $L = 2 \mu\text{m}$ . c) The gap conductance  $G$  shows a linear relationship with the device width.. 159

Figure 6.8:  $R_{xx}$  vs.  $V_{front}$  for  $L = 100 \mu\text{m}$  (a) and  $2 \mu\text{m}$  (b) as  $V_{back}$  is varied in steps of 2 V from 0 to -8 V, with  $B = 0$  T, and  $T = 20$  mK. As  $V_{back}$  is tuned to more negative values, the gap shifts to smaller values of  $k_{cross}$ , and the resistance peaks increase. The difference in the gap conductance,  $\Delta G$ , is shown as a function of  $V_{back}$  (c) and bulk conductivity (d). ..... 161

Figure 6.9: a) Phase diagram of InAs/GaSb double-gated CQWs as a function of the applied electric field  $E_z$  and the Fermi level position  $E_F$ . b) Phase diagram as a function of back gate voltage  $V_{BG}$  and top gate voltage  $V_{TG}$ . c) Top: Four-terminal longitudinal resistance as a function of top and bottom gate at  $T = 300$  mK, revealing the phase diagram of the InAs/GaSb CQW. Inset: Optical image of the Hall bar. Bottom: Resistance along lines L and R. d) In-plane magnetic field dependence of the longitudinal resistance along the same lines L and R..... 163

Figure 6.10: a) Schematic of the InAs/GaSb CQW with Si doping and the potential fluctuations induced by the dopants at the interface. b) Top: Wide conductance plateaus quantized to  $\frac{2E^2}{H}$  for the  $\pi$  bar (red) and  $\frac{4E^2}{H}$  for the Hall bar (blue). Top Inset: Plateau persists to 4 K, but at higher temperatures conductance increases due to delocalized 2D bulk carriers. Bottom: Longitudinal resistance as a function of front gate voltage for longer devices. Bottom Inset: Resistance scales linearly with edge length, indicating a coherence length  $\sim 4.4 \mu\text{m}$ . ..... 166

Figure 6.11: SEM image of one of the long devices. Insets: Optical images of long device (left) and short device (right)..... 168

Figure 6.12: a,b) Differential conductance as a function of temperature in the short device (a) and the long device (b) As expected, conductance decreases with decreasing temperature as the 2D bulk states are gapped out. At low temperatures, both devices exhibit a zero-bias conductance suppression. c,d) Differential noise power as a function of temperature for the short (c) and long (d) devices.. ..... 170

Figure 6.13: a,d) Differential conductance as a function of gate voltage for the short (a) and long (d) devices. b,c,e) Differential current noise with varying gate voltage for the short (b,c) and long (e) devices..... 172

Figure 6.14: a,b) Differential conductance with varying perpendicular magnetic field in the short (a) and long (b) devices. c,d) Differential current noise in the short (c) and long (d) devices. .... 173

Figure 6.15: Current spectral density of an example n-doped Si device, with indications of ranges of  $1/f$ , generation-recombination, and thermal noise..... 174

Figure 6.16: Assumed band positions for unstrained (left) InAs, GaSb, and InSb and in comparison, the effect of strain on InAs and  $\text{Ga}_{0.6}\text{In}_{0.4}\text{Sb}$  (right)..... 177

Figure 6.17: a) Calculated bulk band structure of InAs/ $\text{Ga}_{0.68}\text{In}_{0.32}\text{Sb}$  CQW. b) TEM image of the InAs/ $\text{Ga}_{0.68}\text{In}_{0.32}\text{Sb}$  wafer, showing the crystalline structure is intact throughout the heterostructure. .... 179

Figure 6.18: Arrhenius plots for InAs/GaSb (open squares), InAs/Ga<sub>0.8</sub>In<sub>0.2</sub>Sb (open circles), InAs/Ga<sub>0.75</sub>In<sub>0.25</sub>Sb (filled diamonds), and InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb (filled circles). ..... 180

Figure 6.19: a) Colorized SEM image of an example Corbino device. b) Diagram of device structure. c) Band structure of the CQW. .... 181

Figure 6.20: a) Zero-bias two-terminal resistance as a function of temperature. b,c) Integrated differential current noise in rf bandwidth as a function of temperature (b) and gate voltage (c)..... 184

Figure 6.21: a) Voltage noise minus the zero-bias voltage noise as a function of temperature. b) Comparison of the measured noise to the expected noise based on Equation 6.16.. ..... 185

Figure 6.22: Equivalent circuit (a) and comparison to data (b) of the model described by Equation 6.17, with  $R = 4.34 \text{ k}\Omega$ ,  $F = 5.76 \pm 0.2$ , and  $R_c = 217 \pm 6 \Omega$ . .... 188

Figure 6.23: Fano factors (a) and contact resistance (b) derived from linear fits at high bias  $S_V(I)$  data taken at low frequency for the model described by Equation 6.17..... 190

Figure 6.24: Enhanced noise model of Reklaitis and Reggiani. Electrons tunneling the barrier (a) decrease the barrier energy as seen by electrons coming from the cathode as a result of the positive feedback between space charge in the layer preceding the barrier and the transmission probability for electrons tunneling through the barrier. Consequently, more electrons will succeed in tunneling (b). .... 191

Figure 6.25: a) Comparison of spectra at highest measured current bias for a InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb Corbino device versus a WTe<sub>2</sub> bar structure with edge mode conduction. b) Proposed device configuration for multiple source-drain contact separations along the same edge. c) Another proposed device geometry for comparing noise along the sample edge (using the diagonal contacts) versus through the bulk (using the horizontal or vertical contacts). .... 194

# Chapter 1

## Quantum Transport

Much of the work of this thesis examines systems in which the electrical transport is carried out by a limited number of quantum conduction channels. In this chapter, we introduce how transport and the notion of conductance or resistance change in going from the macroscale to the quantum coherent limit. Beginning with the basics of the macroscopic model, we then consider the limit of coherent, ballistic transport and introduce the Landauer-Büttiker model and the concept of conductance quantization.

### *1.1 Conduction at the Macroscale*

Fundamental to introductory explanations of electrical conduction in bulk systems is Ohm's Law<sup>1</sup>: the current ( $I$ ) through a resistor is proportional to the voltage difference ( $\Delta V$ ) across it, most commonly recognized as

$$\Delta V = IR \quad (1.1)$$

where  $R$  is the resistance. In terms of conductance,  $G$ , Ohm's law implies

$$G = \frac{\sigma W}{L} \quad (1.2)$$

where  $\sigma$  is the electrical conductivity of the material, and  $W$  and  $L$  give the width and length of the sample. Therefore, conductance decreases with increasing length and increases with increasing width of a sample.

Normalizing Equation 1.1 for device geometry, we find

$$\mathbf{j} = \sigma \mathbf{E} \quad (1.3)$$

where  $\mathbf{j}$  is the current density,  $\mathbf{E}$  is the electric field. Conductivity is an inherent property of a material, independent of an individual sample size or geometry. To understand its origin, consider the charge carriers in a material, i.e. electrons in a metal, to be a classical, noninteracting gas, an approach first put forth by Paul Drude in 1900<sup>2</sup>. Assume the electrons are all moving at some nominal speed  $v$ , with random direction such that their average velocity is zero. If the electrons are assumed to not interact with each other, the only collisions will be with stationary defects in the material. On average, an electron will elastically scatter off a defect on a timescale of  $\tau$ , meaning the average distance travelled by an electron before encountering a defect is  $l = v\tau$ , which is known as the elastic mean free path. Looking across the entire length of the sample, the motion of the electrons is diffusive and each individual electron essentially performs a random walk.

If an external electric field  $\mathbf{E}$  or magnetic field  $\mathbf{B}$  is applied across the material, the electrons will respond to the field in between scattering events according to the Lorentz force

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (1.4)$$

For a given electric field  $\mathbf{E}$ , the result is a net drift velocity of the electrons

$$\mathbf{v}_d = \frac{q\mathbf{E}}{m}\tau \quad (1.5)$$

where  $q$  is the electron charge, and  $m$  is its mass. Current density is then

$$\mathbf{j} = n_{3d}q\mathbf{v}_d = \frac{n_{3d}q^2\tau}{m}\mathbf{E} \quad (1.6)$$

where  $n_{3d}$  is the 3d electron density. One can then define the Drude conductivity as

$$\sigma = \frac{n_{3d}q^2\tau}{m}. \quad (1.7)$$



## 1.2 Bulk Conduction in Real Systems

This deliberately simple, classical picture actually often works well in real materials, even though we now know that electrons are primarily governed by quantum mechanics. That is, in real metals or doped semiconductors, single electron states are actually described by wavefunctions that reflect the periodic environment of the crystal lattice.

### 1.2.1 Bloch States

One of the most successful models of these wavefunctions is the Bloch theorem of the allowed electronic states in a periodic lattice. For a particular lattice structure with spatially periodic potential (*Figure 1.1a*),<sup>3</sup>  $V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r})$ , we define a translation operator

$$T_{\mathbf{R}}f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R}). \quad (1.8)$$

Let  $\psi(\mathbf{r})$  be some solution to the time-independent Schrödinger equation,  $H\psi(\mathbf{r}) = E\psi(\mathbf{r})$ . Due to the periodicity of the lattice, this translation operator commutes with the Hamiltonian, meaning  $\psi(\mathbf{r})$  can also be an eigenstate of the translation operator

$$T_{\mathbf{R}}\psi(\mathbf{r}) = c(\mathbf{R})\psi(\mathbf{r}), \quad (1.9)$$

where  $c(\mathbf{R})$  is the eigenvalue of the translation operator. Two translations must commute

$$c(\mathbf{R} + \mathbf{R}') = c(\mathbf{R})c(\mathbf{R}') = c(\mathbf{R}')c(\mathbf{R}). \quad (1.10)$$

Translation must also preserve the normalization of the wavefunction, which implies  $c(\mathbf{R})$  must be some unit-magnitude complex number that carries a phase such that translations add,

$$c(\mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}, \quad (1.11)$$

where  $\mathbf{k}$  is some vector defined in terms of the reciprocal lattice, the Fourier transform of the real-space lattice that defines a complementary momentum-space. The momentum-space vector

is defined in terms of the primitive reciprocal lattice vectors,  $\mathbf{b}_i$ , as  $\mathbf{k} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3$ .

The allowed values of  $x_i$  are set by the boundary conditions on the wavefunction, so that

$$T_{\mathbf{R}}\psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}\psi(\mathbf{r}). \quad (1.12)$$

Equation 1.12 is Bloch's theorem, which reflects the constraints on the single-particle states in a periodic potential.

If we define some function  $u_{\mathbf{k}}(\mathbf{r}) \equiv e^{-i\mathbf{k}\cdot\mathbf{r}}\psi(\mathbf{r})$ , we find it has the same periodicity of the lattice:

$$\begin{aligned} u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) &= e^{-i\mathbf{k}\cdot(\mathbf{r}+\mathbf{R})}\psi(\mathbf{r} + \mathbf{R}) \\ &= e^{-i\mathbf{k}\cdot\mathbf{r}}\psi(\mathbf{r}) \\ &= u_{\mathbf{k}}(\mathbf{r}). \end{aligned} \quad (1.13)$$

If  $u_{\mathbf{k}}(\mathbf{r})$  is a function that is strictly periodic in real space with the same periodicity as the lattice, then a single-particle wavefunction guaranteed to satisfy Bloch's theorem is given by

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}). \quad (1.14)$$

This is known as a Bloch state (*Figure 1.1a*), a plane wave modulated by some function whose periodicity strictly matches that of the lattice. It is important to note that although Bloch states are still labelled by the parameter  $\mathbf{k}$ , this no longer directly corresponds to momentum as in the case of free electrons. The form of Bloch states holds even when the periodic potentials are not weak. These states are referred to as extended states, meaning they extend throughout the lattice in real space, in contrast to the electronic orbitals of a deeply-bound lattice atom, which are localized states. Scattering events can then be defined in terms of the Bloch states as impurities or defects disrupting the periodicity of the Bloch state such that is no longer a fully extended state, but rather has some finite lifetime. An elastic scattering event causes a shift from one value of  $\mathbf{k}$  to another at the same energy.

We now impose boundary conditions on the wavefunction, namely the Born-von Karman boundary conditions,  $\psi(\mathbf{r} + N_i \mathbf{a}_i) = \psi(\mathbf{r})$ , in other words the wavefunction must be periodic after a certain number of lattice sites, where  $\mathbf{a}_i$  are the primitive lattice vectors. Then Bloch's theorem implies

$$e^{iN_i \mathbf{k} \cdot \mathbf{a}_i} = e^{2\pi i N_i x_i} = 1, \quad (1.15)$$

where we have used the definition of  $\mathbf{k}$  in terms of the reciprocal lattice vectors and taken advantage of the orthogonality relationship between  $\mathbf{a}_i$  and  $\mathbf{b}_i$ . Therefore the  $x_i$  are given by

$$x_i = \frac{j_i}{N_i}, j_i = 1, 2, 3 \dots N_i, \quad (1.16)$$

and the allowed values of  $\mathbf{k}$  are

$$\mathbf{k} = \sum_{i=1}^3 \frac{j_i}{N_i} \mathbf{b}_i. \quad (1.17)$$

Therefore, for a lattice with  $N = N_1 N_2 N_3$  sites, there are  $N$  allowed values for  $\mathbf{k}$  that fulfill the boundary conditions. In  $k$ -space, the allowed volume per single-particle state is  $\left(\frac{2\pi}{L}\right)^d$ , where  $L$  is the length of the lattice and  $d$  indicates the dimensionality. If we look at the Brillouin zone of the lattice, a region in  $k$ -space defined by a set of points that are closer to the origin than to any other reciprocal lattice points, we find there are exactly  $N$  allowed  $\mathbf{k}$  values. For each of these  $\mathbf{k}$  values there are multiple solutions to the single-particle Schrödinger equation with different corresponding energy eigenvalues, numbered by what is known as a band index. Beyond the first Brillouin zone, any single-particle state confined by the set boundary conditions exactly corresponds to a state within the first Brillouin zone, just with a different band index (*Figure 1.1b*).

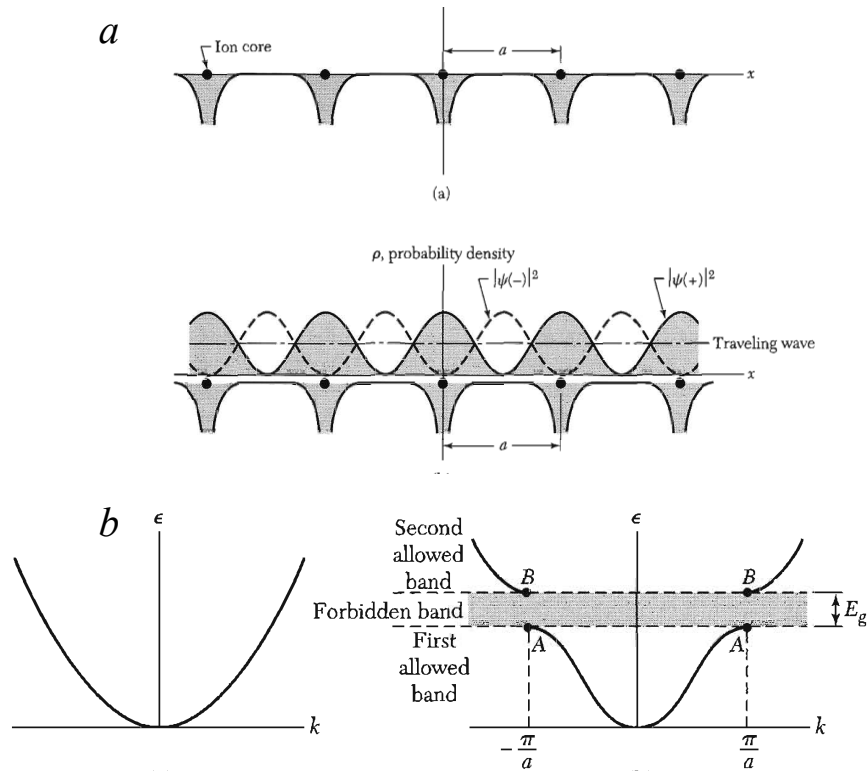


Figure 1.1: [Adapted from Ref. 3] a) Top: an example of a periodic potential (in 1D). Bottom: The single-particle states for electrons in a periodic lattice are described as Bloch states, plane waves modulated by the periodicity of the lattice. b) Left: Energy as a function of wavevector for a free electron. Right: In contrast, when the electron is subject to the constraints of a lattice, the allowed energies form specific bands within the Brillouin zone.

Energy is periodic in  $k$ -space:  $E_n(\mathbf{k} + \mathbf{G}) = E_n(\mathbf{k})$ , where  $\mathbf{G}$  is a reciprocal lattice vector. For each  $n$ , the energies of all single-particle states defined by the allowed values of  $\mathbf{k}$  are said to be within a particular band, and the range in energy covered by the states in the band is the band width.

## 1.2.2 Band Structure and Conduction

The details of the allowed single-particle states, known as the band structure, vary drastically from a metal to a semiconductor to an insulator, and the electronic properties of a material result from its band structure. A metal is defined as a material whose resistivity decreases with decreasing temperature. *Figure 1.2*<sup>4</sup> shows typical band structures for a metal at zero (*Figure 1.2a*) and finite (*Figure 1.2b*) temperature. At  $T = 0$ , there is a well-defined Fermi energy that divides the filled states from the empty ones. For  $T > 0$ , electrons near the Fermi level can be promoted into empty states above the Fermi energy, leaving behind newly empty states called holes. In metals, there are many available states near the Fermi energy, such that electronic excitations can be made with low energetic cost, leading to electronic conduction.

In the limit that the temperature is low relative to the Fermi energy (chemical potential) scale of the electrons, it is valid to consider the single-particle states as full up to within  $k_B T$  of the Fermi energy, where  $k_B = 1.3806 \times 10^{-23} \text{ JK}^{-1}$  is the Boltzmann constant. This defines the Fermi energy as well as the Fermi surface in  $k$ -space. In the limit of a bulk metal, the spacing between single-particle states due to the boundary conditions is negligible. For example, consider a  $1 \text{ cm}^3$  block of sodium, an alkali metal. The spacing between the single-particle states is given by

$$\Delta = \frac{1}{v_{3d} E_F L^d} \quad (1.18)$$

where  $v_{3d}$  is the 3d density of states,  $E_F$  is the Fermi energy, and  $L^d$  is the volume. For sodium,  $v_{3d} = 7.6 \times 10^{46} \text{ J}^{-1} \text{ m}^{-3}$ , and  $E_F = 3.2 \text{ eV} = 5.2 \times 10^{-19} \text{ J}$ , so the spacing for the  $1 \text{ cm}^3$  is  $\Delta = 1.3 \times 10^{-41} \text{ J} = 8.1 \times 10^{-23} \text{ eV}$ , which is much smaller than  $k_B T$  for all experimentally accessible temperatures. For reference, room temperature corresponds to roughly 26 meV. This means the

band is essentially continuous, and the energy cost of creating an electron-hole excitation can be arbitrarily small.

In real metals, both electron-like and hole-like excitations are quasiparticles, a term used to describe low energy (in relation to  $E_F$ ) excitations with well-defined quantum numbers (including spin, charge, band index,  $\mathbf{k}$ ). Quasiparticles are not exact eigenstates of the full many-body Hamiltonian; each quasiparticle state has some finite lifetime  $\tau$ , and therefore has some nonzero uncertainty in energy,  $\Gamma \sim \frac{\hbar}{\tau}$ , where  $\hbar = \frac{h}{2\pi} = 1.0545 \times 10^{-34}$  Js is the reduced Planck's constant. For the quasiparticle model to be a valid description of the low energy excitations, the energetic cost of the excitation must exceed  $\Gamma$ . In thermal equilibrium, this equates to  $\Gamma \ll k_B T$ .

In 1957, Lev Landau<sup>5</sup> demonstrated that within some constraints, the ground state of the interacting many-electron system is still very similar to the non-interacting case. Namely, there is a Fermi sea of filled states separated from empty states by a Fermi surface. Additionally, the low energy excitations are shown to have properties very similar to those of noninteracting quasiparticles and the energy uncertainty-lifetime relationship holds as  $T \rightarrow 0$  if  $\tau^{-1}$  is dominated by electron-electron interactions. This system is called a Fermi liquid (as opposed to the noninteracting Fermi gas). The Fermi liquid model is the standard for describing metals.<sup>6,7</sup> The effects of interactions can be added through a small number of Fermi liquid parameters. The low-energy excitation quasiparticle is not simply a solitary electron elevated above the Fermi energy; it also incorporates the collective response from all other electrons, which slightly modifies its properties.

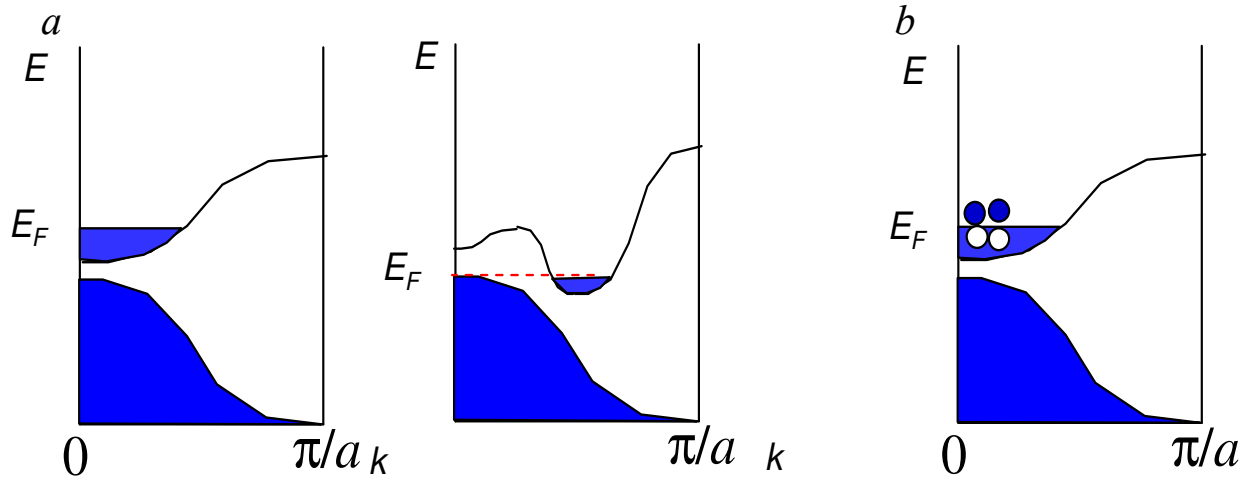


Figure 1.2: [Adapted from Ref. 4] a) Band structure of a metal at  $T=0$ . Filled states are in blue. The Fermi energy  $E_F$  separates the filled states from the empty ones. Conventional metals can arise from either partially filled bands (left) or overlapping bands (right). b) At finite temperature, electrons near the Fermi level can be excited with little energy cost, leading to metallic conduction.

Under the standard Bloch wave picture, the energy in the conduction or valence band depends on  $\mathbf{k}$ . It is often valid to expand this dependence in terms of powers of  $\mathbf{k}$ , and when the primary dependence is quadratic, the coefficient of the  $k^2$  term defines the effective mass

$$E(\mathbf{k}) = Ak^2 \equiv \frac{\hbar^2}{2m^*} k^2. \quad (1.19)$$

In general, the effective mass  $m^*$  can differ strongly from that of a free electron. In a noninteracting picture,  $m^*$  is determined by the details of the lattice, and there are also Fermi liquid corrections to the effective mass. In general, but not always,  $m_e^* > m_e$  due to the effective dragging along of the additional interactions. At the Fermi energy, we can define the Fermi velocity as the slope of the energy as a function of  $k$  at  $E_F$ ,  $v_F = \frac{\partial E_F}{\partial k} = \frac{\hbar k_F}{m^*}$ . We can also define

the Fermi momentum,  $p_F = m^*v_F = \hbar k_F$ . The Fermi velocity and momentum reflect the momentum and group velocity of a fermion at the Fermi surface. In other words, the conduction electrons participating in electrical transport are moving with velocity  $v_F$ , such that the Fermi velocity plays the role of  $v_d$  above.

### 1.2.3 The Fermi-Dirac Function

When considering how the band structure fills up with electrons at finite temperature at equilibrium, it is necessary to use a probabilistic approach. For fermions (spin- $\frac{1}{2}$  particles) we use a probability density function called the Fermi-Dirac (FD) distribution to describe the probability that a particular energy state  $E$  is filled when the system is in equilibrium. As a general approximation, all states above the Fermi energy have a low probability of being filled, and all states below the Fermi energy have a high probability of being filled. By the thermal energy scale  $k_B T$ , the single-particle states much lower than  $k_B T$  from  $E_F$ , must remain fully occupied since there is insufficient thermal energy to excite them above the Fermi level. Therefore, without any external forces, most of the quasiparticle excitations occupy single-particle states within  $k_B T$  of the Fermi energy.

The Fermi-Dirac distribution, the probability that a given single-particle state with energy  $E$  and some spin is occupied, is given by

$$f(E, \mu, T) = \frac{1}{e^{\frac{(E-\mu)}{k_B T}} + 1}, \quad (1.20)$$

where  $\mu$  is the chemical potential. For  $T \ll \frac{E_F}{k_B}$ , it is reasonable to approximate  $\mu \rightarrow E_F$ .

*Figure 1.3a* is the Fermi-Dirac distribution as a function of temperature. In the limit of  $T = 0$ , it is a step function with the step at  $E = \mu$ , and the many-particle system is degenerate. The Fermi temperature of most metals is large, and therefore they may be considered degenerate at



room temperature. For example, copper has  $T_F = \frac{E_F}{k_B} = 82000 \text{ K}$ . This is equivalent to saying the Fermi sphere in k-space is sharply bounded due to the step function nature of the FD.

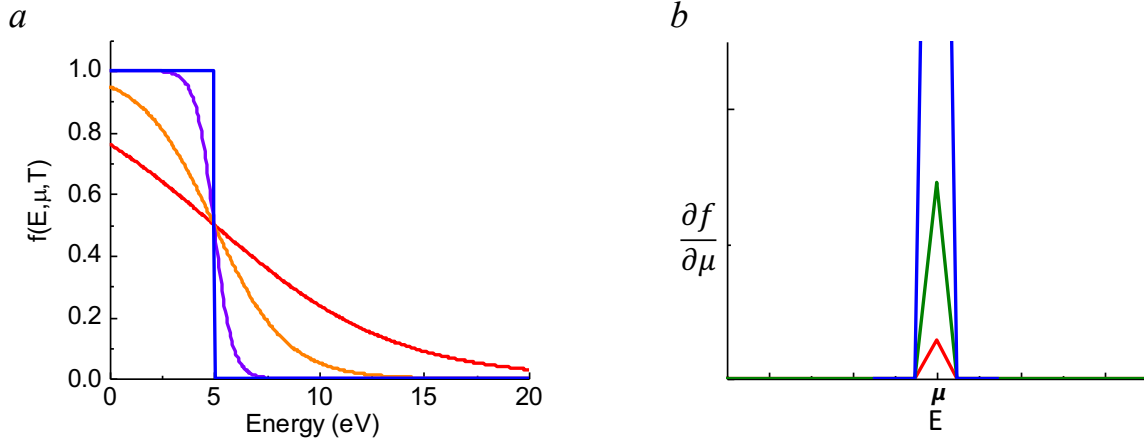


Figure 1.3: a) Fermi-Dirac distribution for  $\mu = 5 \text{ eV}$ . For  $E - \mu \gg k_B T$ , the distribution is essentially a step function. As temperature increases (coldest = blue  $\rightarrow$  hottest = red), the distribution smears out as the probability of finding electrons elevated above the Fermi energy increases. b) The thermal broadening function, a measure of how quickly the occupancy of the single-particle states change with energy. It is peaked around  $E = \mu$  and becomes a Dirac delta as  $T \rightarrow 0$ .

As  $T \rightarrow \infty$ , the FD distribution smears out, such that an increasingly significant fraction of all the electrons are excited above the Fermi energy. We can quantify the thermal smearing by the thermal broadening function

$$\left(\frac{\partial f}{\partial \mu}\right) = -\left(\frac{\partial f}{\partial E}\right) = \frac{1}{4k_B T} \operatorname{sech}^2\left(\frac{E-\mu}{2k_B T}\right). \quad (1.21)$$

As seen in Figure 1.3b, only the occupancy of single-particle states within  $\sim k_B T$  are affected by the thermal smearing. The thermal broadening function is a measure of how rapidly the

occupancy of the single-particle states changes with energy. It is peaked around  $E = \mu$ , and becomes a Dirac delta function in the degenerate limit,  $T = 0$ .

In the framework of these first two sections, conduction is a diffusive process, limited only by the population of scatterers, but how does Ohm's law need to be adjusted to accurately describe transport on significantly smaller scales?

### ***1.3 Length Scales and Transport: From Diffusive to Ballistic***

A diffusive system is one in which the above description of conductivity and electron mobility holds; electrons undergo many scattering events as they transverse a material. As size is reduced to the mesoscopic limit, roughly on the order of micrometers down to the nanoscale, the physics of electronic transport becomes more complicated.<sup>8</sup> Looking at the important length scales in a polycrystalline metal for example, consider a long wire of a metal such as gold, with length  $L$  and diameter  $d$ , with  $L \gg d$ . In the scope of the Drude model, without any inelastic processes, the typical distance an electron will travel within the wire before undergoing a scattering event is the elastic mean free path,  $l_e$ . These elastic scattering events are caused by any disruption of the periodicity in the stacking of the atoms, such as grain boundaries, vacancies in the lattice, or the surface of the metal. In the absence of any inelastic scattering, the grain size of the metal, roughly 20-40 nm for gold, is a good estimate for the elastic mean free path.

Leaving behind the strictly classical approach, it is also important to consider the wave-like nature of the conduction electrons. The effective wavelength of the electrons involved in conduction is the Fermi wavelength,  $\lambda_F$ . The Fermi wavelength determines the dimensionality of the electronic structure. For bulk gold,  $\lambda_F \approx 5 \text{ \AA}$ . Like all waves, electronic waves carry some phase and interact with each other constructively when in phase or destructively when out of

phase. Inelastic scattering events randomize the phase of the electronic waves on a length scale known as the coherence length,  $L_\phi$ . For bulk gold at room temperature,  $L_\phi \sim 1-2$  nm.

Consider an electron traveling along the wire for a very short length  $l < L_\phi$ , so that it encounters no scattering. This is referred to as ballistic conduction. In the classical model, where resistance arises from scattering events, one might naively think the resistance would drop to zero. Experimentally, however, it has been repeatedly shown that the two-terminal resistance never reaches zero, even in a point contact between two metals. As no scattering occurs within the ballistic channel, resistance must arise at the interface of the channel and the electrodes. Contact resistance was measured directly in metals in 1969 by Y.V. Sharvin and N.I. Bogatina.<sup>9</sup> To understand how the voltage is dropped in the case of a ballistic channel between two contacts, a quantum mechanical approach is necessary.

#### ***1.4 The Landauer-Büttiker Model***

To build the quantum mechanical picture of electronic transport in a ballistic conductor, imagine a ballistic one-dimensional channel (*Figure 1.4a*) with length  $L$  between left and right contacts in the  $x$  direction, similar to the model formalized by S. Datta<sup>8</sup>. Electrons are free to move along  $L$  and can occupy a nearly continuous band of states. The narrow conductor is essentially a waveguide, allowing a discrete series of 1d subbands, spaced apart in energy, corresponding to different transverse modes that carry current through the channel. *Figure 1.4b* shows the dispersion relations, the energy of the electron waves as a function of wavenumber, of the subbands involved in conduction in the channel. If the spacing between bands is large compared to  $k_B T$ , where  $k_B$  is the Boltzmann constant and  $T$  is the environmental temperature,

then a well-defined number of bands will be occupied. We can define a function  $M(E)$  to count the number of occupied modes with energies less than  $E$ .

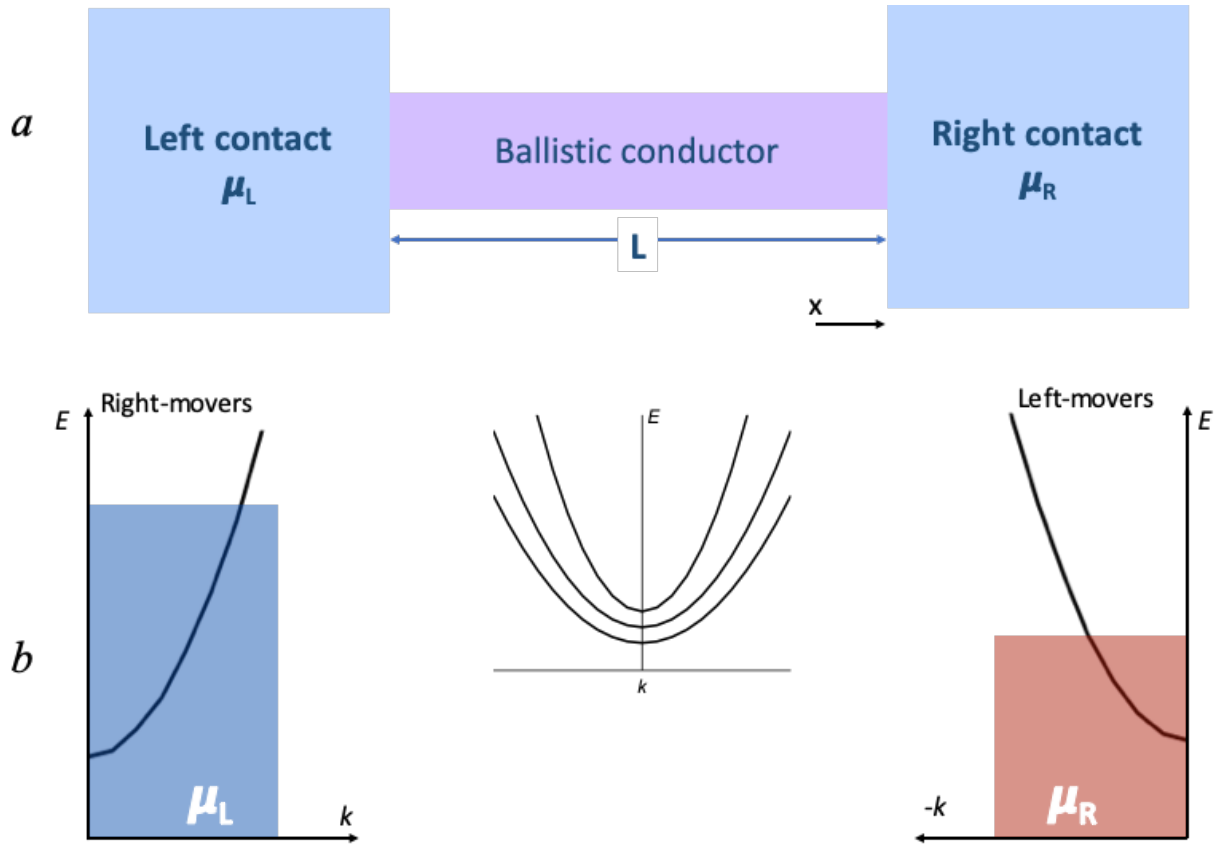


Figure 1.4: a) One-dimensional ballistic conductor, capped by two contacts at chemical potentials  $\mu_L$  and  $\mu_R$ . b) Dispersion curves of right-moving and left-moving electrons.

The left and right contacts have chemical potentials  $\mu_L$  and  $\mu_R$ , respectively. With perfect transmission from contact to channel and no scattering, right-moving and left-moving electrons do not interact and move across the channel with the energy with which they were injected. Thus, all right-moving electrons ( $k > 0$ ) originated at the left contact, and all left-moving electrons ( $k < 0$ ) originated at the right contact. If a voltage difference, in other words a

difference in chemical potentials, is established between the two contacts, the net current is the difference in flux between right-moving and left-moving electrons. The occupation probability for the right-moving carriers is given by the Fermi function for the energy, temperature, and chemical potential of the left contact  $f_+(E, T, \mu_L)$ . For electrons as charge carriers, the right-moving current is the flux per unit time of right movers, summed over all occupied states:

$$I_+ = \frac{-e}{L} \sum_k v(E(k)) f_+(E, T, \mu_L) M(E) \quad (1.22)$$

where  $v(E(k)) = \frac{1}{\hbar} \frac{\partial E}{\partial k}$  is the semiclassical velocity of the electrons, derived from the slope of the dispersion curve. In 1d  $k$ -space, wavefunctions are spaced by  $\frac{2\pi}{L}$ . So for each  $k$ , there are  $\frac{L}{2\pi}$  states to occupy, and the density of states is  $\frac{L}{2\pi} \frac{dk}{dE}$ . We can convert to an integral over energy using the relation  $\frac{\partial E}{\partial k} dk \rightarrow dE$ . The current due to right-moving electrons is then

$$I_+ = \frac{-2e}{h} \int_0^\infty f_+(E, T, \mu_L) M(E) dE \quad (1.23)$$

where the factor of 2 arises from the spin- $\frac{1}{2}$  degeneracy of the electrons. Repeating for carriers that originated from the right contact, the left-moving current is

$$I_- = \frac{-2e}{h} \int_0^\infty f_-(E, T, \mu_R) M(E) dE. \quad (1.24)$$

The total current  $I$  is the difference between the two,  $I = I_+ - I_-$ . At zero temperature,  $f_+$  and  $f_-$  are simple step functions at  $\mu_L$  and  $\mu_R$ , respectively, so that

$$I = \frac{-2e}{h} (\mu_L - \mu_R) M = \frac{2e^2}{h} \left( \frac{\mu_L - \mu_R}{-e} \right) M = \frac{2e^2}{h} V M \quad (1.25)$$

where  $V = \frac{\mu_L - \mu_R}{-e}$  is the voltage difference between the left and right contacts. When the chemical potential of the left contact is higher than that of the right, more charge carriers are

moving right than left, and a net positive current results. The two-terminal conductance across  $M$  ballistic channels, known as the Landauer formula,<sup>10,11</sup> is given by

$$G = \frac{2e^2}{h} M. \quad (1.26)$$

The major differences between the Landauer formula for conductance and Ohm's law are firstly that there is some inherent resistance that is independent of the sample length, and secondly that the conductance does not decrease linearly with the sample width. Rather, conductance depends on the number of transverse channels involved in transport. The remarkable conclusion of the Landauer formula is that if we can tune the number of channels through which charge carriers can traverse, we should see the two terminal conductance change in discrete steps of the conductance quantum,  $G_0 = \frac{2e^2}{h} \approx \frac{1}{12906} \Omega$ .

Generalizing to the case of multiple channels, the role of scattering in the conductance of atomic-scale junctions can be thought of in terms of transmission rates. The total conduction is given by the full Landauer formula

$$G = \frac{2e^2}{h} \sum_i \tau_i \quad (1.27)$$

with  $\tau_i$  denoting the transmission probability of each channel involved in the conduction. Büttiker later generalized Landauer's model to include the case of multi-lead geometries and the influence of magnetic field.<sup>12</sup> The full model is widely known as Landauer-Büttiker formalism. The ideal case works under the following assumptions: 1) coherent, ballistic transport, with no inelastic scattering events; 2) length scales less than the Fermi wavelength; and 3) zero temperature, so that there is no smearing of the Fermi-Dirac distributions.

Since all modes involved are ballistic, the finite resistance must arise at the interfaces between the channel and the contacts. Though there are only a few transverse modes through the

channel, there are many more available modes in each contact. Only a small number of these are able to couple to the modes that propagate through the ballistic channel, however, so all wavepackets that couple poorly are reflected back. In terms of the chemical potentials of each piece of the system, we recall that right-moving carriers remain at the chemical potential of the left contact,  $\mu_L$ , until they reach the right contact, and left-moving carriers remain at the chemical potential of the right contact,  $\mu_R$ , until they reach the left contact. The ballistic channel is characterized by the average of  $\mu_L$  and  $\mu_R$ , meaning the applied voltage must be dropped at the contacts, where the change in chemical potential occurs. In real systems, electron-electron interactions smear the ideally sharp cutoff between  $\mu_{avg}$  and the contact chemical potentials.

### ***1.5 Experimental evidence of conduction quantization***

Conductance quantization was first demonstrated experimentally in two-dimensional electron gases (2DEGs) at GaAs/AlGaAs interfaces in the 1980s.<sup>13,14</sup> *Figure 1.5*, adapted from the work by van Wees *et al.*, depicts clear steps in the two-terminal conductance of a 2d electron gas in a GaAs/AlGaAs heterostructure as a gate voltage is applied to pinch off the constriction. These measurements were performed at very low temperature, 0.6 K, but at higher temperatures, the Fermi-Dirac distributions of the electrons broaden, which leads to a smearing of the steps.

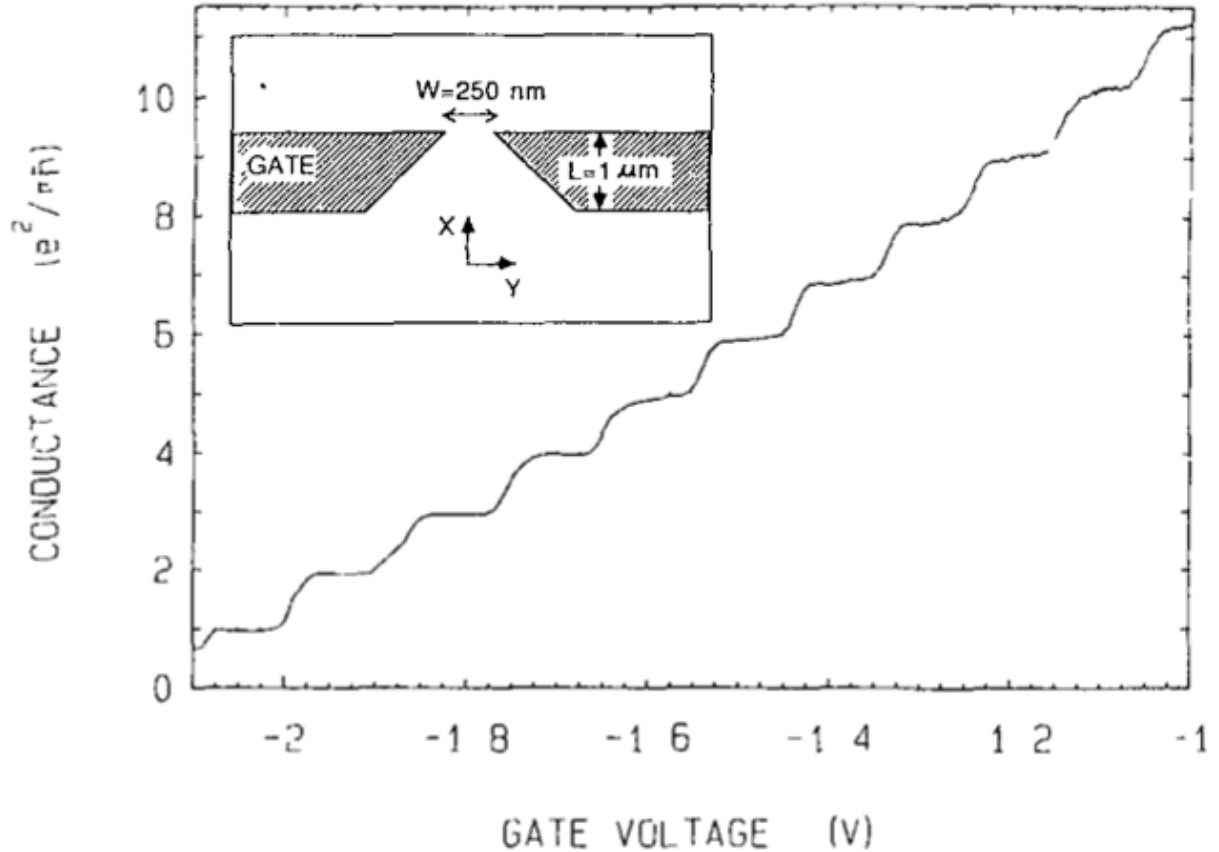
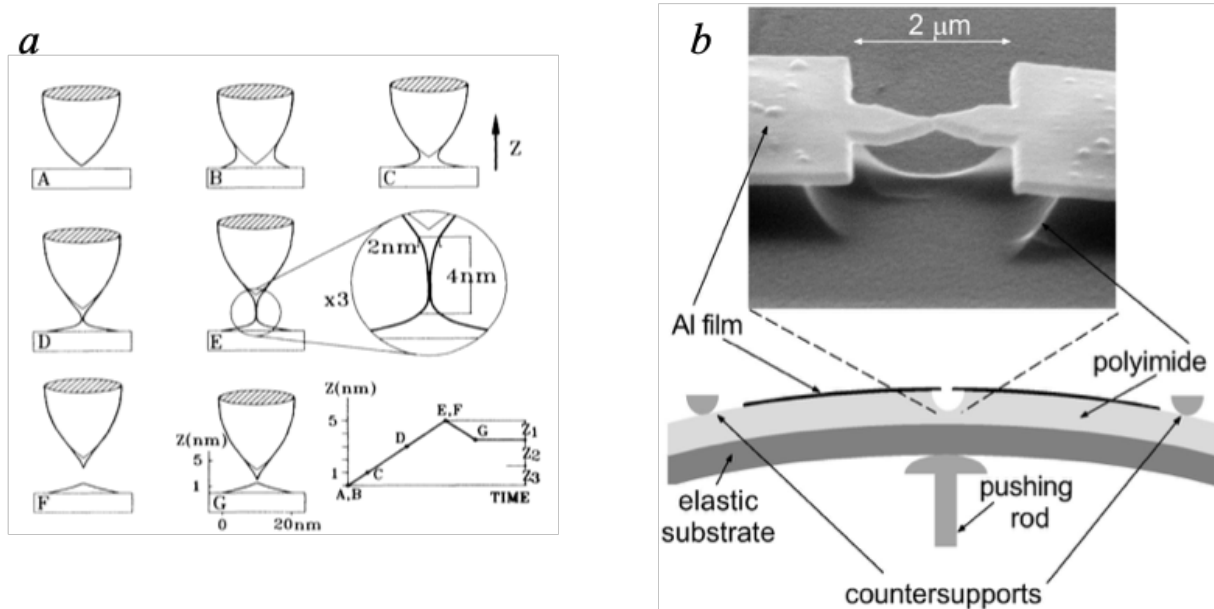


Figure 1.5: [Adapted from Ref. 13] Two-terminal conductance of GaAs/AlGaAs 2DEG as the width is tuned by a top gate.

Steps in conductance have also been observed in atomic-scale metal junctions, even up to room temperature since  $k_B(T = 300 \text{ K}) \ll E_F$  for most metals. Mechanical approaches are used to observe this phenomenon in metals, either by scanning tunneling microscopy (STM)<sup>15–18</sup> or mechanical break junctions (MBJ).<sup>19–25</sup> In both cases, junctions of just a few or a single atom are created by deforming a larger contact. In STM-style experiments, a metal tip is brought into and out of contact with a metal film relatively slowly while measurements of current and conductance of the junction capture the transition from macroscopic to atomic-scale connection. In MBJ setups, a wider metal junction is stretched to elongate the junction until it is only a few



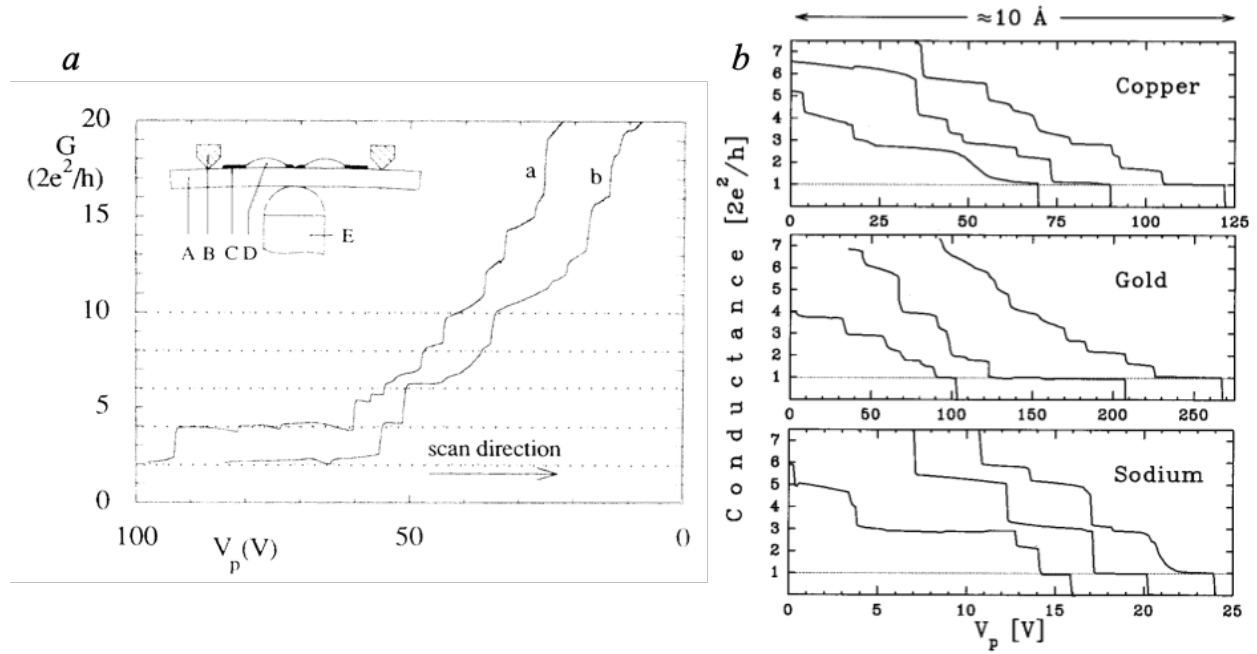
atoms wide. Once the contact has broken the process is reversed to form a new junction. In both measurement types, many contact breaking-reforming cycles are performed to collect enough data for statistical analysis, particularly conductance histograms. *Figure 1.6a* and *b*, adapted from References 11 and 17 respectively, depict the processes for STM and MBJ measurements.



*Figure 1.6: a) [Adapted from Ref.16] An example of an STM-style measurement. In this case, a mound of gold is deposited onto a gold substrate by the tip by applying a large voltage pulse while the tip is within tunneling distance to the substrate, then slowly pulling the tip away. b) [Adapted from Ref. 22] An example of a MBJ setup. A three-point bending mechanism is used to stretch the aluminum film that sits on top of a phosphorbronze substrate. The micrograph shows a suspended aluminum microbridge.*

Peaks in the conductance histogram and steps in conductance versus time or distance denote the most geometrically stable atomic configurations as the junction breaks. Though conductance traces sometimes show steps at non-integer values of the conductance quantum,

conductance histograms, as ensemble-averaged data, give a clearer picture of conductance quantization behavior, typically displaying peaks at integer values of  $G_0$ . *Figure 1.7* shows examples of conductance steps seen during experiments of MBJs in platinum,<sup>19</sup> copper, gold, and sodium.<sup>26</sup> *Figure 1.8 a* and *b* are conductance histograms for gold,<sup>25</sup> sodium, and copper.<sup>20</sup>



*Figure 1.7: Examples of conductance steps seen as the width of metallic atomic-scale junctions are varied. X-axis is the piezo voltage (piezo-controller is seen as E in inset of a), controlling the strain placed on the junction. a) [Adapted from Ref. 19] Platinum in vacuum at 1.2K. b) [Adapted from Ref. 26] copper, gold, and sodium at helium temperatures.*

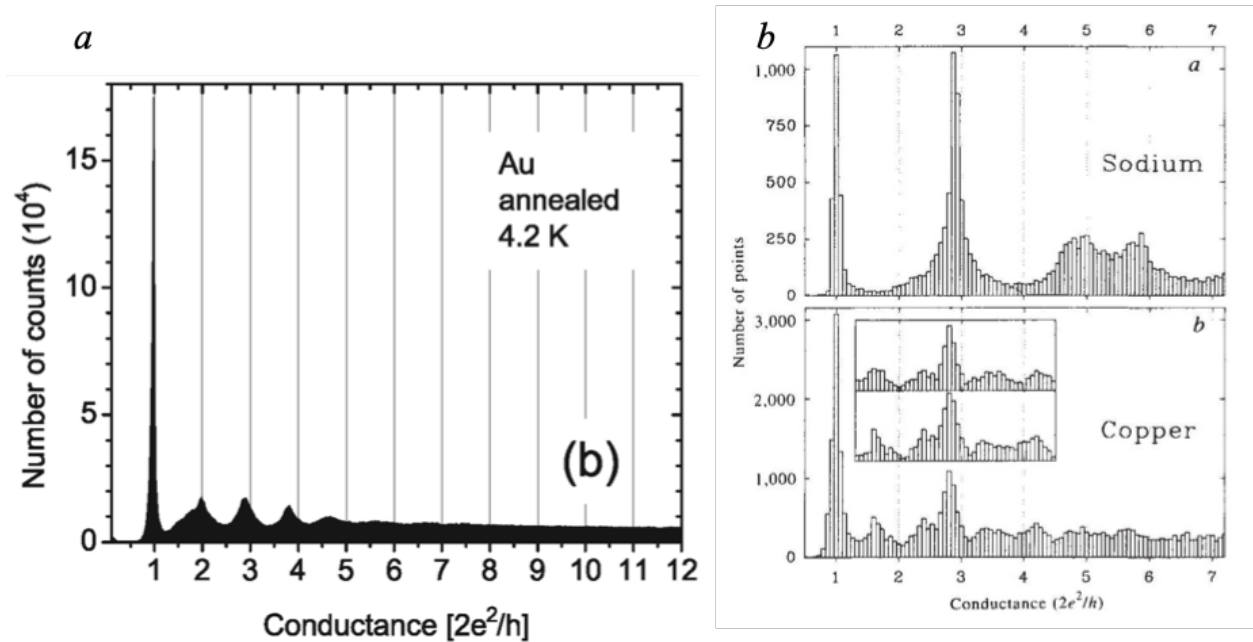


Figure 1.8: Examples of conductance histograms collected for gold, sodium, and copper.

a)[Adapted from Ref. 25] Conductance histogram for annealed gold at 4.2K collected using the MBJ technique. b)[Adapted from Ref. 20] Conductance histogram for sodium and copper at 4.2K also collected using the MBJ technique.

These two types of plots alone are not definitive proof of conductance quantization, however, only the presence of preferred stable geometries. Indeed, atomic force microscopy (AFM) measurements have shown that abrupt changes in the conductance accompany shifts in the atomic arrangement.<sup>27</sup> An important study by E. Scheer *et al.* showed that the number of quantum channels involved in transport is directly related to the number of valence electrons of the atoms.<sup>22,23</sup> Gold, for example, has electron configuration  $[\text{Xe}]4f^{14}5d^{10}6s^1$ , so a single atom chain would contain only one quantum channel. It was shown by Yanson *et al.* that Au single-atom chains display a single nearly fully transparent channel for any chain length.<sup>18</sup> For other metals, however, such as aluminum, conductance histograms denote only stable atomic

configurations. Conductance histograms of Al show peaks near 1, 2, and  $3G_0$ , despite there being three valence electrons and thus multiple channels contributing to conductance.<sup>28</sup> It is therefore necessary to be wary of assuming the number of channels involved in atomic-scale transport in metals from conductance histograms alone.

## ***1.6 Back to Ohm's Law***

Given the experimental evidence supporting the Landauer-Büttiker model, how do we reconcile the quantum transport picture of conductance with the macroscopic Ohm's law? To begin, we need to determine the transmission probability of a conductor of length  $L$ .<sup>8</sup> Let the conductor be a 1d channel with two barriers, with transmission probabilities  $T_1$  and  $T_2$ , such as a constriction in an otherwise 2d electron gas. To find the transmission probability to pass through both barriers,  $T_{12}$ , we need to consider all possible paths for an electron, including both the direct route and paths with any number of reflections. For simplicity, only consider the classical properties of the electrons and ignore phase differences between paths. Let  $R_1$  and  $R_2$  be the probabilities of reflection off the first and second barriers, respectively, with  $R_{1(2)} = 1 - T_{1(2)}$ . In the case of a direct transmission through the channel,  $T_{12} = T_1T_2$ . The simplest case of a path with reflection is transmission through the first barrier, reflection off the second, reflection off the first, then transmission through the second. This two-bounce scenario has probability  $T_1R_2R_1T_2$ . *Figure 1.9* depicts the first two possible routes and their respective probabilities for an electron traversing a conductor with two barriers.

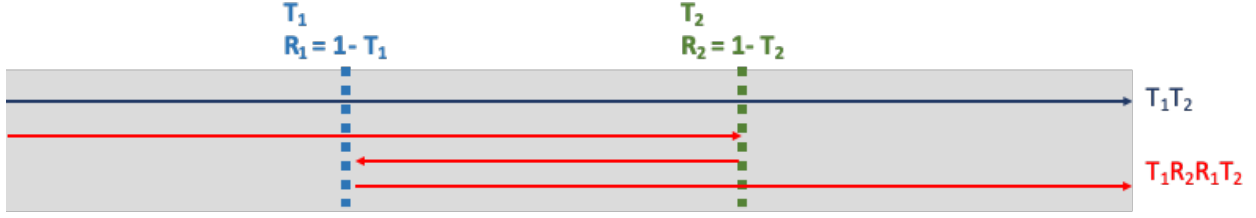


Figure 1.9: Possible paths for an electron travelling through a conductor with two barriers with transmission probabilities  $T_1$  and  $T_2$ .

Summing over the infinite series all possible bounce possibilities,

$$T_{12} = T_1 T_2 + T_1 T_2 R_1 R_2 + T_1 T_2 R_1^2 R_2^2 + \dots = T_1 T_2 \frac{1}{1 - R_1 R_2} \quad (1.28)$$

In terms of transmission probabilities only we arrive at the expression

$$\frac{1 - T_{12}}{T_{12}} = \frac{1 - T_1}{T_1} + \frac{1 - T_2}{T_2} \quad (1.29)$$

Generalizing to  $N$  scatterers in the conductor, each with transmission probability  $T$

$$\frac{1 - T_N}{T_N} = N \frac{1 - T}{T} \quad (1.30)$$

so

$$T_N = \frac{T}{N(1 - T) + T} \quad (1.31)$$

If  $\nu = \frac{N}{L}$  is the linear density of scatterers in the conductor, then

$$T(L) = \frac{L_0}{L + L_0} \quad (1.32)$$

where  $L_0 \equiv \frac{T}{\nu(1 - T)}$  is essentially the mean free path,  $L_m$ , the average distance an electron travels before scattering. Since the probability of an electron scattering off a single scatterer is  $1 - T$ , if we assume  $T \sim 1$ ,

$$(1 - T)\nu L_m \sim 1 \rightarrow L_m \sim \frac{1}{\nu(1 - T)} \sim L_0. \quad (1.33)$$

For a wide conductor with many modes, the number of modes (spaced by  $\frac{2\pi}{W}$ ) is proportional to the width  $W$ ,  $M \sim \frac{k_F W}{2\pi}$ . The conductance can then be written as

$$G = e^2 \frac{k_F W}{h\pi} T = e^2 \frac{k_F W}{h\pi} \frac{L_0}{L+L_0}. \quad (1.34)$$

Using the definition of the crystal momentum  $\hbar k$ , and the 2d density of states  $\nu_{2d} = \frac{m}{\pi\hbar^2}$ , the conductance is

$$G = e^2 \nu_{2d} \frac{v_F L_0}{\pi} \frac{W}{L+L_0} \quad (1.35).$$

Recalling the Einstein relation for conductivity in terms of the diffusion constant  $\sigma = e^2 \nu_{2d}(E_F) D$ , where  $D$  is the diffusion constant, let  $D = \frac{v_F L_0}{\pi}$ , so that

$$G = \frac{\sigma W}{L+L_0} \text{ and } G^{-1} = \frac{L+L_0}{\sigma W} \quad (1.36)$$

which can be separated into a contribution from the contacts,  $G_c^{-1} = \frac{L_0}{\sigma W}$ , and from the main conductor,  $G_s^{-1} = \frac{L}{\sigma W}$ , which is recognizable as Ohm's law.

# Chapter 2

## Electronic Noise

While current, voltage, and conductance measurements are crucial to understanding the steady-state characteristics of a system, they can fail to give the full picture. Electronic noise, the stochastic fluctuations about the time-averaged value of the voltage, current, or resistance, can help fill in the gaps. For example, if we could have a time stamp of every charge carrier traversing a particular system, we would then have a complete picture of how charge transport works in that system at a given voltage. The current is just the first moment of the charge versus time distribution, the average charge per time. The current noise, the mean square fluctuations about that average, is the second moment. Very rarely it is possible to measure higher moments, though the theoretical ideal would be full counting statistics, namely either having the entire charge v. time statistics, or equivalently, all higher order moments of the current. Due to its dynamic nature and origin from different physical properties, noise is able to carry information that cannot be inferred from current or conductance curves. In general, electronic noise can arise from fluctuations in current, voltage, or resistance, but typically measurements focus on current  $\langle(\delta I)^2\rangle$  or voltage fluctuations  $\langle(\delta V)^2\rangle$  about the average value. Noise is most commonly referred to in terms of spectral density, the noise power distribution across a given frequency range, and is written in units of  $A^2/Hz$  for current noise and  $V^2/Hz$  for voltage noise.

Different types of electronic noise have distinctive origins and frequency responses. The main types of electronic noise relevant to this thesis are Johnson-Nyquist thermal noise<sup>29,30</sup> (*Section 2.1*),  $1/f$  or flicker noise<sup>31,32</sup> (*Section 2.2*), and shot noise<sup>33</sup> (*Section 2.3*).

## 2.1 Johnson-Nyquist Noise

Johnson-Nyquist (J-N) thermal noise was discovered by J.B. Johnson<sup>29</sup> and theoretically modelled by H. Nyquist<sup>30</sup>, both in 1928. Both found that the mean-square voltage fluctuations across a device due to thermal agitation of charge carriers are proportional to the resistance and the absolute temperature (*Figure 2.1*). Johnson-Nyquist noise is observed in all types of resistors and is considered an equilibrium noise; in other words, it is present even when the device is in equilibrium with no external potential difference.

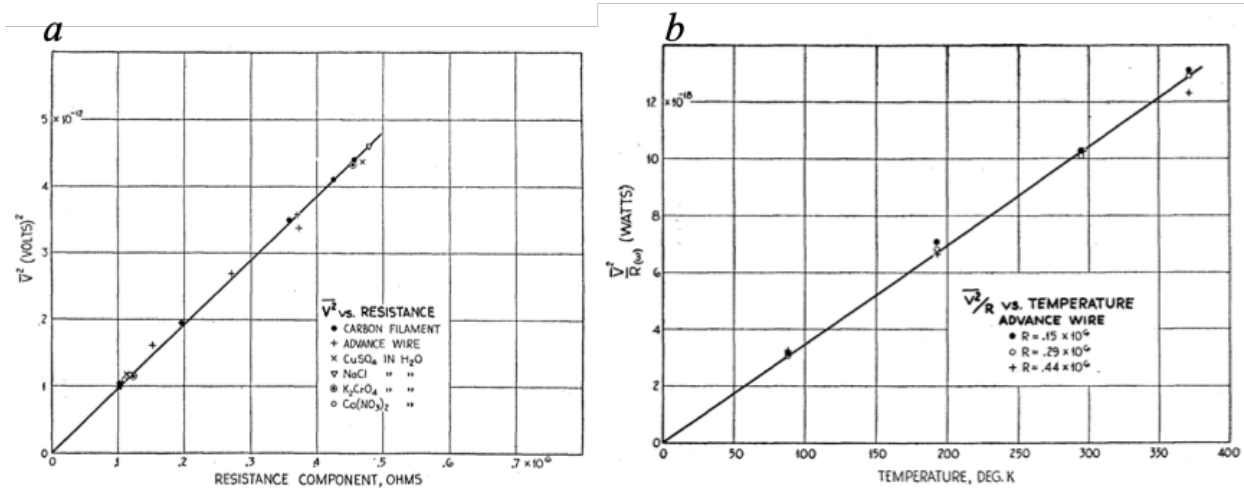


Figure 2.1: [Adapted from Ref. 29] a) Voltage-squared as a function of resistance for several types of conductors. b) Apparent power versus temperature. Johnson found a linear dependence of mean-square voltage fluctuations on both resistance and temperature.

Nyquist formalized J-N noise using a model of two conductors of equal resistance  $R$  and held at the same temperature  $T$  (*Figure 2.2a*).<sup>30</sup> Thermal fluctuations in the first resistor set up a current in the circuit, whose value is given by  $I = \frac{\varepsilon}{2R}$ , where  $\varepsilon$  is the electromotive force due to the fluctuations. The power dropped over the second conductor due to  $\varepsilon$  is  $P = I^2R$ . The second law of thermodynamics, that no system can spontaneously produce mechanical work from



thermal energy, implies that since the conductors are at the same temperature, power dropped over the first conductor due to fluctuations in the second is equal to the power dropped over the second.

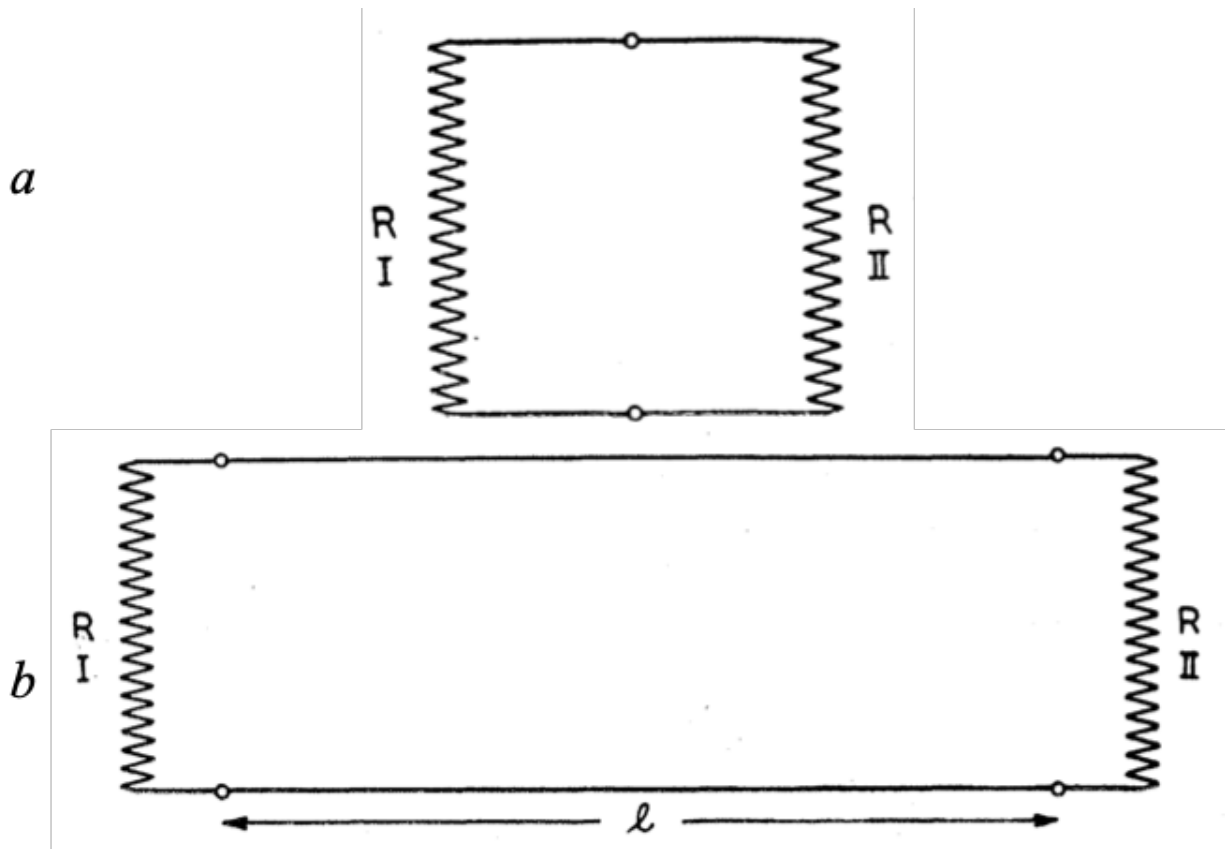


Figure 2.2: [Adapted from Ref. 30] Nyquist's system of two resistors used to theoretically derive Johnson-Nyquist noise. a) Two resistors of equal resistance  $R$  held at the same temperature. The power dropped over the second resistor due to voltage fluctuations in the first is  $\frac{\varepsilon^2}{4R}$ , where  $\varepsilon$  is the electromotive force due to thermal fluctuations in the first resistor. b) A model of the two resistors now separated by a long, non-dissipative transmission line can be used to arrive at the voltage noise spectral density  $\delta S_V = 4k_B T R$ .

Now consider the case where the two resistors are connected by a long, non-dissipative transmission line of length  $l$  (*Figure 2.2b*) with switches at each point of connection between line and conductor. The line has inductance  $L$  and capacitance  $C$  per unit of line such that  $\sqrt{\frac{L}{C}} = R$ , implying there is no reflection at the end of the line. Let the system be in thermal equilibrium, such that there is equal power flowing from the first conductor to the second and vice versa. Now electrically isolate the transmission line from the two conductors by opening all switches. Any energy that was in the transmission line at the moment the switches were opened is now trapped, forming standing waves within the line. The frequency of the natural modes are  $\frac{nv}{2l}$ , where  $n$  is an integer, with  $n \geq 2$ , and  $v$  is the velocity of propagation. Focusing on a frequency range of width  $df$ , the number of cavity standing modes, or degrees of freedom, within this range is  $\frac{2ldf}{v}$ , as long as  $l$  is long enough to ensure there are many modes. We can then ascribe the average energy per mode as  $k_B T$ , based on the equipartition theorem for a one-dimensional harmonic oscillator. Note, this assumes  $k_B T$  is large compared to the mode spacing; modes with energies large compared to  $k_B T$  are not occupied, as in the Planck blackbody spectrum.

So the total energy in the frequency range  $df$  is  $E = \frac{2ldf}{v} k_B T$ , and since there is no reflection, this is equal to the total energy transferred during the transit time  $t = l/v$  by both conductors. In terms of power, the average power delivered to the transmission line by each conductor within  $df$  is  $P = \frac{1}{2} \frac{E}{t} = k_B T df$ . If the total current is given by  $\frac{V}{2R}$ , the original expressions for current and power give  $P = I^2 R_{tot} = \frac{V^2}{4R}$ . If we then denote the voltage-squared within  $df$  as  $V^2 df$ , or divide through by  $df$  for the voltage noise spectral density,  $S_V$ , we arrive at the well-known expression for Johnson-Nyquist noise

$$\delta S_V = 4k_B T R, \quad (2.1)$$

or as current noise (converted by dividing through by  $R^2$ )

$$S_I = 4k_B T G \quad (2.2)$$

where  $G = \frac{1}{R}$  is the two-terminal conductance.

Johnson-Nyquist noise is a particular example of the fluctuation-dissipation theorem,<sup>34,35</sup> which relates the thermal mean-square fluctuations of some physical variable of a system to the dissipative response of the system to perturbations that effect that variable. Resistance dissipates electrical energy by turning it into heat, and the corresponding fluctuation is J-N thermal noise. The thermal fluctuations of electrons and atoms within a resistor on a loop of wire cause fluctuations in the current around the wire. J-N noise converts the thermal energy into electrical energy, which is the reverse of the resistance.

J-N noise obeys Equation 2.1 for frequencies  $f \ll \frac{k_B T}{\hbar}$ , but begins to roll off steeply to zero at higher frequencies. In most measurements, it is safe to assume J-N noise is white, as  $\frac{k_B}{\hbar} \sim \frac{10^{-23}}{10^{-34}} \sim 10^{11}$  or roughly 20 GHz for  $T \sim 1$  K. It is also worth noting that for a system with complex impedance, only the real component  $Re\{Z\}$  contributes to the thermal noise.<sup>36</sup> In other words, a pure capacitor or inductor does not produce thermal noise.

## 2.2 *1/f* Noise

*1/f* noise, also known as flicker noise, was first described by J.B. Johnson in a 1925 report on current fluctuations, then referred to as the Schottky effect, in the thermionic emission of tungsten filaments (*Figure 2.3a*).<sup>37</sup> Though in this first account, Johnson merely noted that the fluctuations were up to 100 times larger than expected at low frequencies, he later described the inverse dependence on frequency in a letter to Nature in 1927.<sup>38</sup> *1/f* noise is ubiquitous,

appearing in completely unrelated systems, from gold nanocontacts<sup>39</sup> to music and speech<sup>40</sup> (Figure 2.3a and b). The wide diversity of included structures is not fully understood and renders any attempts at a universal model impossible.

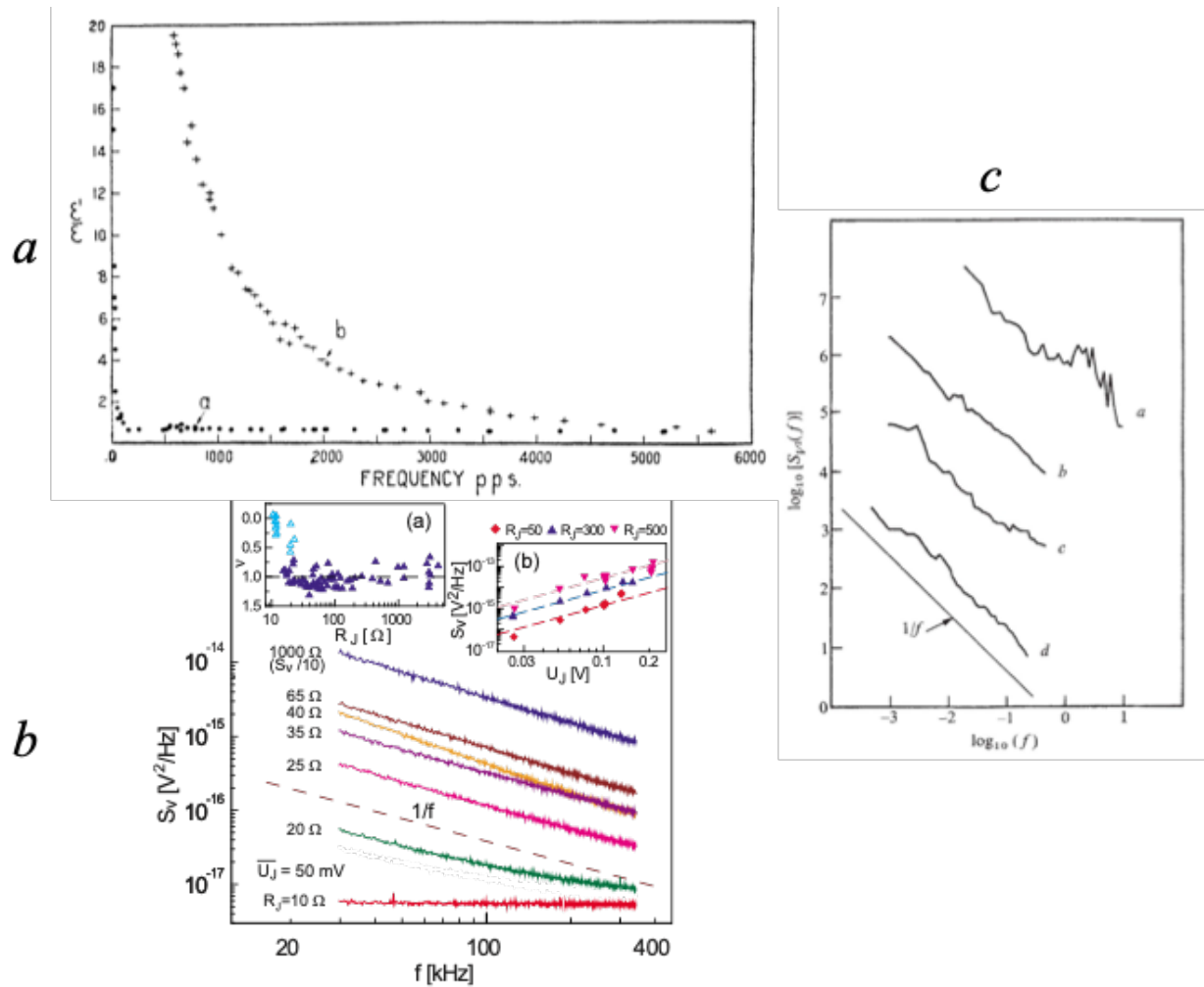


Figure 2.3: a) [Adapted from Ref. 37] J.B. Johnson found excess noise at low frequencies in the thermionic emission of tungsten filaments. b) [Adapted from Ref. 39] Voltage noise versus frequency in electromigrated gold junctions taken at a bias of 50mV. c) [Adapted from Ref. 40]  $1/f$  noise found in a. Scott Joplin piano rags, and b. classical, c. rock, and d. news and talk radio stations.

In electronic systems, at least,  $1/f$  noise is attributed to resistance fluctuations. Sources of resistance fluctuations include shifts in atomic motions or two-level fluctuations,<sup>31,32,41,42</sup> charge trapping,<sup>43</sup> mobility fluctuations,<sup>44,45</sup> and more. Following from the classical Ohm's law, resistance fluctuations lead to voltage fluctuations that are quadratic in bias or current, since  $\langle(\delta V)^2\rangle = \langle(\delta R)^2\rangle \cdot I^2$ . In nearly all resistors, resistance fluctuations exist even in the absence of a driving current,<sup>32</sup> but  $1/f$  noise is still considered a nonequilibrium noise because it is only revealed once a system is driven out of equilibrium by an applied bias or current.

Hooke's law, though it is generally agreed to have no universal origin, has been used for decades to describe  $1/f$  noise:

$$\frac{S_R(f)}{R^2} = \frac{\alpha}{Nf} = \frac{S_I(f)}{I^2} \quad (2.3)$$

where  $\alpha$ , known as the Hooge parameter, is a material-dependent constant, and  $N$  denotes the number of fluctuating particles. For bulk conductors,  $N$  is usually taken to be the total number of electrons. As  $N$  is dependent on volume, and thus resistance, spectral density tends to scale inversely with system size, which is consistent with local rather than long-range noise sources. In general, one can simply think of the spectral density  $S(f) \propto f^{-\beta}$ , with  $\beta$  ordinarily in the range of 0.8 - 1.4.

### **2.3 Shot Noise**

The remainder of Chapter 2 will be dedicated to shot noise, a nonequilibrium noise that is the basis for the studies in thesis. Shot noise arises due to the discrete nature of charge carriers; though there is some average current, there will be fluctuations in the actual arrival times of electrons at the drain. Walter Schottky first discovered the presence of shot noise in vacuum diodes in 1918<sup>46</sup> and defined the current noise spectral density in the classical limit:

$$S_I = 2e \langle I \rangle, \quad (2.4)$$

where  $e$  is the effective charge of the conductive carriers, and  $\langle I \rangle$  is the time-averaged current. As evidenced from the formula, shot noise has no frequency dependence (*i.e.* it is white noise), and it is only generated when a system is driven out of equilibrium by an applied bias or current. This definition of the current noise spectral density holds for systems in which electronic transport is characterized by uncorrelated, non-interacting tunneling processes. Namely, when the arrival events of electrons at the drain can be described by Poissonian statistics.

## ***2.4 Poissonian Statistics***

A Poisson probability distribution (*Figure 2.4*) describes a set of events with some mean rate, but the actual timing of an individual event is uncorrelated to the timing of any other event in the set. Therefore, the actual timing between successive events will fluctuate. Examples include, the number of goals scored during a game between two teams, the number of men killed by a horse kick each year in the Prussian army,<sup>47</sup> and the number of yeast cells used in brewing Guinness.<sup>48</sup>

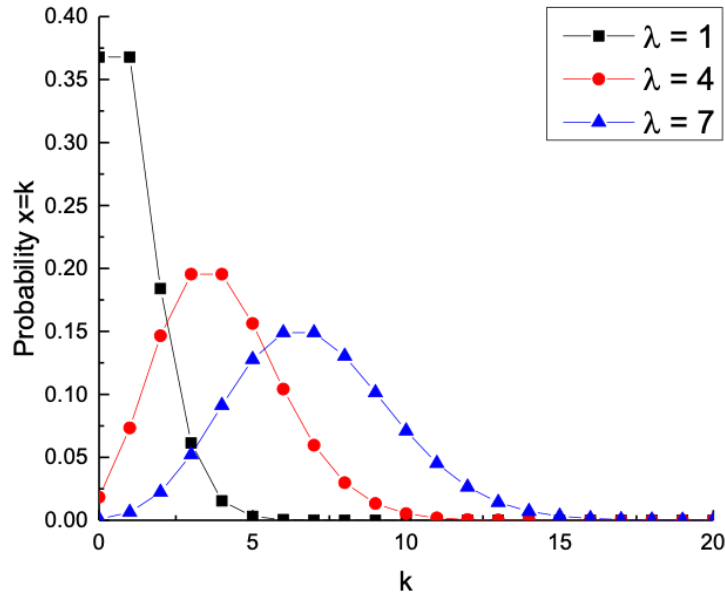


Figure 2.4: Examples of Poissonian distributions,  $P(x = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ , for three different average number of events.

For a given time interval, an event can occur any number of times, or not at all. The probability that there are  $k$  events during a certain interval is given by

$$P(x = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad (2.5)$$

where  $\lambda$  is the average number of events per interval. In terms of an average rate  $r$  instead of average number of events, this becomes

$$P(k \text{ events in time } t) = e^{-rt} \frac{(rt)^k}{k!}. \quad (2.6)$$

Return to Schottky's example of electron discharge into a vacuum to determine the mean square deviation for a given time interval  $\tau$ .<sup>46</sup> If  $I_0$  is the time-averaged current over all time intervals and  $I_\tau$  is the average current for interval  $\tau$ , then the deviation of  $I_\tau$  from  $I_0$  is given by

$$\Delta I = I_\tau - I_0. \quad (2.7)$$

The ultimate goal is to derive the mean square deviation  $\overline{\Delta I^2}$ . Let  $N\tau$  be the number of electrons discharged within interval  $\tau$  on average. In actuality, the number of electrons discharged will vary, and for a certain interval of time  $\tau$ , let the real number be  $n_\tau$ . The mean deviation in number of electrons discharged is then

$$\Delta N = n_\tau - N\tau. \quad (2.8)$$

For a Poissonian distribution, the variance of the set is equal to the mean, so that

$$\overline{\Delta N^2} = N\tau, \quad (2.9)$$

If  $I_0 = \frac{eN\tau}{\tau} = eN$  and  $I_\tau = \frac{en_\tau}{\tau}$ , where  $e$  is the fundamental charge of an electron, then

$$\Delta I = \frac{e\Delta N}{\tau}, \quad (2.10)$$

and

$$\overline{\Delta I^2} = \frac{e^2}{\tau^2} \overline{\Delta N^2} = \frac{e^2}{\tau^2} N\tau = \frac{e^2 N}{\tau} \quad (2.11)$$

or in terms of  $I_0$ ,

$$\overline{\Delta I^2} = \frac{eI_0}{\tau}. \quad (2.12)$$

From this we can see the mean-square deviation of the current is directly proportional to the average current multiplied by the charge of the carriers. The traditional expression for the current spectral density,  $S_I = 2eI$ , can then be found via a Fourier transformation to the frequency domain. A more detailed derivation can be found in Reference 49.<sup>49</sup>

## 2.5 The Fano Factor

Many processes and interactions can cause shot noise to be suppressed below or enhanced beyond the  $2eI$  limit (some of which will be discussed in upcoming sections). The Fano factor, named for Ugo Fano and his statistical analysis of ionization yield from radiation,<sup>50</sup>



is a descriptor of how measured noise compares to the perfectly non-interacting, Poissonian case. The current spectral density is then written as  $S_I = F \cdot 2eI$ , where  $e$  can also be taken to be  $e^*$ , the effective charge of the carriers, which may be greater than or less than  $1e$ . As the effective charge and the average current are contained in the  $2eI$  term, the Fano factor then carries different, independent information about the system. For example, in the case of a ballistic, or quasi-ballistic few-channel junction, the Fano factor can provide the single-particle transmittances of the channels.<sup>51–53</sup> The Fano factor can also reveal electron-electron or electron-phonon interactions.<sup>54,55</sup>

## ***2.6 Shot Noise of a Non-Interacting Single Barrier***

The simplest case of a system that produces shot noise is transport through a single barrier with no interactions. Though basic, this example is very useful for elucidating the partition nature of shot noise and the results are still applicable in many experimental systems, such as metallic atomic-scale junctions or 2DEGs. Consider a single barrier placed between two metallic leads, assuming for now only one transmission channel. The “barrier” in this case is simply a means of deriving the transmission probability for an electron passing through the system of interest. The actual microscopic structure is unimportant for the derivation.

Begin with the situation in which a one-dimensional beam of particles is incident on the barrier. At zero temperature, there is no smearing of the Fermi-Dirac distributions, and the only source of noise is shot noise arising from the fluctuations in number of particles being transmitted or reflected by the barrier. Framing the transport using the Landauer-Büttiker model, the current noise spectral density should be of the form

$$S_I = 2eVG_0 \langle \Delta n_T^2 \rangle \quad (2.13)$$

where  $\langle \Delta n_T^2 \rangle$  is the mean-square fluctuation of the transmitted intensity of the beam of particles. To determine  $\langle \Delta n_T^2 \rangle$  in terms of the transmission probability of the barrier, let  $T$  be the probability an incident particle is transmitted and  $R$  be the probability the particle is reflected. If the beam is incident on the barrier with an intensity  $\langle n \rangle = 1$ , then the transmitted and reflected intensities are  $\langle n_T \rangle = T$  and  $\langle n_R \rangle = R$ . The mean square fluctuations of the transmitted and reflected particles are  $\langle (\Delta n_T)^2 \rangle = \langle (n_T - \langle n_T \rangle)^2 \rangle$  and  $\langle (\Delta n_R)^2 \rangle = \langle (n_R - \langle n_R \rangle)^2 \rangle$ , respectively. Since transmission and reflection are inversely related, a positive fluctuation for transmission is a negative fluctuation for reflection,  $\Delta n_T = -\Delta n_R$ . So then

$$\begin{aligned}
\langle (\Delta n_T)^2 \rangle &= \langle (\Delta n_R)^2 \rangle = -\langle \Delta n_T \Delta n_R \rangle \quad (2.14) \\
&= -\langle (n_T - \langle n_T \rangle) \cdot (n_R - \langle n_R \rangle) \rangle \\
&= -\langle n_T n_R - n_T \langle n_R \rangle - n_R \langle n_T \rangle + \langle n_T \rangle \langle n_R \rangle \rangle \\
&= -(n_T n_R - 2\langle n_T \rangle \langle n_R \rangle + \langle n_T \rangle \langle n_R \rangle).
\end{aligned}$$

$n_T$  and  $n_R$  are orthogonal, so  $n_T n_R = 0$ , leaving

$$\langle (\Delta n_T)^2 \rangle = \langle n_T \rangle \langle n_R \rangle = TR = T(1 - T). \quad (2.15)$$

Equation 2.13 then becomes

$$S_I = 2eVG_0 T(1 - T), \quad (2.16)$$

or to generalize to the case of multiple channels with transmission coefficients  $\tau_i$ , the final form is

$$S_I = 2eVG_0 \sum_i \tau_i (1 - \tau_i). \quad (2.17)$$

It follows then that the Fano factor, the Landauer noise divided by the Poissonian noise is given by

$$F = \frac{2eVG_0 \sum_i \tau_i (1 - \tau_i)}{2eI} = \frac{2eVG_0 \sum_i \tau_i (1 - \tau_i)}{2eVG_0 \sum_i \tau_i} = \frac{\sum_i \tau_i (1 - \tau_i)}{\sum_i \tau_i}. \quad (2.18)$$

Consider briefly how this expands to macroscopic tunnel junctions. Macroscopic tunnel junctions are essentially two leads with many transmission channels, separated by a single, reasonably opaque barrier, so that the tunneling conductance divided by the number of channels is small compared to  $\frac{2e^2}{h}$ . This is equivalent to having many single-channel junctions, all with  $\tau_i \ll 1$ , in parallel. This results in a Fano factor of 1 in the limit of many poorly transmitting channels.

From the Equations 2.17 and 2.18 definitions of the Fano factor and shot noise, one can see that shot noise is suppressed to zero for fully transparent ( $\tau_i = 1$ ) or fully reflective ( $\tau_i = 0$ ) channels due to the  $\tau_i(1-\tau_i)$  term. In the case of a single channel,  $F=1-\tau$ , and  $S_I$  approaches the Poissonian limit as  $\tau \rightarrow 0$ . Most experimental cases are rarely perfectly single-channel or completely transparent/reflective, so typically shot noise is only mostly or partially suppressed. Noise suppression due to high channel transmission has been confirmed experimentally both in 2DEGs<sup>56,57</sup> and atomic-scale metallic<sup>51,58,59</sup> and molecular junctions<sup>52</sup> (Figure 2.5).

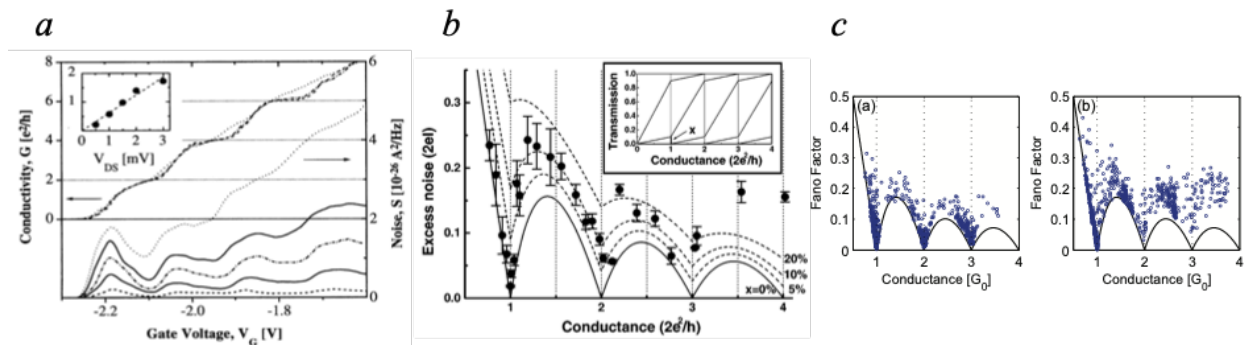


Figure 2.5: a) [Adapted from Ref. 57] Conductivity and current noise as gate voltage adjusts the width of a 2DEG. Positions of shot noise suppression correspond to steps in the conductance. b) [Adapted from Ref. 51] Excess noise in 27 gold contacts. Solid line is the case of a single partially transmitted mode. c) [Adapted from Ref. 53] Fano factor of ensembles of silver (left) and gold (right) atomic contacts. In both, noise suppression is seen near integer values of  $G_0$ .

In atomic-scale gold junctions, noise is largely suppressed when the conductance is near  $1G_0$ , confirming conductance quantization, not simply stable atomic configurations that coincidentally give  $G \sim \frac{2e^2}{h}$ . However, in other metals with multiple channels involved in conduction, such as aluminum which has three quantum channels for a single atom contact, noise is much greater (Figure 2.6).<sup>58</sup>

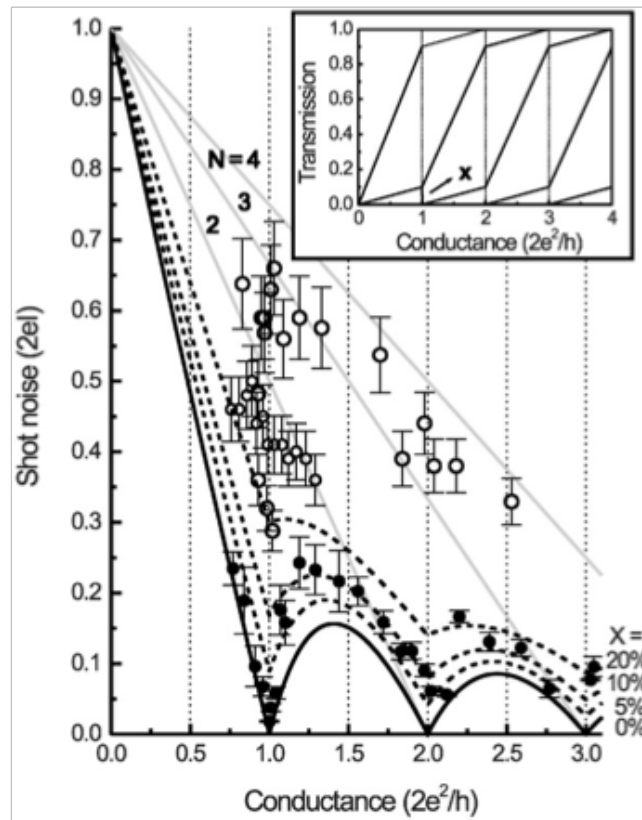
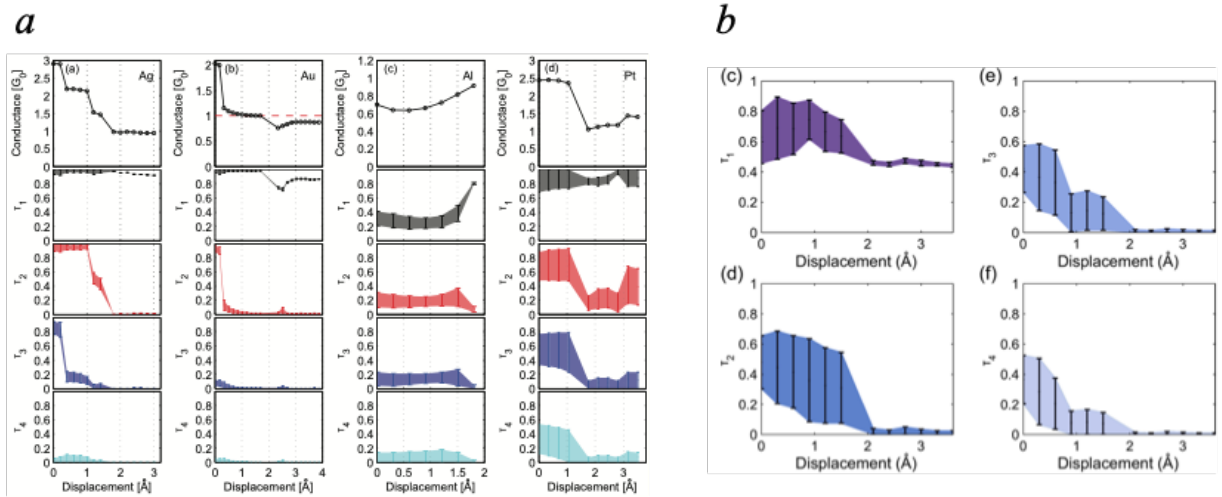


Figure 2.6: [Adapted from Ref. 58] Fano factor for current noise measured in gold (filled circles) and aluminum (open circles). While gold shows suppression at integer values of  $G_0$  due to fully transmitting channels, aluminum, which has multiple channels involved in conduction, even at  $1G_0$  does not show the same trend.

Given the Landauer model definitions for conductance,  $G = G_0 \sum_i \tau_i$ , and Fano factor,  $F = \frac{\sum_i \tau_i (1 - \tau_i)}{\sum_i \tau_i}$ , it is possible to derive the transmission probabilities for a given system of a few quantum channels if the shot noise level and conductance are well known. In systems with only one or two channels, the transmission coefficients may be found precisely,<sup>60–62</sup> and in cases of more channels, statistical analysis can still place bounds on their values (*Figure 2.7*).<sup>53,63</sup> This process was used in the work of Chapter 5 in atomic-scale gold junctions at 77 K and will be discussed further.



*Figure 2.7: a) [Adapted from Ref. 53] Conductance traces and calculated transmission coefficient ranges for silver, gold, aluminum, and platinum point contacts formed by the mechanical break junction method. b) [Adapted from Ref. 63] Transmission coefficient estimates for the four most contributing channels in nickel oxide break junctions.*

The above derivations in the zero-temperature limit require some adjustment when adding in finite temperature, namely the addition of Johnson-Nyquist noise and the smearing of

the electronic Fermi-Dirac distributions. In the finite temperature limit, the total current noise spectral density, including thermal and shot noise, is given by

$$S_I = 4k_B T G_0 \sum_i \tau_i^2 + 2eV \coth\left(\frac{eV}{2k_B T}\right) G_0 \sum_i \tau_i (1 - \tau_i). \quad (2.19)$$

In the limit of  $T \rightarrow 0$ , we recover the Landauer shot noise, Equation 2.17. In the limit of  $V=0$ , we are left with only the Johnson-Nyquist noise. In the high bias limit, since  $\lim_{eV \gg 2k_B T} \coth\left(\frac{eV}{2k_B T}\right) = 1$ , we find that the noise scales linearly with bias as in the zero temperature case. In the low-bias limit, however, the coth term causes a curvature of  $S_I(V)$  around zero bias that sharpens with decreasing temperature (*Figure 2.8a*).<sup>64</sup> This curvature is a useful feature for shot noise thermometry and can be used to discern sample temperature even when the absolute magnitude of the noise is uncalibrated.<sup>65</sup> *Figure 2.8b* demonstrates how shot noise can be used for thermometry. The zero bias current noise intercept indicates the Johnson-Nyquist noise, and at high bias, the current noise is linearly fit to  $2eI$ . If one extrapolates the high bias linear fit to the point where it intersects with the Johnson-Nyquist noise at positive and negative bias, the voltage span between the two intersection points is  $\frac{4k_B T}{e}$ .

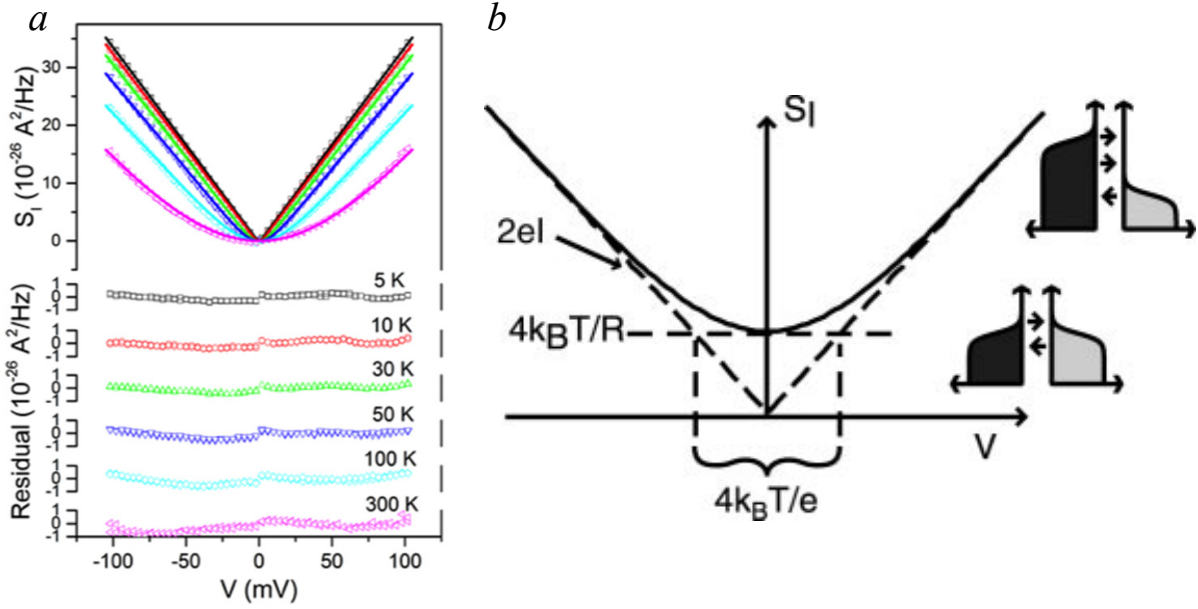


Figure 2.8: a) [Adapted from Ref. 64] Shot noise in an Au/hBN/Au tunnel junction as a function of temperature. The curvature about zero-bias sharpens as temperature decreases. b) [Adapted from Ref. 65] Zero bias curvature of shot noise can be used for thermometry.

## 2.7 Shot Noise of an $N$ -Barrier System

Returning to the model of a one-dimensional channel with a single barrier between two leads, consider what happens if instead there are two barriers. Several experiments have demonstrated shot noise for a double-barrier system.<sup>66–69</sup> For transmission probabilities  $T_1$  and  $T_2$ , in the very asymmetric limit ( $T_1 \ll T_2$  or  $T_2 \ll T_1$ ) full Poissonian shot noise is found because the dominant barrier generates all the noise, but in the symmetric case ( $T_1 \cong T_2$ ) noise is reduced by half. Theoretical approaches to the two barrier system first derived by Chen and Ting, using a quantum mechanical method,<sup>70</sup> and by Davies *et al.*,<sup>71</sup> using classical formulations, arrived at the same equation for the Fano factor,

$$F = \frac{T_1^2 + T_2^2}{(T_1 + T_2)^2}. \quad (2.20)$$

This definition of  $F$  explains well the asymmetric and symmetric regimes.

Beyond the two barrier case, de Jong and Beenakker<sup>72</sup> derived the solution to the  $n$ -barrier tunneling problem. Assuming each barrier has the same transmission probability  $T$ , they find

$$F = \frac{1}{3} \left( 1 + \frac{n(1-T)^2(2+T)-T^3}{[T+n(1-T)]^3} \right). \quad (2.21)$$

As  $n \rightarrow \infty$ ,  $F \rightarrow \frac{1}{3}$ , independent of the value of  $T$ , which is the universal limit of noise in a diffusive conductor. Multiple theories, using both quantum and classical approaches have arrived at the  $F = \frac{1}{3}$  limit for shot noise in diffusive systems.<sup>72-76</sup> The  $1/3$  suppression of shot noise is common across diffusive conductors, regardless of the exact details of the particular conductor.

The above model assumes phase coherence; the length scale was longer than the mean free path, but still shorter than the inelastic scattering length. The case of  $N$  uncorrelated noise sources was considered by Landauer using a model of a series of vacuum diodes (*Figure 2.9*).<sup>77</sup> The  $N$  diodes are all assumed to have large enough voltage to allow every electron emitted by each cathode to cross over to the anode. Therefore the current in each diode is independent of the voltage across it. Each diode exhibits independent and uncorrelated shot noise. The fluctuating current in the diode charges up the corresponding capacitance, and the fluctuating charge is then distributed to the rest of the circuit. We can consider a noise current generator in parallel with each diode (*Figure 2.9*), with each generator producing mean-square noise current  $\langle i_n^2 \rangle_{\Delta f} = 2eI\Delta f$ , where  $\Delta f$  is the frequency band of interest, and  $I$  is the externally maintained current. The noise of each generator is fed to a parallel circuit. One branch is the capacitor attributed with the particular diode generating the noise, and the other is the remaining  $(N-1)$  capacitors in series. The actual diode currents are unaffected by the fluctuating capacitive



voltages, and therefore, they are excluded from the differential signal current. The current flow through the external circuit due to one of the noise generators is equal to that through the  $(N-1)$  series capacitors

$$i_{ext} = i_n \frac{Y_{N-1}}{Y_1 + Y_{N-1}}, \quad (2.22)$$

where  $Y_{N-1}$  is the admittance of the  $(N-1)$  series capacitors and  $Y_1$  is that of a single capacitor. So then the mean-square noise of the external current is

$$\langle i_{ext}^2 \rangle_{\Delta f} = \frac{1}{N^2} \langle i_n^2 \rangle_{\Delta f}. \quad (2.23)$$

Accounting for all  $N$  uncorrelated noise generators, the total noise current is

$$\langle i^2 \rangle_{\Delta f} = \frac{1}{N} 2eI\Delta f. \quad (2.24)$$

Equation 2.24 arises due to the nature of shot noise being a measure of stochastically independent events.

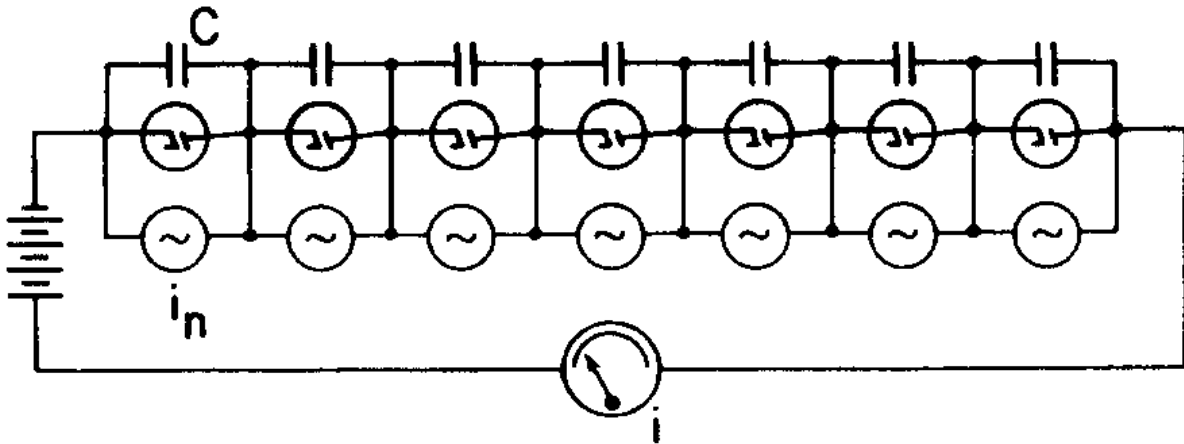


Figure 2.9: [Adapted from Ref. 77] Shot noise in the incoherent limit modelled as a chain of  $N$  diodes in series, each with its own capacitance. Each combination of diode and capacitor is associated with a parallel noise current generator,  $i_n$ . The noise by each individual generator is measured in the external circuit.

## 2.8 Effects of Interactions on Shot Noise

As mentioned in *Section 2.5*, interactions can cause the Fano factor to deviate away from 1. The above definitions for the Fano factor were derived excluding interaction effects. When the size of the system of interest is comparable to the inelastic mean free path, however, electron-electron and electron-phonon scattering can greatly modify the shot noise. *Figure 2.9*, adapted from the work of Steinbach *et al.*,<sup>54</sup> outlines the Fano factors for current noise in a metallic resistor of varying length  $L$  and resistance  $R$ . The length scales shown reflect the elastic mean free path  $l$ , the phase coherence length  $L_\phi$ , the electron-electron inelastic scattering length,  $L_{e-e}$ , and the electron-phonon inelastic scattering length,  $L_{e-ph}$ . The electron-electron inelastic scattering length is the length scale over which inelastic collisions between electrons cause their thermalization, and likewise, the electron-phonon inelastic scattering length is the length scale over which the conduction electron temperature relaxes to the phonon temperature.<sup>54</sup>

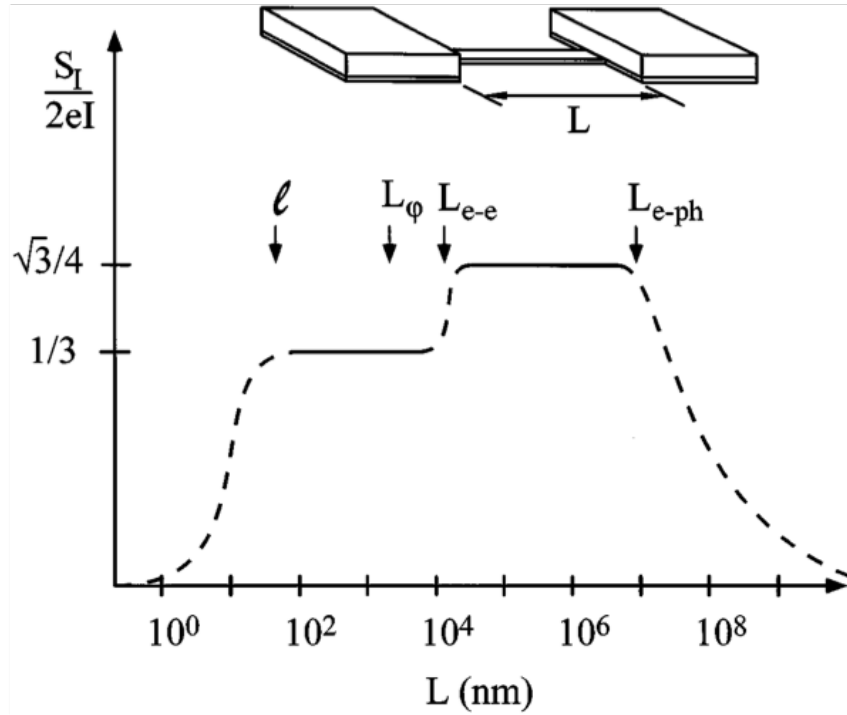


Figure 2.10: [Adapted from Ref. 54] Fano factor as a function of resistor length, broken into regimes in terms of the elastic mean free path,  $l$ , the phase coherence length  $L_\phi$ , the electron-electron inelastic scattering length,  $L_{e-e}$ , and the electron-phonon inelastic scattering length,  $L_{e-ph}$ .

Phonons are quasiparticles that are the quantization of crystal lattice vibrational modes, or displacements of the atomic lattice from its equilibrium position, in bulk solids. They are bosons, and therefore, there can be many phonons occupying the same mode. The number of phonons for a given mode and temperature is given by Bose-Einstein statistics,  $\langle n \rangle = \frac{1}{e^{\frac{\hbar\omega}{k_B T}} - 1}$ , and the amplitude of the vibrations increases with an increasing number of phonons due to constructive interference. Phonons are able to scatter electrons by imposing a finite lifetime on the Bloch states of the electrons in the lattice. The scattering of conduction electrons by phonons

is the primary reason metals have resistivity at room temperature ( $\sim 26$  meV) as typical phonon energies are on the order of 1-100 meV. For example, gold optical phonons have an energy of approximately 17 meV. At lower temperatures, however, as  $k_B T \ll \hbar\omega$ , the occupation number  $n$  goes to zero; in other words, the vibrational states freeze out.

In the model of the variable-length resistor, assume the resistor is connected on either end to ideal thermal and electrical reservoirs that impose an electron temperature  $T$  and chemical potentials  $\mu_L$  and  $\mu_R$  such that  $eV = -(\mu_L - \mu_R)$ , where  $V$  is the applied bias across the resistor.

When  $L \ll l$ , the system is in the ballistic or point-contact regime, and shot noise is described by the Landauer model from *Section 2.6*. On this length scale, the resistor itself produces no shot noise, indicated by the dashed line in *Figure 2.10* going to zero. Shot noise in this case is generated by the barriers set by transmission into the contacts. For  $l \ll L \ll L_\phi$  and for  $L_\phi \ll$

$L \ll L_{e-e}$ ,  $F = \frac{1}{3}$ . In the absence of inelastic scattering by phonons or other electrons, the electronic distribution in the resistor is an average of the Fermi-Dirac distributions of the source and drain. The result is a position-dependent, two-step non-Fermionic distribution that converges to the Fermi-Dirac distributions at either end. On the length scale  $L_{e-e} \ll L \ll L_{e-ph}$ , the interacting hot electron regime, electrons now interact with one another inelastically, and the two-step distribution becomes smeared. Electrons become hot via Joule heating, thermalizing their energy via electron-electron inelastic scattering, and they can only cool down by thermal conduction to the source and drain.<sup>78-80</sup> Theoretical predictions<sup>78,81,82</sup> set  $F = \frac{\sqrt{3}}{4}$  for this regime due to the local heating effect, which was confirmed experimentally by Steinbach *et al.*<sup>54</sup>

As  $L$  continues to grow, electron-phonon interactions become important, and shot noise decays with increasing length. For macroscopic systems, when  $L \gg L_{e-ph}$ , electrons have thermalized to the phonon temperature via inelastic scattering, and shot noise is fully

suppressed.<sup>33,73</sup> Beenakker and Büttiker modelled the suppression of shot noise in a macroscopic conductor by first considering a conductor with three contacts: source (1), drain (2), and a voltage probe (3) between the source and drain.<sup>73</sup> The total conductance for this system is

$$G = G_0 \left( T_{21} + \frac{T_{23}T_{31}}{T_{31}+T_{32}} \right) \quad (2.22)$$

where  $T_{21}$  represents coherent transport from source to drain. The second term represents incoherent transport in which an electron reaches the voltage probe (contact 3) and is then replaced by an electron from the contact 3 reservoir, which is out of phase with the original electron. The result is a decoherence effect. Assuming completely incoherent transport ( $T_{21} = T_{12} = 0$ ) and an ideal voltmeter ( $I_3 = 0$ ), shot noise is reduced by half due to the inelastic scattering. It follows that as length is increased well beyond  $L_{e-ph}$  and many decohering events occur, shot noise is ultimately fully suppressed.

The  $I/N$  reduction factor derived by Landauer in the previous section<sup>77</sup> is closely related to the Beenakker and Büttiker model. Landauer argues it arises due to the fact that in the diode chain (*Figure 2.9*), as well as in the many reservoir limit of the Beenakker-Büttiker case,<sup>73</sup> an electron traveling through the whole sample is not an event that occurs within some well-defined time limit. The re-emission of an electron from a cathode is not closely correlated to the arrival event at an anode. Analogously, the transmission or reflection of an electron at one reservoir is not clearly linked to similar events at other reservoirs.

# Chapter 3

## Topological Insulators

Topological insulators (TIs) are of great interest due to their potential ability to achieve the quantum coherent transport discussed in Chapter 1 on the macroscopic scale. TIs are materials that have a bulk band gap like normal insulators, but also exhibit “topologically-protected” ballistically conductive edges (two-dimensional TIs) or surface states (three-dimensional TIs). These properties arise due to a combination of spin-orbit coupling and time-reversal symmetry.

Many other quantum phases in condensed matter systems, including crystalline solids, magnets, and superconductors, are described by some mode of spontaneous symmetry breaking based on the Landau-Ginzburg theory of phase transitions.<sup>83</sup> Landau-Ginzburg theory assumes that phase transitions are described by some order parameter, such as magnetization, the behavior of which characterizes the nature of the transition. In the “normal” state, the order parameter is zero, but it gains some finite value when the system is in the ordered state, coinciding with an abrupt alteration (usually a reduction) in the symmetry of the system. For example, magnets break rotational symmetry and crystalline solids break translational symmetry.

In 1980, however, von Klitzing *et al.* demonstrated a new quantum state that could not be described by the Landau-Ginzburg model.<sup>84</sup> In the quantum Hall (QH) state, described in detail in *Section 3.2*, two-dimensional materials at very low temperatures under high magnetic fields exhibit exceptionally precise quantized Hall conductance,  $\sigma_H$ , with values either in integer or fractional multiples of  $\frac{e^2}{h}$ . This quantization is remarkably immune to random disorder in

samples. Quantum Hall effect (QHE) systems also exhibit dissipationless longitudinal charge transport along the sample edges. The quantum Hall state was the first example of a quantum state that does not break any symmetries. Instead, Thouless *et al.* used the concept of topological order to characterize this 2D phenomenon.<sup>85</sup> The QH state is considered a topological phase in that specific fundamental properties, including the quantized Hall conductance and the number of gapless boundary modes, are unaffected by smooth changes in material parameters and can only be altered by quantum phase transitions as described below.

### ***3.1 Topological Order***

In mathematics, topological order was introduced to group together geometrical objects into general classes. The topological invariant common to a particular group is a characteristic that remains unchanged under continuous transformations. For example, 2D objects are classified by their number of holes. A perfect circle and an ellipse are topologically equivalent because one can be smoothly deformed into the other without creating any holes. In physics, the concept of smooth deformations can be considered in terms of the Hamiltonian for a many-particle system.<sup>86</sup> If the general Hamiltonian of the system describes the energy gap separating the ground state from the excited state, a smooth deformation is a change in the Hamiltonian that does not close the bulk gap. Thus, in the way that an object with one hole cannot be smoothly deformed into an object with two holes, one gapped state cannot be deformed into another gapped state in a different topological class without undergoing a quantum phase transition in which the system becomes gapless.

Thouless and his group employed topological arguments to describe the quantum Hall effect by accumulating all possible wavefunctions that could describe electrons in a 2D material

and representing the entire set as a curved surface.<sup>85</sup> The topological invariant of the surface is a stable number describing the set of wavefunctions, such as the Hall conductance. The robustness of the QHE and the quantization of the Hall conductance is due to the stability of the topological invariant; while small changes in the material may change individual electronic wave functions, it is much more difficult to alter the set as a whole.

### ***3.2 The Quantum Hall Effect***

To understand the idea of topological protection and edge states in 2D topological insulators, we can learn important lessons from the quantum Hall effect. The quantum Hall effect (QHE) is the quantum mechanical version of the Hall effect and is observed in systems at low temperatures and strong external magnetic fields. Charge transport in a QHE system is carried out by chiral modes at the edges of a Hall bar, and the Hall conductance takes on quantized values with remarkable precision. The quantum Hall effect was later expanded to the quantum spin Hall effect (QSHE), in which conduction is carried out by parallel spin-momentum-locked edge modes, known as helical modes. At a single edge, the left-moving mode has opposite spin of the right-moving mode. The two helical edge states form a Kramers pair, protected by time reversal symmetry. The QSHE forms the basis for the model of a two-dimensional topological insulator, often called quantum spin Hall insulators.

#### ***3.2.1 The Classical Hall Effect and the Discovery of the Quantum Hall Effect***

The Hall effect, discovered by Edwin Hall in 1879,<sup>87</sup> describes the potential difference in a conductor that arises perpendicular to the flow of current due to the presence of an external magnetic field. *Figure 3.1* shows an example of a Hall effect measurement setup. Current  $I$



flows in the  $x$ -direction through a 2D conductor of width  $W$ . Two pairs of voltage probes are situated on either side of the conductor, each separated by length  $L$ . Potential differences are measured both longitudinally, between two probes on the same side,  $V_{xx}$ , and transversely, between two probes on opposite sides,  $V_{xy}$ . The conductivity or resistivity of a 2D electron system can readily be measured from  $V_{xx}$ . The electric field experienced by the charge carriers in the absence of a magnetic field is given by

$$\mathbf{E} = \frac{V_{xx}}{L}, \quad (3.1)$$

and the current density is

$$\mathbf{j} = \frac{I}{W}. \quad (3.2)$$

The longitudinal conductivity in the absence of a magnetic field is then

$$\sigma_{xx} = \frac{j}{E} = \frac{IL}{WV_{xx}}. \quad (3.3)$$

The longitudinal conductivity is equivalent for the upper or lower sets of contacts unless a magnetic field  $\mathbf{B}$  is applied perpendicular to the plane of the conductor. In that case, charge carriers will be deflected based on the Lorentz force,  $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ , where  $q = -e$ , the electron charge, and  $\mathbf{v}$  is the longitudinal drift velocity of the charges. Depending on the sign of  $\mathbf{B}$ , electrons will either be deflected toward the top or bottom edge of the conductor. Since the electrons cannot escape, an accumulation of charge builds up, resulting in a transverse electric field  $E_y$  that opposes the Lorentz force. A steady state is reached when enough charge has built up to exactly balance the transverse electric field and the Lorentz force, and again, current only flows in the  $x$ -direction. The resulting potential difference between the top and bottom edge,  $V_{xy}$ , is known as the Hall voltage.

The steady state condition is

$$E_y = Bv. \quad (3.4)$$

If  $n$  is the 2D charge carrier density, then  $v = \frac{j}{ne}$ . The Hall resistance, defined as the transverse potential difference,  $V_{xy} = E_y W$ , divided by the current,  $I = jW$  and normalized for the magnetic field is

$$R_H = \frac{WE_y}{WjB} = \frac{1}{ne}. \quad (3.5)$$

The Hall effect, then, is helpful in deducing the composition of a conductive 2D system. In terms of conductivity we have

$$\sigma_{xx} = \frac{IL}{V_{xx}W} = ne\mu \quad (3.6)$$

$$\text{and } \sigma_{xy} = \frac{ne}{B},$$

where  $\mu = \frac{v}{E}$  is the electron mobility.

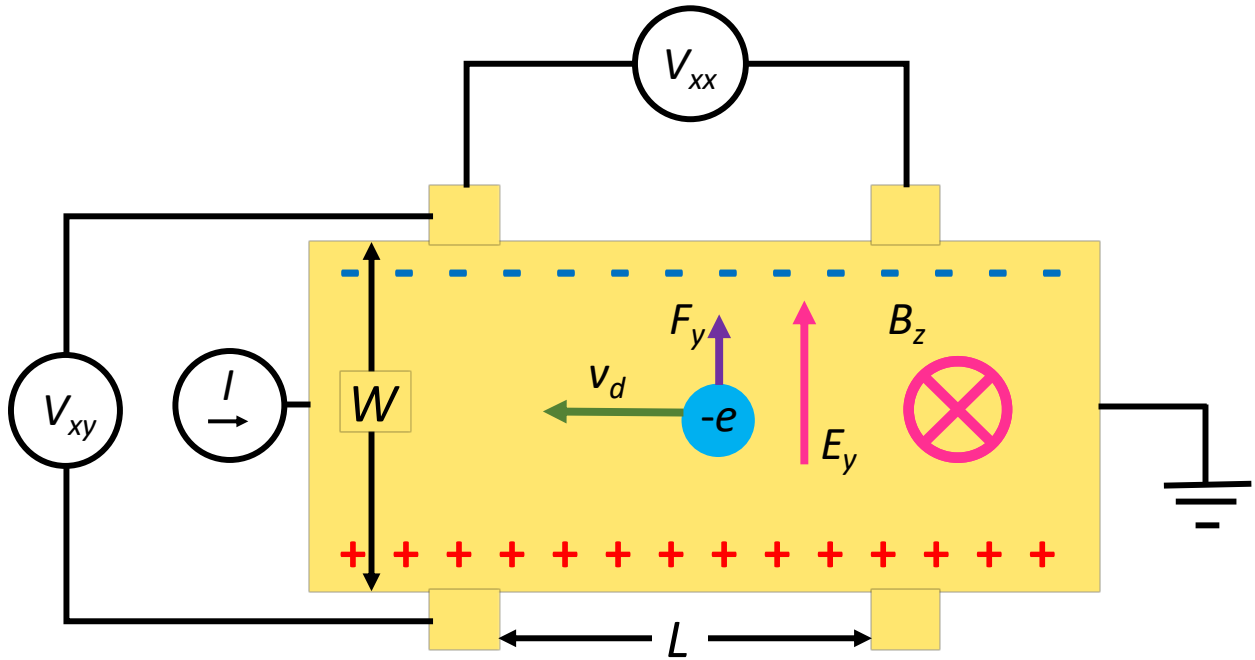


Figure 3.1: Example of a typical Hall bar setup.

If  $B$  is sufficiently large, then cyclotron orbits of the conduction electrons are possible. Cyclotron orbits describe the circular motion, sometimes superimposed with lateral motion, of electrons in a stationary, uniform magnetic field due to the Lorentz force. The angular frequency of cyclotron motion is given by  $\omega_c = \frac{eB}{m_e}$ , where  $B$  is the magnitude of the magnetic field,  $e$  is the electron charge, and  $m_e$  is the electron mass. For a very clean system, free from defects or impurities, electrons may undergo many orbits between scattering events, and the resulting phase-coherent transport requires a quantum mechanical approach.

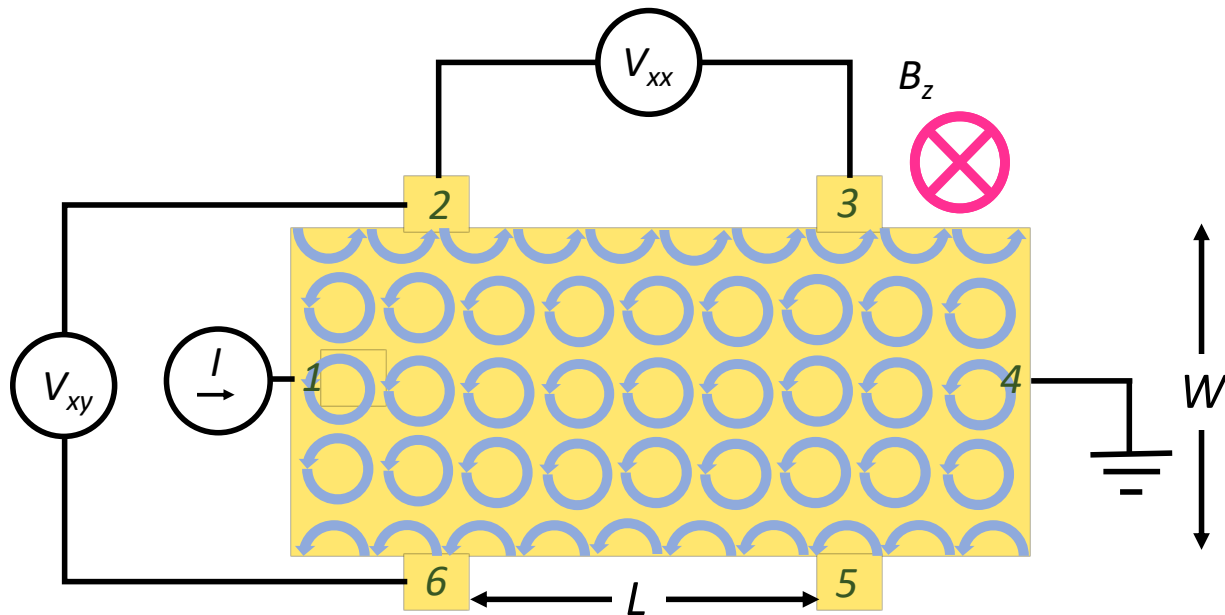


Figure 3.2: Under very large perpendicular magnetic field, electrons in a Hall bar begin to move in cyclotron orbits with chiral transport along the edges.

First, though, consider the more simplified picture in Figure 3.2. In the bulk of the Hall bar, electrons can no longer carry the longitudinal current, but at the top and bottom, carriers move via “skipping orbits” by bouncing off the edges. This mode of transport in which a particle is restricted to move in a single direction bouncing along a straight line is referred to as chiral.

Due to the disparity in charge accumulation, carriers at the top edge move in the opposite direction of carriers at the bottom. Since no transport occurs in the bulk, the two edges are spatially separated by an insulating barrier and cannot scatter into one another or to any other states. Due to this long-range coherence, the chiral edge transport can be described by the Landauer-Büttiker model. Based on the labelling in *Figure 3.2*, the transmission coefficients are  $\tau_{i \rightarrow i+1} = 1$  and zero otherwise. Using the Landauer-Büttiker picture, the current flowing out of probe  $i$  is given by

$$I_i = \frac{e^2}{h} \sum_j (\tau_{ji} V_j - \tau_{ij} V_i), \quad (3.7)$$

where  $\tau_{ji}$  is the transmission coefficient from probe  $j$  to probe  $i$ ,  $V_j$  is the voltage of probe  $j$ , and vice versa. In Equation 3.7, spin has been neglected for now; accommodating for spin requires an additional sum over possible spin orientations. As no current flows into or out of probes 2,3,5, or 6, the voltages  $V_i$  at each terminal equilibrate, so that  $V_1 = V_2 = V_3$  and  $V_4 = V_5 = V_6 = 0$ .  $V_1$  is determined from the current  $I$  and the overall source-drain conductivity. In terms of a noninteracting 1D channel in the Landauer-Büttiker model, we can reasonably assume the conductivity is  $\frac{e^2}{h}$ . Then,  $V_{xx} = 0$ , and  $V_{xy} = \frac{Ih}{e^2}$ . This is the limit of the quantum Hall effect, when the longitudinal and transverse resistivities are functions of the magnetic field or carrier density given by  $\rho_{xx} = 0$  and  $\rho_{xy} = \frac{h}{ne^2}$ , where  $n$  is an integer. Similar to conductance measurements in atomic scale contacts, the value of  $n$  in QH measurements can reveal the number of chiral edge channels involved in transport.

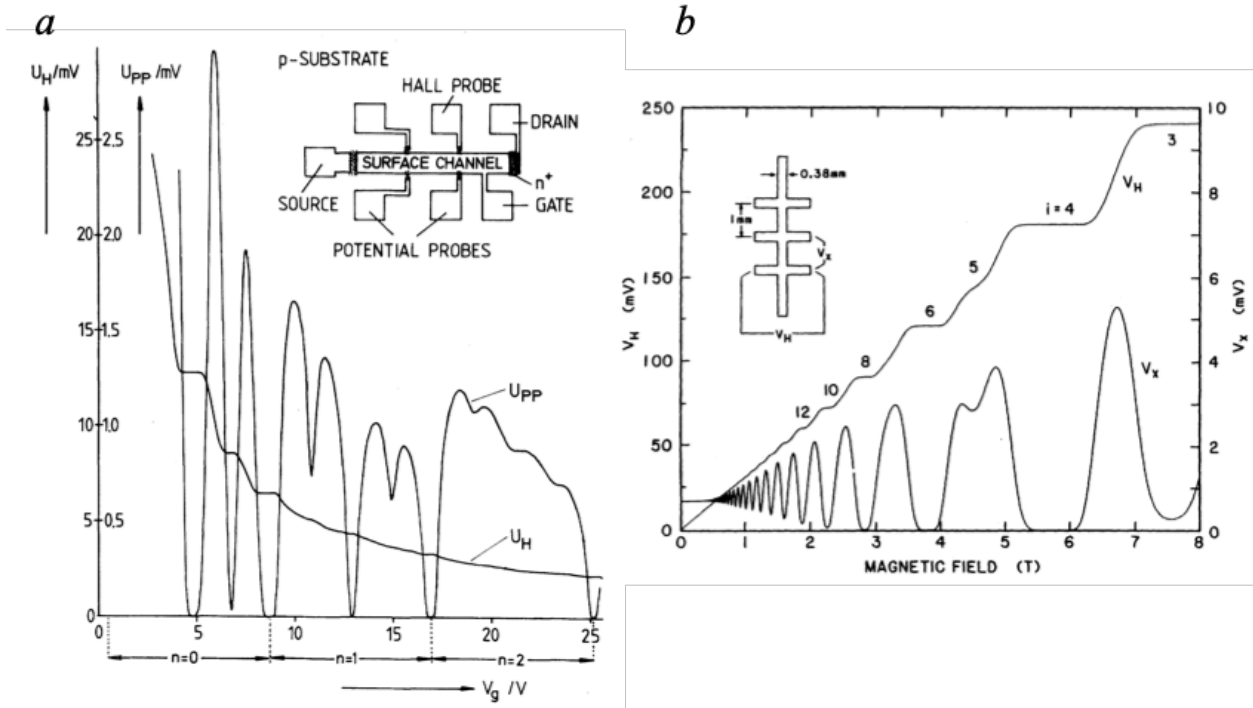


Figure 3.3: a) [Adapted from Ref. 89] First evidence of the Quantum Hall effect. b) [Adapted from Ref. 88] Another example of oscillatory longitudinal resistance and quantized Hall resistance by Cage *et al.*<sup>88</sup>

Von Klitzing *et al.* were the first to observe this phenomenon in a 2D inversion layer of silicon metal-oxide-semiconductor field effect transistor at a temperature of 1.5 K and magnetic field of 18 T.<sup>84</sup> They found that as a function of the gate voltage, the longitudinal resistance showed strong oscillatory behavior, while the Hall resistance exhibited plateaus corresponding to the minima in the longitudinal resistance (*Figure 3.3*). Even from the initial measurements, the QHE shows incredible precision, with parts-per-billion measurements now being standard, and the QHE is now the basis of a metrological standard.<sup>89</sup>

### 3.2.2 Quantum Approach to the Quantum Hall Effect

Consider a 2D electron system, with infinite length in  $x$  and width  $W$  in  $y$ . Confinement in  $y$  is enforced by an infinite potential

$$V(y) = \begin{cases} 0, & |y| < \frac{W}{2} \\ \infty, & \text{otherwise} \end{cases}. \quad (3.8)$$

Apply some transverse magnetic field  $\mathbf{B} = B\hat{z}$ . The Hamiltonian in this case for an electron of mass  $m$  and charge  $-e$  is given by

$$\hat{H} = \frac{1}{2m} (\hat{\mathbf{p}} - e\mathbf{A})^2 + V(y) \quad (3.9)$$

where  $\hat{\mathbf{p}}$  is the canonical momentum operator, and  $\mathbf{A}$  is the vector potential, which is related to the magnetic field by

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (3.10)$$

The Hamiltonian is gauge invariant, meaning the physical properties are not influenced by the specific gauge chosen. The simplest choice, then, is the Landau gauge, which sets the vector potential as

$$\mathbf{A} = \begin{pmatrix} -By \\ 0 \\ 0 \end{pmatrix} \quad (3.11)$$

so that  $\mathbf{B} = \nabla \times \mathbf{A} = B\hat{z}$ . The Hamiltonian then becomes

$$\hat{H} = \frac{1}{2m} (\hat{\mathbf{p}} - e\mathbf{A})^2 + V(y) = \frac{1}{2m} (\hat{p}_x + eBy)^2 + \frac{1}{2m} \hat{p}_y^2 + V(y). \quad (3.12)$$

Since  $x$  does not appear in the equation, the  $x$ -momentum operator  $\hat{p}_x$  commutes with the Hamiltonian, meaning it can be replaced by its eigenvalue  $\hbar k_x$ , and we can rewrite the Hamiltonian

$$\hat{H} = \frac{1}{2m} \hat{p}_y^2 + \frac{1}{2} m \left( \frac{eB}{m} \right)^2 \left( y + \frac{\hbar k_x}{eB} \right)^2 + V(y). \quad (3.13)$$

When  $|y| < \frac{W}{2}$  so that  $V(y) = 0$ , this is the exact Hamiltonian for the quantum Harmonic oscillator ( $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$ ) centered about  $y = -\frac{\hbar k_x}{eB}$  with cyclotron frequency  $\omega_c = \frac{eB}{m}$ . So in the limit that the electron's wavefunctions are confined within  $\pm \frac{W}{2}$ , the energy spectrum of the electrons can be described by discrete values  $E_n = \left(n + \frac{1}{2}\right)\hbar\omega_c$ , for  $n \geq 0$ . Each set of wavefunctions with the same  $n$  value is known as a Landau level.

In units of length, the width of the ground state of the simple harmonic oscillator is on the order of  $\sqrt{\frac{\hbar}{m\omega_c}}$ . In order for the Landau level spectrum to be valid within the interior of the conductor, this width, called the magnetic length  $l_B$ , must be much smaller than the width of the channel,  $l_B = \sqrt{\frac{\hbar}{eB}} \ll W$ . For perspective,  $\sqrt{\frac{\hbar}{eB}} \approx 26$  nm for  $B = 1$  T. If  $|k_x|$  is large enough, however, the center point of the magnetic potential will be outside of the well edges at  $y = \pm \frac{W}{2}$ , forming a nearly triangular well and raising the ground state energy (*Figure 3.4a,b*).<sup>90</sup> As a function of  $y$ , the levels increase to some large value or even to infinity, depending on the details of  $V(y)$  at the edges (*Figure 3.4c*).

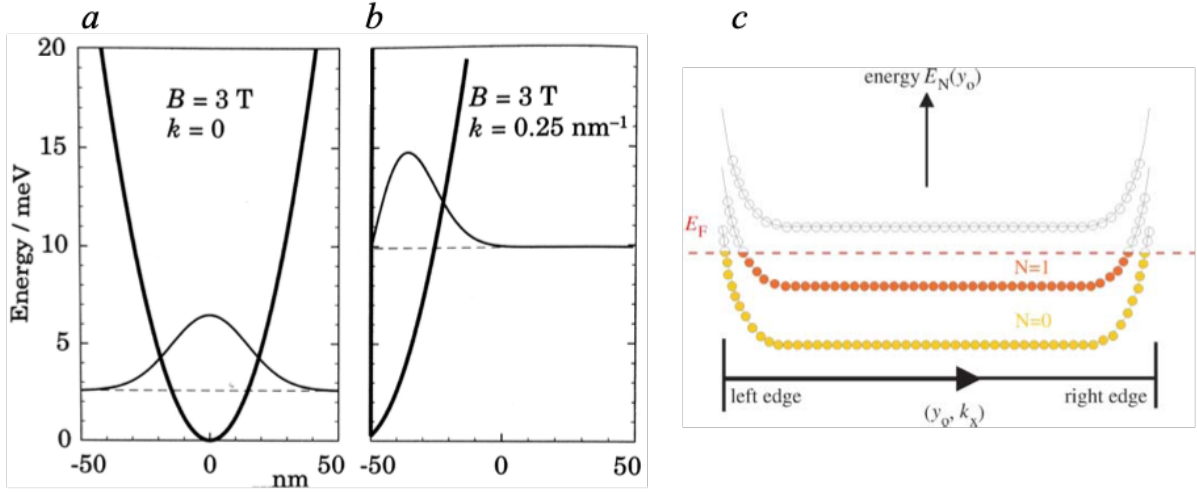


Figure 3.4: a,b) [Adapted from Ref. 90] Magnetic potential and hard wall potential as a function of  $y$  for two values of  $k_x$ , including the wavefunction and energy level of the ground state for the Hamiltonian from Equation 3.13. c) [Adapted from Ref. 89] Ideal Landau levels for a system with boundaries. Fully occupied levels toward the center of the device rise in energy close to the edges, resulting in metallic modes near the intersection of the Landau levels with the Fermi energy.

Within the interior, the Landau levels are highly degenerate (Figure 3.4c). At a given energy,  $k_x$  can be any value that fulfills two requirements. Firstly, for a finite length in  $x$ , given by  $L$ ,  $k_x$  will assume integer values of  $\frac{2\pi}{L}$ , or  $k_x = N \frac{2\pi}{L}$ , where  $N$  can be any positive or negative integer. Secondly,  $k_x$  must be small enough to keep the center of the magnetic potential within the interior. Together, these limits state

$$\left| \frac{2\pi\hbar N}{eBL} \right| < \frac{W}{2} \text{ or } |N| < \frac{eB}{2\hbar} LW. \quad (3.14)$$

The degeneracy per unit area ( $\frac{N}{LW}$ ), then, is  $\frac{eB}{h}$ , counting both positive and negative values of  $N$ .

In terms of magnetic flux, if  $\Phi = BLW$  is the magnetic flux through unit area  $LW$ , each electron



occupies in real space the area of one flux quantum,  $\Phi_0 = \frac{h}{e}$ , and the degeneracy per Landau level is

$$N = \frac{\Phi}{\Phi_0}. \quad (3.15)$$

In the Landau gauge, it is simple to add in an electric field along the  $y$ -direction by adding an electric potential  $\phi = \mathcal{E}y$ , which changes the Hamiltonian to

$$\hat{H} = \frac{1}{2m} \hat{p}_y^2 + \frac{1}{2} m \left( \frac{eB}{m} \right)^2 (y - y_0)^2 + e\mathcal{E}y + \frac{1}{2} m \frac{\mathcal{E}^2}{B^2}, \quad (3.16)$$

where the new center of the harmonic oscillator is  $y_0 = -\frac{\hbar k_x}{eB} - \frac{m\mathcal{E}}{eB^2}$ .

The total energy of this new displaced harmonic oscillator, including contributions from both the harmonic oscillator and drift motion is

$$E = \hbar\omega_c \left( n + \frac{1}{2} \right) - e\mathcal{E} \left( \frac{k_x \hbar}{eB} + \frac{m\mathcal{E}}{eB^2} \right) + \frac{1}{2} m \frac{\mathcal{E}^2}{B^2}. \quad (3.17)$$

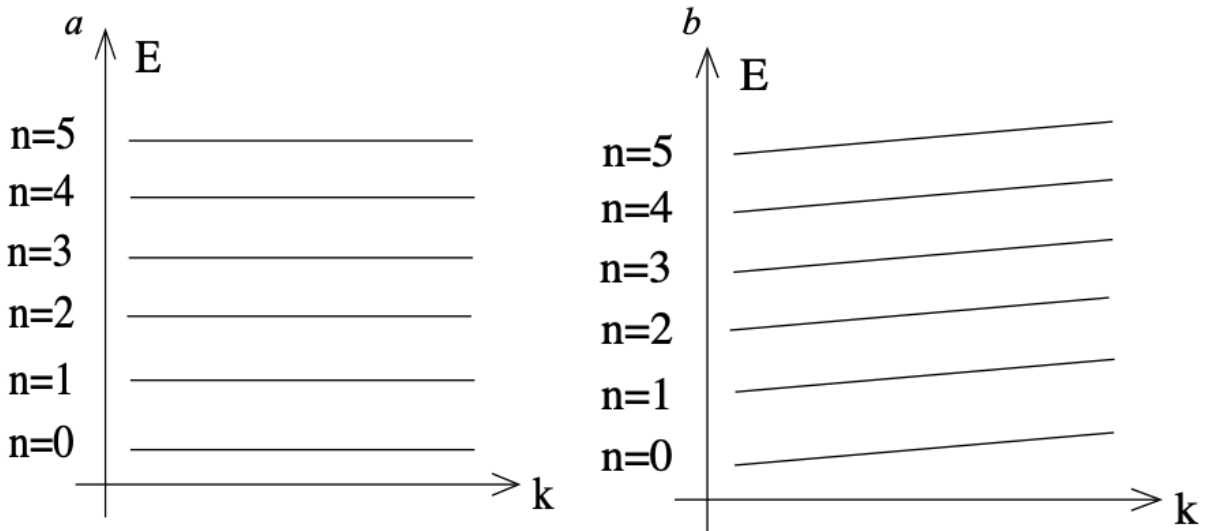


Figure 3.5: [Adapted from Ref. 91] Density of states of ideal Landau levels in case of zero (a) and nonzero (b) electric field.<sup>91</sup>

Under an electric field, the degeneracy in each Landau level has now been lifted, and the energy in each level now depends linearly on  $k_x$  (*Figure 3.5b*). Since the energy now has a momentum dependence, states now begin to drift in the  $x$ -direction. The group velocity is given by

$$v_x = \frac{1}{\hbar} \frac{\partial E}{\partial k_x} = \frac{\mathcal{E}}{B}. \quad (3.18)$$

Therefore, a wavepacket with momentum  $k_x$  is centered about position  $y_0 = -\frac{k_x \hbar}{eB} - \frac{e\mathcal{E}}{m\omega_c^2}$  with potential energy  $-e\mathcal{E} \left( \frac{k_x \hbar}{eB} + \frac{e\mathcal{E}}{m\omega_c^2} \right)$ , and kinetic energy  $\frac{1}{2} m v_x^2$ .

To determine the transverse and longitudinal conductivities or resistivities, we need to determine the expression for current in this framework. Begin by deriving the conductivity for a single free particle. Given the drift velocity  $v_x = \frac{\mathcal{E}}{B}$ ,

$$I_x = -en \sum_{k_x} \frac{\mathcal{E}}{B} \quad (3.19)$$

for  $n$  filled Landau levels. The sum over  $k_x$  just returns the degeneracy per Landau level,  $N = \frac{BLW}{\Phi_0}$ . Divide through by the area  $LW$  to get the current density

$$J_x = \frac{-en\mathcal{E}}{\Phi_0}. \quad (3.20)$$

Given the relation between the current density and the electric field

$$\mathbf{J} = \sigma \mathbf{E} \quad (3.21)$$

where  $\sigma$  is the conductivity tensor

$$\begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix}, \quad (3.22)$$

we know

$$\vec{\mathcal{E}} = \begin{pmatrix} 0 \\ \mathcal{E} \end{pmatrix} \rightarrow \vec{J} = \begin{pmatrix} \frac{en\mathcal{E}}{\Phi_0} \\ 0 \end{pmatrix}, \quad (3.23)$$

so

$$\sigma_{xx} = 0 \text{ and } \sigma_{xy} = \frac{en}{\Phi_0} \quad (3.24)$$

or in terms of resistivity

$$\rho_{xx} = 0 \text{ and } \rho_{xy} = \frac{\Phi_0}{en} = \frac{h}{e^2n}. \quad (3.25)$$

$\rho_{xy} = \frac{h}{e^2n}$  is the exact resistivity seen on the quantum Hall plateau. The prefactor  $n$  is called the filling factor, appropriately named for indicating the number of filled Landau levels.

### 3.2.3 Edge Modes

As described above, in a fixed magnetic field, particles will move in cyclotron orbits, with all electrons orbiting in the same direction (still neglecting spin). The particles at the edge, however, must collide with the boundary. As they cannot bounce back in the other direction, they travel forward along the wall in a skipping motion described as chiral. Particles at either edge travel in opposite directions so that in the absence of a longitudinal electric field, the total current vanishes. Such edge states are well known at the interface between an integer QH system and vacuum.<sup>92</sup>

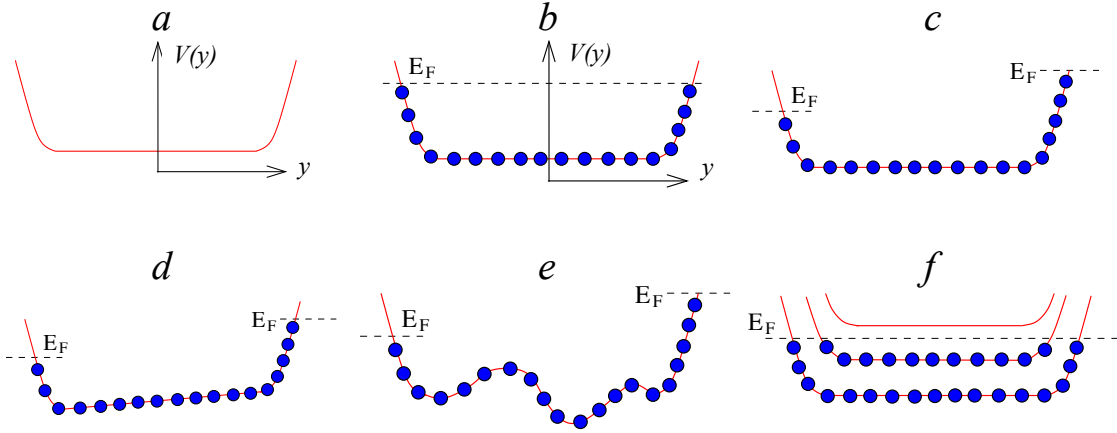


Figure 3.6: [Adapted from Ref. 91] a) Potential similar to Equation 3.8. b) Available states are filled, setting the Fermi level. c) Applying a potential difference across the two sides of the sample. d) Accounting for the tilt of the Landau level under an electric field. e) Arbitrary potential across the interior of the sample. f) Generalizing the model to multiple levels.

The chiral edge modes can also be derived from the quantum approach. Figure 3.6a gives a picture of a potential generally similar to that described in Equation 3.8, flat within the interior, but rising sharply at the edges. The Hamiltonian is still  $\hat{H} = \frac{1}{2m} \hat{p}_y^2 + \frac{1}{2} m \left(\frac{eB}{m}\right)^2 \left(y + \frac{\hbar k_x}{eB}\right)^2 + V(y)$ .

If the potential is smooth over distances on the order of the magnetic length  $l_B = \sqrt{\frac{\hbar}{eB}}$ , then for each state, we can Taylor expand the potential around its position  $y'$ .

$$V(y) = V(y') + \frac{\partial V}{\partial y}(y - y') + \dots \quad (3.26)$$

Dropping quadratic and higher terms and neglecting the constant term, we have a linear potential, the same as the energy from Equation 3.17 for Landau levels in an electric field. The drift velocity in  $x$  is now given by

$$v_x = \frac{1}{eB} \frac{\partial V}{\partial y}. \quad (3.27)$$

Each wavefunction, denoted by its  $k_x$ , is centered about a different  $y' = \frac{k_x \hbar}{eB}$  with a different drift velocity. Within the interior,  $\frac{\partial V}{\partial y}$  is negligible, and electron motion is suppressed. At each edge, the modes are chiral, with  $v_x > 0$  at the top (positive  $y$  edge) and  $v_x < 0$  on the bottom.

Now consider what happens as the available states are filled (*Figure 3.6b*) by introducing a chemical potential and thereby setting the Fermi level,  $E_F$ . Though the states are labelled by their  $x$ -momentum, we have seen they can also be thought of in terms of their position in  $y$ . In (*Figure 3.6b*), it is clear the states in the interior are insulating because all the states in the band are filled. The edges, however, lie near the Fermi level, and can therefore carry current. If we introduce a potential difference,  $\Delta\mu$ , on the two sides of the sample (*Figure 3.6c*), we can now fill more states on one side versus the other. The current is given by summing over all filled states, or approximating by integrating over  $y$

$$I_x = -e \int \frac{dk}{2\pi} v_x(k) = \frac{e^2 B}{h} \int dy \frac{1}{eB} \frac{\partial V}{\partial y} = \frac{e}{h} \Delta\mu. \quad (3.28)$$

The Hall voltage,  $V_H = \frac{\Delta\mu}{e}$ , then gives the conductivity

$$\sigma_{xy} = \frac{I_x}{V_H} = \frac{e^2}{h}, \quad (3.29)$$

which is equivalent to the Hall conductance found in Equation 3.24 for a single Landau level.

This picture implies all current is carried by the edges since in the bulk, the Landau level is flat, but this is not quite accurate. Consider *Figure 3.6d*, where we have accounted for the tilting of the Landau level due to the electric field or *Figure 3.6e*, in which  $V(y)$  takes on a random form. As long as  $V(y)$  is smooth over length scales on the order of  $l_B$  and does not

extend above  $E_F$ , though, Equation 3.29 still holds and current is still carried by the edge modes. This is vital to the robustness of the QHE.

Extending to multiple Landau levels, as long as  $E_F$  lies between two Landau levels such that we have  $n$  filled levels, we can expect a picture like *Figure 3.6f*. In this case, there are  $n$  chiral modes on each edge. These chiral modes are incredibly robust. Since all states on an edge move in the same direction, in order to scatter a left-moving electron into a right-moving electron, it would have to traverse the width of the sample. The width is very large, and so scattering is largely suppressed. Chiral modes, then are immune to scattering by impurities.

### ***3.2.3 The Role of Disorder and Laughlin's Gedankenexperiment***

The derivations for the QHE have so far neglected disorder. The origin of the quantized conductance plateaus, though, is partially due to the inherently imperfect nature of experimental samples. To understand the role of disorder in the QHE, consider the gedankenexperiment (thought-experiment) constructed by R.B. Laughlin in 1981.<sup>93</sup> Laughlin proposed that since  $\sigma_{xy}$  is quantized, it should be impervious to the details of the sample geometry and one should be able to smoothly deform the rectangular shape. Imagine taking the conducting channel from the previous section and wrapping it around itself into a circular ribbon (*Figure 3.7*)<sup>94</sup> so that it forms a loop with circumference  $L$ . Let the axes be such that the  $x$ -axis is along the circumference of the loop, and the  $y$ -axis is parallel to the axis of the cylinder. There is still an external magnetic field normal to the surface, and the Fermi level still lies in a gap between Landau levels with  $n$  filled levels below it. If a persistent current flows around the loop, a voltage difference arises between the two edges of the ribbon due to the Hall effect.

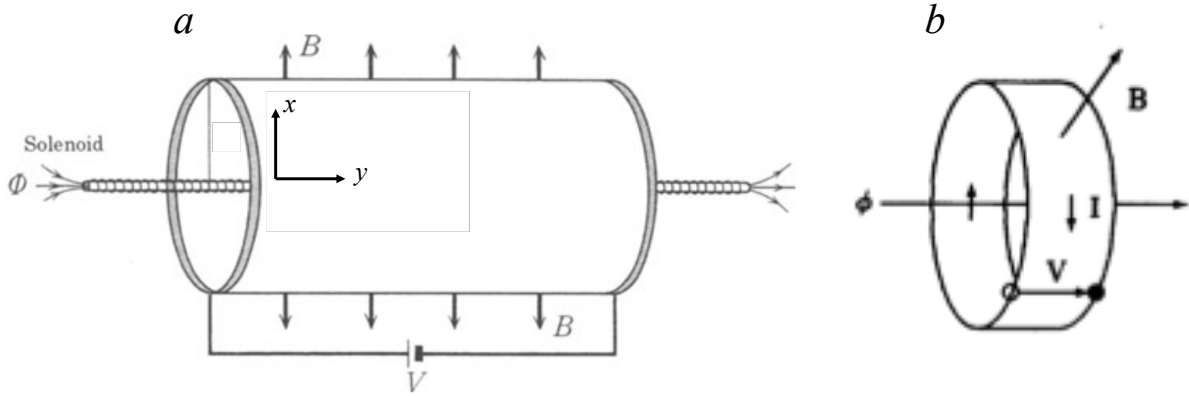


Figure 3.7: a) [Adapted from Ref. 94] and b) [Adapted from Ref. 95] Laughlin's gedankenexperiment.

Including the effects of the transverse electric field  $\mathcal{E} = \frac{V}{w}$ , we regain the Hamiltonian from Equation 3.16 for the inclusion of a transverse electric field in the conductor

$$\hat{H} = \frac{1}{2m} \hat{p}_y^2 + \frac{1}{2} m \left( \frac{eB}{m} \right)^2 (y - y_0)^2 + e\mathcal{E}y + \frac{1}{2} m \frac{\mathcal{E}^2}{B^2}, \quad (3.30)$$

with again,  $y_0 = -\frac{k_x \hbar}{eB} - \frac{e\mathcal{E}}{m\omega_c^2}$ , as the guiding center. Since the values of  $k_x$  are linearly spaced apart in steps  $\frac{2\pi}{L}$ , the spacing of the locations of the wavefunctions in  $y$  is  $\frac{\hbar}{eBL}$ .

In this ring geometry, the states extend all around the loop in  $x$  and connect back to themselves. Therefore any wavefunction solution must satisfy  $\psi(x, y) = \psi(x + L, y)$ . This condition becomes difficult to satisfy if we now apply a magnetic flux  $\Delta\Phi$  through the center of the ring. Even if the flux does not produce a magnetic field at the ribbon, it affects the Hamiltonian by adding a vector potential  $\Delta\mathbf{A} = \frac{\Delta\Phi}{L} \hat{x}$ . The result is a dynamical phase that adds a factor of  $e^{ie\Delta Ax/\hbar}$  to the wavefunctions. This can be treated as a gauge transformation.<sup>94</sup> In

order to satisfy the connectivity condition on the wavefunctions, however, the only allowed transformations are those where  $\Delta\Phi = n\frac{h}{e} = n\Phi_0$ , where  $n$  is an integer.

As a gauge transformation, this operation must map the system back to itself. In terms of the Hamiltonian, this equates to adding an extra term to the vector potential so that  $eBy \rightarrow eB\left(y - \frac{\Delta\Phi}{BL}\right)$ , which only changes the guiding center of the cyclotron orbits. Adding a flux of  $I\Phi_0$  moves the state by  $\frac{h}{eBL}$ , which resolves how to satisfy  $\psi(x, y) = \psi(x + L, y)$ : The addition of a flux quantum simply moves the wavefunction over to the neighboring state in the  $y$  direction.

If  $n$  Landau levels are filled,  $n$  electrons are moving along the ring, and  $n$  electrons are transferred from one edge to the other. The current carried along a single channel is given in terms of the drift velocity calculated in Equation 3.18,  $I = e\frac{\varepsilon}{B}$ . In his 1998 Nobel Prize lecture<sup>95</sup>, Laughlin proves that the current operator is formally the derivative of the Hamiltonian with respect to the vector potential so that by the Hellman-Feynman theorem ( $\frac{dE_\lambda}{d\lambda} = \langle \psi_\lambda | \frac{d\hat{H}_\lambda}{d\lambda} | \psi_\lambda \rangle$ ) the current is just the adiabatic differentiation of the energy with respect to the magnetic flux,  $I = \frac{dE}{d\Phi}$ . If the loop is large enough so that Aharonov-Bohm oscillations are suppressed, and the current changes negligibly as the flux is inserted, the adiabatic derivative can be replaced by a differential,  $I = \frac{\Delta E}{\Delta\Phi}$ . If  $\Delta E = neV$  and  $\Delta\Phi = \frac{h}{e}$ , then the Hall conductance is recovered

$$\frac{I}{V} = \frac{ne^2}{h} = \sigma_{xy}. \quad (3.31)$$

Through this geometry of connected extended states, the Hall conductance was obtained without any details about the edge modes; it was an inevitable feature of the Hamiltonian.



Now consider what happens when disorder is introduced. From general perturbation theory, any perturbation that does not preserve symmetry will break degeneracies. Therefore, we expect that adding a small amount of disorder will split the degeneracy of the Landau levels. Indeed, when disorder is included in the model, the density of states of the Landau levels are no longer delta functions; they are now broadened due to spatial variations. As long as the width of the levels is small relative to their separation in energy, though, we still obtain the same quantized conductance.

If the disorder perturbation is strong enough, however, many of the extended quantum states will become localized, restricted to lie within some region in  $y$ . *Figure 3.8a* is an example of a segment of the ribbon with a local maximum and minimum in the potential due to impurities. Particles in the vicinity of the maximum move anti-clockwise around it, and those near the minimum move clockwise around it. In both cases, the particles are trapped, and cannot contribute to charge transfer. Equipotentials that cross from one side of the sample to the other are uncommon, but they are guaranteed to exist at the edge of the sample. *Figure 3.8c* shows the new density of states in terms of the localized and extended states. States at the far ends of a band are localized, and there may even be a finite density of states between Landau levels. Only states near the center of a band are extended.

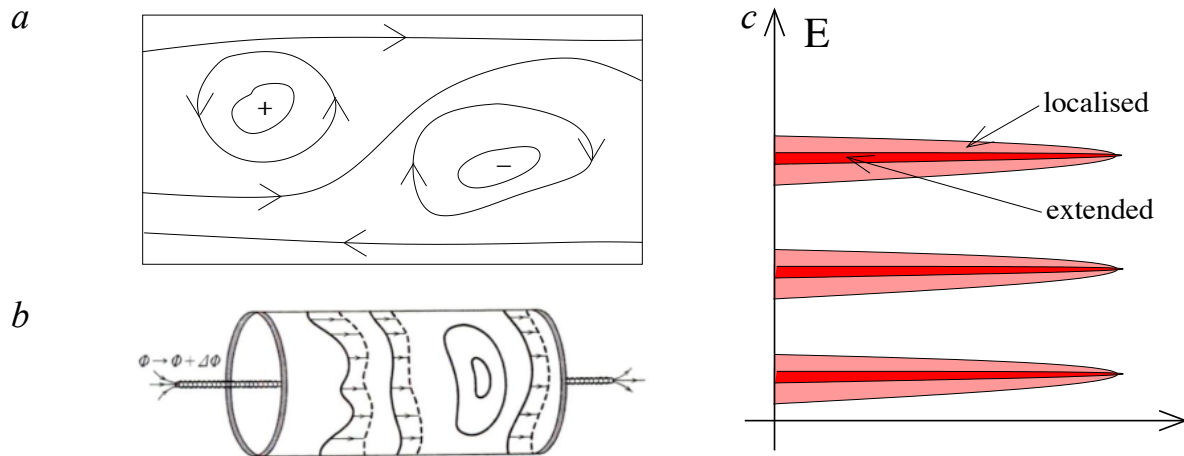


Figure 3.8: a) [Adapted from Ref. 91] A portion of the ribbon from Laughlin’s gedankenexperiment including disorder. Particles circle local maxima and minima in the interior, so that only states on the edge are continuous around the entire ribbon. b) [Adapted from Ref. 94] The shift of wavefunctions along  $y$  will skip over localized states caused by disorder. c) [Adapted from Ref. 91] The density of states become broader in the presence of disorder, leading to a combination of localized states in the interior and extended states around the edges.

Within the localized states, the self-connecting phase condition does not apply, and continuous gauge transformations are allowed. Therefore, when flux is added, the states do not move; instead they only change in phase.<sup>93</sup> Only the extended states are able to move charge from one side of the sample to the other, so the “conveyer-belt” of moving electrons will simply skip over these localized states, and the overall result for conductivity is the same (Figure 3.8b).

To understand how this leads to plateaus in the conductance, consider the case if all extended states in a given Landau level are filled and  $\mathbf{B}$  is decreased for a fixed  $n$ . Each Landau level can accommodate fewer electrons ( $N \sim B$ ), so the Fermi level increases. Unlike the no-

disorder case with delta function density of states, where electrons would jump up to the next Landau level, electrons simply start occupying the localized states instead. The localized states do not contribute to the conductivity, though, so the conductivity is unchanged, resulting in a plateau. This may seem to contradict the original derivation assuming all states in the Landau level contribute to the current, but R.E. Prange<sup>96</sup> found that the current carried by the extended states increases in an exact amount to compensate for the lack of current by the localized states. Therefore the conductivity remains quantized even in the presence of disorder.

### 3.3 Topological Invariants

#### 3.3.1 The TKNN Model of the Quantum Hall Effect

To distinguish the quantum Hall state from ordinary insulators, in 1982, Thouless, Kohmoto, Nightingale, and den Nijs (TKNN) formulated an explanation of the QHE in terms of topology.<sup>85</sup> Their approach applies the Kubo formula, an expression describing the linear response of an observable to a time-dependent perturbation,<sup>34,97</sup> to calculate the Hall conductivity of a 2D system. Gapped band structures of a 2D system can be classified topologically by grouping classes of the Bloch Hamiltonian that can be continuously deformed into one another without closing the band gap. These classes are characterized by a topological invariant called the Chern number.

Begin with the Hamiltonian from Equation 3.13,  $\hat{H} = \frac{1}{2m}\hat{p}_y^2 + \frac{1}{2}m\left(\frac{eB}{m}\right)^2\left(y + \frac{\hbar k_x}{eB}\right)^2 + V(y)$ , and label the many-electron eigenstates as  $|\Psi_n\rangle$ , with energies  $E_n$ . In the ground state,  $|\Psi_0\rangle$ , no current flows until an electric field  $\mathcal{E}_y$  is applied. Then, perturbation theory can be used to determine the resulting current density  $j_x$  and conductivity  $\sigma_{xy}$ . As before, the perturbation to the Hamiltonian is  $e\mathcal{E}y$ , so to first order, the ground state is shifted by

$$|\Psi_0\rangle \rightarrow |\Psi_0\rangle + \sum_{n>0} \frac{\langle\Psi_n|e\mathcal{E}y|\Psi_0\rangle}{E_0-E_n} |\Psi_n\rangle = |\Psi_0\rangle + \sum_{n>0} \frac{|\Psi_n\rangle\langle\Psi_n|}{E_0-E_n} e\mathcal{E}y|\Psi_0\rangle. \quad (3.32)$$

The current density is then

$$\langle j_x \rangle = \left( \langle\Psi_0| + \langle\Psi_0|e\mathcal{E}y \sum_{m>0} \frac{|\Psi_m\rangle\langle\Psi_m|}{E_0-E_m} \right) \hat{j}_x \left( |\Psi_0\rangle + \sum_{n>0} \frac{|\Psi_n\rangle\langle\Psi_n|}{E_0-E_n} e\mathcal{E}y|\Psi_0\rangle \right). \quad (3.33)$$

Dropping  $\mathcal{E}^2$  terms to keep the calculation in first order, and since  $\langle\Psi_0|\hat{j}_x|\Psi_0\rangle = 0$  (no current flows in the unperturbed ground state),

$$\langle j_x \rangle = e\mathcal{E} \sum_{n>0} \frac{\langle\Psi_0|y|\Psi_n\rangle\langle\Psi_n|\hat{j}_x|\Psi_0\rangle + \langle\Psi_0|\hat{j}_x|\Psi_n\rangle\langle\Psi_n|y|\Psi_0\rangle}{E_0-E_n}. \quad (3.34)$$

We can then put everything in terms of the current density by substituting the definition  $\hat{j}_y =$

$\frac{e}{m} \hat{p}_y$  into the commutation relation  $[\hat{H}, y] = -\frac{i\hbar}{m} \hat{p}_y$ , and

$$\langle\Psi_n|[\hat{H}, y]|\Psi_0\rangle = \langle\Psi_n|(E_n - E_0)y|\Psi_0\rangle = \langle\Psi_n|-\frac{i\hbar}{e}\hat{j}_y|\Psi_0\rangle. \quad (3.35)$$

Substituting Equation 3.35 back into Equation 3.34 and dividing through by  $\mathcal{E}$ , we have the Hall conductivity in the form of a current-current correlation function<sup>98</sup>

$$\sigma_{xy} = -i\hbar \sum_{n>0} \frac{\langle\Psi_0|\hat{j}_x|\Psi_n\rangle\langle\Psi_n|\hat{j}_x|\Psi_0\rangle + \langle\Psi_0|\hat{j}_x|\Psi_n\rangle\langle\Psi_n|\hat{j}_x|\Psi_0\rangle}{(E_0-E_n)^2}. \quad (3.36)$$

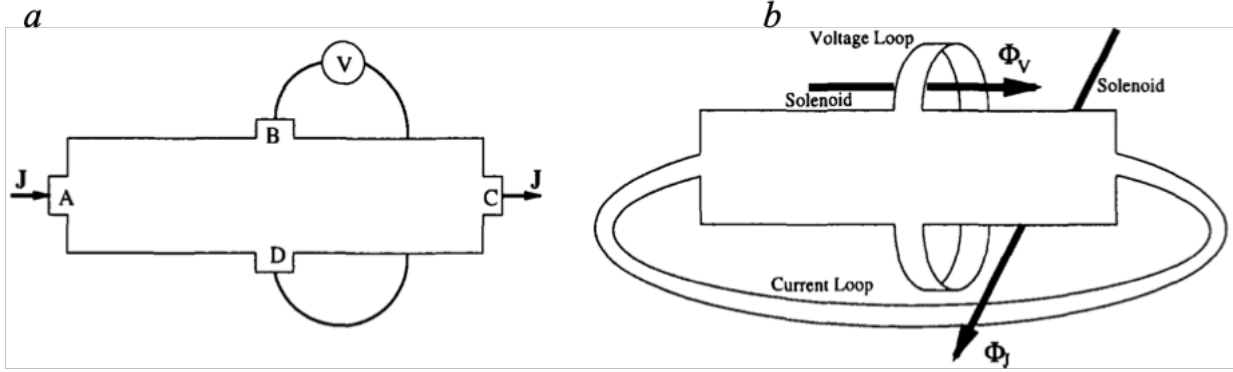


Figure 3.9: [Adapted from Ref. 98] Quantum Hall systems with realistic boundary conditions a) A simple Hall bar with source and drain on left and right and voltage leads on the top and bottom. b) An equivalent geometry where current and voltage leads are replaced by loops of the same material as the Hall bar with flux threaded through each.

From here, Avron and Seiler related the TKNN formulation to Laughlin's gedankenexperiment for a real Hall bar geometry<sup>99,100</sup> such as that in Figure 3.9a.<sup>98</sup> Figure 3.9b is an equivalent situation, in which the current and voltage elements are replaced with loops of the same material as the bar. Two different fluxes are threaded through the loops,  $\Phi_J$  through the current loop and  $\Phi_V$  through the voltage loop.<sup>101</sup> Now, the application of the electric field  $\mathcal{E}$  is achieved by uniformly increasing  $\Phi_V$  to produce an electromotive force, and the resulting current is seen by observing  $\Phi_J$ . Again, we can use the relation  $I = \frac{\Delta E}{\Delta \Phi}$ , to find the current density

operators  $\hat{j}_x$  and  $\hat{j}_y$  by  $\hat{j}_x = \frac{\partial \hat{H}}{\partial \Phi_J}$  and  $\hat{j}_y = \frac{\partial \hat{H}}{\partial \Phi_V}$ . By first order perturbation theory we have

$$\left\langle \frac{\partial \Psi_0}{\partial \Phi_J} \right\rangle = \sum_{n>0} \frac{\langle \Psi_n | \frac{\partial \hat{H}}{\partial \Phi_J} | \Psi_0 \rangle}{E_0 - E_n} |\Psi_n\rangle \text{ and } \left\langle \frac{\partial \Psi_0}{\partial \Phi_V} \right\rangle = \sum_{n>0} \frac{\langle \Psi_n | \frac{\partial \hat{H}}{\partial \Phi_V} | \Psi_0 \rangle}{E_0 - E_n} |\Psi_n\rangle. \quad (3.37)$$

Plugging back into Equation 3.36, and using the identity operator  $\sum_n |\Psi_n\rangle\langle\Psi_n|$ , the Hall conductivity becomes

$$\sigma_{xy} = i\hbar \left( \left\langle \frac{\partial\Psi_0}{\partial\Phi_V} \left| \frac{\partial\Psi_0}{\partial\Phi_J} \right\rangle - \left\langle \frac{\partial\Psi_0}{\partial\Phi_J} \left| \frac{\partial\Psi_0}{\partial\Phi_V} \right\rangle \right). \quad (3.38)$$

Just as in Laughlin's original gedankenexperiment, the system is invariant to changes in either flux by  $\frac{h}{e}$ . This is equivalent to saying  $\Phi_J$  and  $\Phi_V$  are periodic, and  $(\Phi_J, \Phi_V)$  space is shaped like a toroid with surface area  $\frac{h^2}{e^2}$ . Integrating over all  $(\Phi_J, \Phi_V)$  space, we find the flux-averaged value for the Hall conductance

$$\sigma_{xy} = i\hbar \frac{e^2}{h^2} \int_0^{h/e} d\Phi_J \int_0^{h/e} d\Phi_V \left( \left\langle \frac{\partial\Psi_0}{\partial\Phi_V} \left| \frac{\partial\Psi_0}{\partial\Phi_J} \right\rangle - \left\langle \frac{\partial\Psi_0}{\partial\Phi_J} \left| \frac{\partial\Psi_0}{\partial\Phi_V} \right\rangle \right). \quad (3.39)$$

We can rewrite the integrand as the curl of the vector

$$\frac{1}{2} \left( \left\langle \frac{\partial\Psi_0}{\partial\Phi_J} \left| \Psi_0 \right\rangle - \left\langle \Psi_0 \left| \frac{\partial\Psi_0}{\partial\Phi_J} \right\rangle, \left\langle \frac{\partial\Psi_0}{\partial\Phi_V} \left| \Psi_0 \right\rangle - \left\langle \Psi_0 \left| \frac{\partial\Psi_0}{\partial\Phi_V} \right\rangle \right) = -i\text{Im} \left( \left\langle \Psi_0 \left| \frac{\partial\Psi_0}{\partial\Phi_J} \right\rangle, \left\langle \Psi_0 \left| \frac{\partial\Psi_0}{\partial\Phi_V} \right\rangle \right), \quad (3.40)$$

so

$$\sigma_{xy} = \frac{e^2}{2\pi h} \int d^2\Phi \text{Im}(\nabla_\Phi \times \langle \Psi_0 | \nabla_\Phi \Psi_0 \rangle). \quad (3.41)$$

Applying Stoke's theorem, which relates a surface integral to an integral over a curve by

$\iint \nabla \times \mathbf{F} \cdot d\mathbf{S} = \oint \mathbf{F} \cdot d\mathbf{s}$ , we have

$$\sigma_{xy} = \frac{e^2}{2\pi h} \text{Im} \oint (\langle \Psi_0 | \nabla_\Phi \Psi_0 \rangle \cdot d\Phi). \quad (3.42)$$

The term  $\text{Im} \oint (\langle \Psi_0 | \nabla_\Phi \Psi_0 \rangle \cdot d\Phi)$  is the negative of Berry's phase for a path around the border of  $\Phi$  space. Berry's phase<sup>102</sup> appears in cyclic adiabatic evolutions. If we start with a Hamiltonian based on a set of parameters, then vary the parameters, but ultimately bring them back to their starting values, we have traced out a closed path in the space of those parameters. The adiabatic theorem states that if we started out in the ground state, we will end up back at the ground state, only with a change in phase

$$|\psi_0\rangle \rightarrow e^{i\gamma} |\psi_0\rangle \quad (3.43)$$

This change in phase becomes important in considering the interaction between one state that went through the adiabatic change and another that did not. The inclusion of phase means that now, interference between the two states, whether constructively or destructively, could have physical meaning. There are two contributions to the phase  $e^{i\gamma}$ : a dynamical phase  $e^{-iEt/\hbar}$ , present for every eigenstate, even the ones that do not go through the change; and Berry's phase, the phase gained after taking a closed path in parameter space. Berry's phase is independent of the time taken to undergo the change in parameters, but does depend on the path taken through parameter space.

In Equation 3.42, the path around the border of  $\Phi$  space must be a single-valued and therefore an integer multiple of  $2\pi$  in order to map around the loop back onto the original values. Thus, we recover the Hall conductance

$$\sigma_{xy} = \frac{e^2}{2\pi h} (2\pi n) = \frac{ne^2}{h}. \quad (3.44)$$

### 3.3.2 The Chern Number

The important result of the TKNN formulation is that the quantized Hall conductance is a fundamental property of the system. The value of  $n$  in Equation 3.44 is invariant to smooth changes of the Hamiltonian, and therefore defines an equivalence class for gapped Hamiltonians. Thus,  $n$  is a topological invariant.<sup>100</sup>

As discussed in *Section 3.1*, two-dimensional objects are classified in topology by their number of holes because two surfaces with the same number of holes can be smoothly deformed into one another. The number of holes is therefore the topological invariant for the class. The Chern number, or Chern invariant, was introduced by mathematician Shiing-Shen Chern in 1946

to characterize classes of complex vector bundles.<sup>103</sup> Chern argued that the genus of a surface can be found by integrating the local curvature over its entire area. In terms of the Hall conductivity in the TKNN formulation, this is similar to integrating Berry's curvature over the area of parameter space (Equation 3.41) or integrating Berry's phase around the area of parameter space (Equation 3.42).

In real space, the Chern number defines a class of gapped Hamiltonians, and a system cannot smoothly deform between regions of different Chern numbers. In order to move across the boundary, the gap with one value of  $n$  must close and then open as a gap with a different value of  $n$ . At the boundaries, though, gapless conducting states must exist. This leads to a unique feature of these states called bulk-boundary correspondence, which relates the topological structure of the bulk to the presence of gapless chiral boundary modes. Bulk-boundary correspondence states that the number of chiral edge modes is equal to the difference in the Chern number across the interface<sup>104</sup>

$$N_{edge} = \Delta n. \quad (3.45)$$

This is readily seen at the edge of a QH system, where the transition from Chern number (or filling factor)  $n$  to vacuum (Chern number of 0) leads to  $n$  1D chiral edge modes. This statement holds for any shape or local details of the Hall bar, demonstrating how the topological nature of the QHE relates to real space and verifies its robustness.

### ***3.4 Topological Insulators: The Quantum Spin Hall Effect***

While the QHE is a phenomenon of rich physics, the required low temperatures and high magnetic fields are experimentally demanding. It was therefore desirable to find equivalent



systems with nontrivial Chern numbers or other topological invariants. The class of materials that fulfill these criteria are known as topological insulators.

### 3.4.1 Time Reversal Symmetry

One property that is crucial to achieving topological invariants without high magnetic fields is time-reversal symmetry (TRS). Time reversal symmetry essentially requires that a system be identical under the reverse flow of time. Let  $\psi(\mathbf{x}, t) = e^{-\frac{iEt}{\hbar}} \psi(\mathbf{x}, 0)$  be the time evolution of the wavefunction  $\psi(\mathbf{x}, 0)$ , where  $e^{-\frac{iEt}{\hbar}}$  is the time evolution factor. Then  $\psi(\mathbf{x}, -t)$  is the time-reversed conjugate of  $\psi(\mathbf{x}, t)$ , found by running the system with time flowing backwards, or by reversing all the velocities or momenta in the system.<sup>105</sup> In the absence of spin, time reversal, denoted by the operator  $\hat{T}: t \rightarrow -t$ , effectively turns the wavefunction to its complex conjugate

$$\hat{T}\psi(\mathbf{x}, t) = \psi(\mathbf{x}, -t) = \psi^*(\mathbf{x}, t). \quad (3.46)$$

When accounting for spin,  $\hat{T} = \hat{K}\sigma_y$ , where  $\hat{K}$  is the complex conjugation operator, and  $\sigma_y$  is the Pauli spin operator

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (3.47)$$

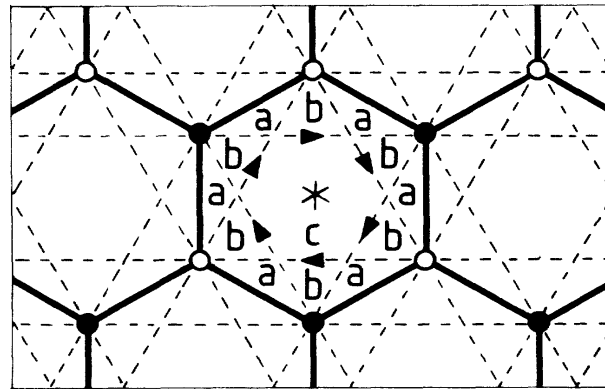
It can be shown that the Hamiltonian is invariant under  $\hat{T}$  whether spin is included or not.<sup>105</sup> The time reversal operator also commutes with the Hamiltonian as a consequence of energy conservation.

### 3.4.2 The Haldane Model

In 1988, F.D.M. Haldane worked out a model for achieving quantized Hall conductance, in other words a nonzero Chern number, in the absence of a magnetic field.<sup>106</sup> He formulated his model using graphene, a 2D system of carbon atoms in a hexagonal lattice (*Figure 3.10*). The hexagonal lattice of graphene is often broken down into two triangular sublattices,  $A$  and  $B$ . The resulting band structure has two special points in reciprocal space,  $\mathbf{K}$  and  $\mathbf{K}' = -\mathbf{K}$ , located at the corners of the Brillouin zone. At  $\mathbf{K}$  and  $\mathbf{K}'$ , the conduction and valence bands touch, and near these points, the electronic dispersion resembles the linear dispersion of massless relativistic particles, described by a 2D massless Dirac Hamiltonian<sup>107,108</sup>

$$\hat{H} = \hbar v_F \mathbf{q} \cdot \boldsymbol{\sigma}, \quad (3.48)$$

where  $v_F$  is the Fermi velocity,  $\boldsymbol{\sigma}$  are the Pauli matrices in  $x$  and  $y$ , and  $\mathbf{q} = \mathbf{k} - \mathbf{K}$  is the momentum relative to the  $\mathbf{K}$  or  $\mathbf{K}'$  point.



*Figure 3.10: [Adapted from Ref. 106] Hexagonal lattice of graphene. The  $A$  and  $B$  sublattices are denoted by the closed and open circles, respectively.*

The degeneracy of the Dirac points is protected by both inversion and time-reversal symmetries. Breaking the symmetries would allow the degeneracy to be lifted. For example,

inversion symmetry is broken if the two atoms in the unit cell are inequivalent. If we change the relative energies of the two sublattices by adding a small momentum term  $\hbar v_F q_z$ , then near  $\mathbf{K}$ , we regain a massive term in  $z$ , so that the full Hamiltonian becomes

$$\hat{H} = \hbar v_F \mathbf{q} \cdot \boldsymbol{\sigma} + m \sigma_z, \quad (3.49)$$

where  $m = \hbar v_F K_z$ . This changes the spectrum from being completely linear, gapless, and therefore massless, to having an effective mass at  $K$ . Furthermore the dispersion  $E(\mathbf{q}) = \pm \sqrt{|\hbar v_F \mathbf{q}|^2 + m^2}$  now has an energy gap of  $2|m|$ . Time-reversal symmetry requires the Dirac point at  $\mathbf{K}'$  to also have a mass with the same magnitude and sign  $m' = m$ . From a topological perspective, when integrating the Berry curve over the Brillouin zone, contributions from  $\mathbf{K}$  and  $\mathbf{K}'$  cancel, and the result is an ordinary insulator.<sup>109</sup>

Instead of breaking inversion symmetry, Haldane envisioned lifting the degeneracy by breaking the time-reversal symmetry.<sup>106</sup> He realized it was possible to break time-reversal symmetry with a magnetic field that is zero on average, but respects the full symmetry of the lattice, so that the masses at  $\mathbf{K}$  and  $\mathbf{K}'$  are not equal. In this case, the system is not a trivial insulator, but rather a quantum Hall state with quantized conductance.

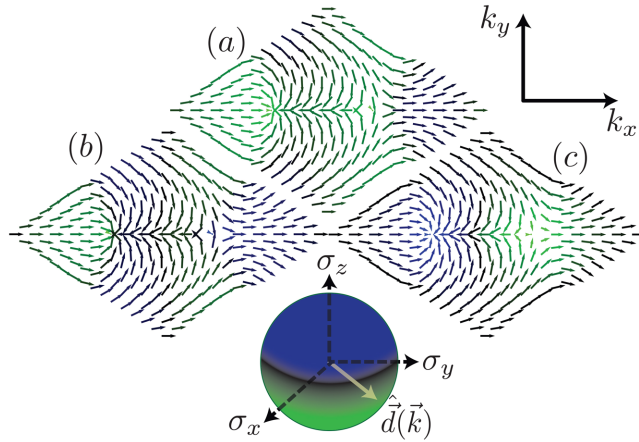


Figure 3.11: [Adapted from Ref. 110] Due to the periodic conditions of the Brillouin zone, the unit vector must wrap around the sphere an integer number of times.

The Chern number can distinguish between similar or seemingly-identical band structures by calculating the integral of the Berry curvature in the Brillouin zone. Due to the periodic boundary conditions of the Brillouin zone, the unit vector  $d\mathbf{k}$  must wrap around the unit sphere an integer number of times, and the number of trips around gives the Chern number (Figure 3.11).<sup>110</sup> When  $m = m' = 0$ , momentum is confined to the  $x$ - $y$  plane, with a unit and opposite-direction winding around each Dirac point. For small but finite  $m, m'$ , momentum is finite everywhere and  $\mathbf{q}(\mathbf{K})$  hits  $z_{max}$  or  $z_{min}$  (the north or south poles of the unit sphere) depending on the sign of  $m$ . Each Dirac point, then, contributes  $\pm \frac{e^2}{2h}$  to  $\sigma_{xy}$ .<sup>104</sup> In the case of  $m = m'$ , the contributions cancel. When  $m = -m'$ , however, they add, resulting in a nonzero, quantized Chern number and full quantum Hall edge state behavior without a net magnetic field.

Haldane concluded his paper by emphasizing his proof was simply a toy model, “unlikely to be directly physically realizable,”<sup>106</sup> but revived interest in his ideas have led to proposals for its realization in buckled hexagonal lattices such as silicene<sup>110</sup> and iron-based honeycomb

ferromagnetic insulators.<sup>111</sup> In 2014, Jotzu *et al.* were able to experimentally demonstrate a Haldane insulator in ultracold fermionic atoms in a periodically modulated optical honeycomb lattice.<sup>112</sup>

### 3.4.3 $Z_2$ Topological Invariant

The Hall conductivity is odd under time-reversal symmetry, and therefore the topologically nontrivial situation described above can only be obtained by breaking TRS. The spin-orbit interaction, however, allows for a different topological class of insulating band structures that preserves TRS.<sup>113,114</sup> When graphene was experimentally discovered in 2004,<sup>115,116</sup> interest in the Haldane model was renewed. Building on theoretical models of the spin Hall effect,<sup>117–119</sup> the quantum spin Hall effect (QSHE) (*Section 3.4.4*) was discovered,<sup>113,114,120,121</sup> which can be modeled as two complementary Haldane models.<sup>86,104</sup>

This new topological class is largely dependent on spin-orbit coupling. The spin-orbit interaction is a relativistic effect connecting a particle's spin with its motion inside a potential. An electron moving with a velocity with a component perpendicular to an electric field  $\mathbf{E}$  will, in its rest frame, experience a magnetic field,  $\mathbf{B} = -\frac{v}{c^2} \times \mathbf{E}$ , that couples to its spin by the Zeeman effect. The perturbation to the Hamiltonian is given by  $U = -\frac{e\hbar}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} = -\frac{e\hbar}{2m^2 c^2} \boldsymbol{\sigma} \cdot (\mathbf{E} \times \mathbf{p})$ , where  $\boldsymbol{\sigma}$  is the magnetic moment of the electron due to its spin and  $\mathbf{p}$  is its momentum. In the laboratory reference frame, this is seen as a degeneracy splitting between the two spin directions proportional to the momentum. For spin- $\frac{1}{2}$  particles, the TRS operator is antiunitary,  $\hat{T}^2 = -1$ , so spin-orbit coupling respects TRS; the TRS operator flips both spin and momentum, resulting in the same spin-orbit energy.

This leads to an important constraint on the edge states of a time-reversal symmetric 2D system known as Kramers theorem. Kramers theorem states that for every energy eigenstate of a time-reversal symmetric system with half-integer total spin, there must be at least one other eigenstate of the same energy. Without spin-orbit interactions, this is just the degeneracy between up and down spins, but when spin-orbit interactions are included, the consequences are more complex. TRS-invariant Bloch Hamiltonians must satisfy the condition

$$\hat{T}\hat{H}(\mathbf{k})\hat{T}^{-1} = \hat{H}(-\mathbf{k}). \quad (3.50)$$

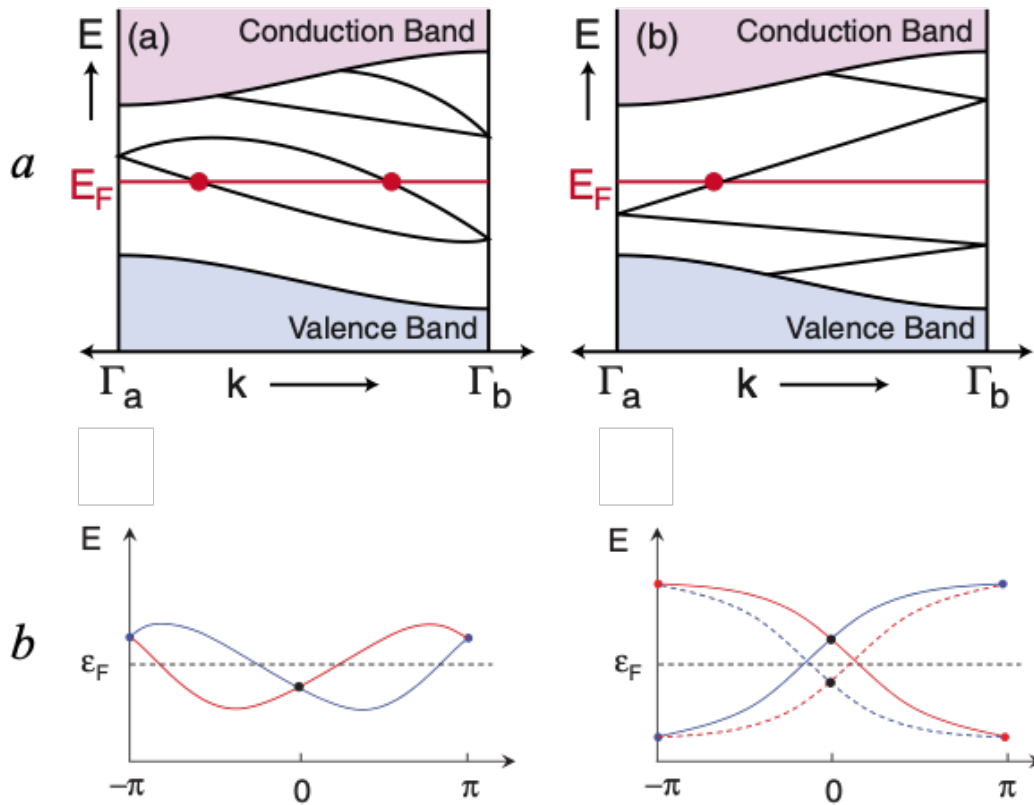


Figure 3.12: a) [Adapted from Ref. 104] Two options for electronic dispersion between two boundary Kramers degenerate points. On the left, an even number of surface states cross the Fermi level, resulting in an ordinary insulator. On the right, an odd number of states cross the Fermi level, which leads to topologically protected metallic boundary states. b) [Adapted from Ref. 121] Another example of the energy dispersions. On the left is a 1D TR invariant system. Due to the Kramers degeneracy at  $0$  and  $\pm\pi$ , the energy spectrum intersects the Fermi level  $4n$  times. On the right, the helical states on one boundary of a QSH system are represented by the solid lines. At  $k = 0$ , the edge states are Kramers partners, while at  $k = \pm\pi$ , they merge into the bulk and pair with the edge states at the other boundary (dashed lines). The red and blue lines represent the two partners of a Kramers pair.

We can define new topological equivalence classes of Hamiltonians that satisfy this constraint in terms of those that can be smoothly deformed without closing the bulk energy gap. The TKNN invariant is  $n = 0$ , but we now have a new invariant with two possible values,  $\nu = 0$  or  $1$ .<sup>113</sup> *Figure 3.12* plots the two distinct possibilities for edge states for a TRS 2D insulator as a function of crystal momentum  $k$  along the edge. By TRS, in *Figure 3.12a*, the displayed half of the Brillouin zone  $0 < k_x < \frac{\pi}{a}$  is the mirror image of the other half,  $-\frac{\pi}{a} < k_x < 0$ , with the accompanying spin-flip. The shaded regions at the top and bottom are the conduction and valence bands of the bulk, separated by an energy gap. The details of the Hamiltonian dictate whether or not there are states bound to the edge within the gap. If there are states bound to the edges, then by Kramers theorem they must be twofold degenerate at the TRS invariant momenta  $k_x = 0, \pm \frac{\pi}{a}$ . In other words, at exactly  $k_x = 0$  and at the boundary of the Brillouin zone  $k_x = \pm \frac{\pi}{a}$ , the time-reversed states are at the same momentum. Away from these points, spin-orbit coupling splits the degeneracy, but there are two ways by which the Kramers pair of states can recombine. In *Figure 3.12, left*, they connect as pairs; in other words, the Fermi energy always intersects an even number of bands. In *Figure 3.12, right*, there is an odd number of crossings.

In the case of an even number intersections it is possible to smoothly deform the Hamiltonian to push the edge states out of the gap. The edge states can be eliminated by forcing all bound states out of the gap. If there is an odd number of intersections, however, there will always be at least one Kramers pair in the gap. Therefore, the presence of edge states is a topological property,<sup>113,114</sup> and we can define an equivalence class of Hamiltonians. In the same way that the Chern number groups Hamiltonians by an integer ( $Z$ ), these two new classes are defined by the  $Z_2$  topological invariant. The  $Z_2$  topological invariant,  $\nu$ , can be either 0 or 1, and is defined as the number of Kramers pairs at the Fermi energy in the gap modulo 2.



$$\nu = N_K \text{mod}(2). \quad (3.51)$$

There are multiple formulations of  $\nu$  based on bulk properties, including spin-charge separation,<sup>122</sup> symmetry properties of the Brillouin zone,<sup>123</sup> and more.<sup>113,124–129</sup> Determining  $\nu$  becomes simpler if the crystal has extra symmetry, such as when a 2D system conserves the perpendicular spin  $S_z$ . In this case, the up and down spins have independent Chern numbers,  $n_\uparrow$  and  $n_\downarrow$ , and TRS requires,  $n_\uparrow + n_\downarrow = 0$ . The difference, however,  $\frac{n_\uparrow - n_\downarrow}{2} = n_\sigma$  defines a quantized spin Hall conductivity,<sup>130</sup> and the  $Z_2$  invariant is given by

$$\nu = n_\sigma \text{mod} 2. \quad (3.52)$$

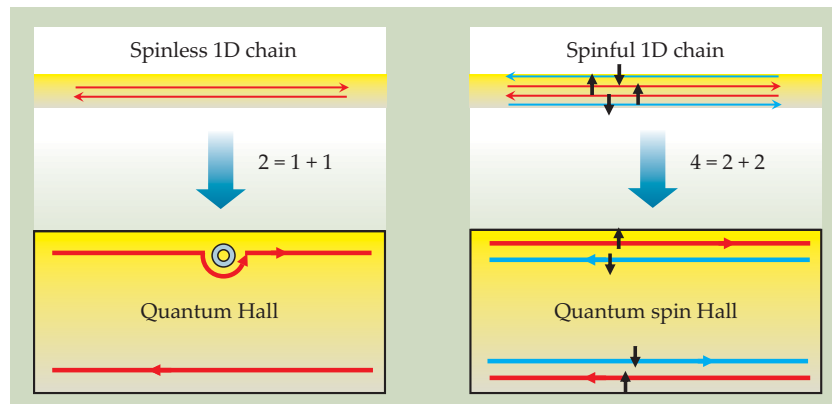
The value of the  $Z_2$  invariant in this case is robust even when the addition of terms that do not conserve  $S_z$  are incorporated, which causes  $n_\uparrow$  and  $n_\downarrow$  to lose their significance as independent numbers.

### 3.4.4 The Quantum Spin Hall Effect

The quantum spin Hall effect (QSHE) is an experimentally demonstrated example of nontrivial  $\nu = 1$  topological order. In 2005, Kane and Mele<sup>113,114</sup> developed the formulation for the QSH state as the combination of two copies of the Haldane model. In the QSHE, there is a single Kramers pair of states at the edge of the system with opposite momenta and spins.<sup>113,120</sup> The spin up electrons exhibit the chiral integer QHE, while the spin down electrons exhibit an anti-chiral integer QHE. In effect, this is a double quantum Hall effect, where each edge is partnered with its opposite spin to respect time reversal symmetry. Edge modes in the QSHE are called helical since their spin is correlated to their direction of motion.<sup>131</sup>

*Figure 3.13* is a side by side comparison of the chiral quantum Hall effect and the helical quantum spin Hall effect. In both cases, the edge modes are topologically protected. The

difference, however, is that while the QHE has only clockwise or counterclockwise conduction, the QSHE has both right- and left-moving states on each side, with the oppositely moving states having opposite spin. In the simple geometry of *Figure 3.13*, the two-terminal conduction for the QHE setup is  $\frac{ne^2}{h}$  for integer  $n$ , but for the QSHE it is always  $\frac{2e^2}{h}$  due to the doubling of both right- and left-moving states.



*Figure 3.13: [Adapted from Ref. 86] Comparison of the QHE (left) and the QSHE (right).*

Though an edge in the QSHE has both forward and backward movers, scattering by nonmagnetic impurities is forbidden. A quick analogy suggested by Qi and Zhang<sup>132</sup> is instructive for understanding the QSHE's invulnerability to nonmagnetic impurities. Consider a lens coated with some antireflection material (*Figure 3.14a*). Light reflected from the top and bottom surfaces interfere destructively, so that there is zero net reflection and therefore perfect transmission. Similarly, an electron can be reflected by an impurity and the different possible paths for reflection interfere. As *Figure 3.14b* shows, an electron in a QSH edge state can take either a clockwise or a counterclockwise turn about a nonmagnetic impurity. During this turn, the electron's spin rotates by an angle of  $\pi$  or  $-\pi$ . Thus, the two paths, related by TRS, differ by a total angle of  $2\pi$ , or a full rotation of the electron spin. A quantum mechanical property of a

spin- $\frac{1}{2}$  particle is that its wavefunction gains a negative sign when it has undergone a full  $2\pi$  rotation. Therefore, the two backscattering paths interfere destructively, and perfect transmission is protected. If an impurity has some magnetic moment, however, it can break the TRS between the two edge modes, and the two wavefunctions no longer interfere destructively.

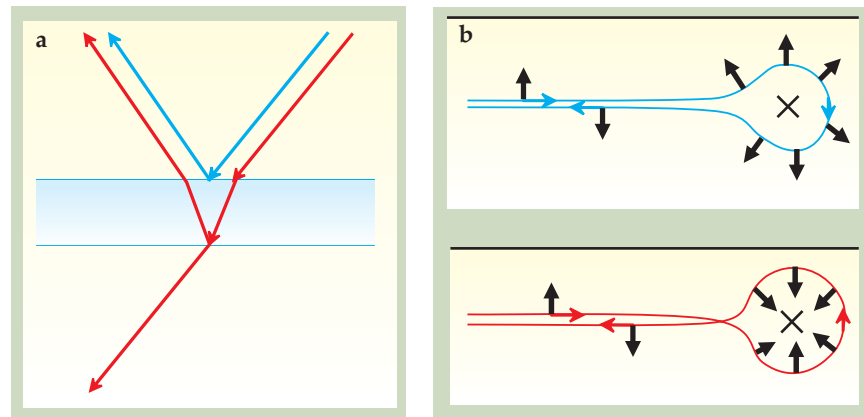


Figure 3.14: [Adapted from Ref. 86] a) Photons traveling through a lens with antireflective coating. Light reflecting from the top and bottom surfaces interfere destructively. b) Likewise, the possible paths for an electron scattering off an impurity interfere destructively such that ballistic transmission is protected.

This picture only applies when there is an odd number of QSH edge state pairs, primarily only a single pair.<sup>113,131,133</sup> If there were two forward-moving and two backward-moving states on a single edge, it would be possible for an electron to be scattered from a forward-moving state to a backward-moving state without flipping its spin, which would destroy the perfect destructive interference, and dissipation could occur. Therefore, in order for the QSH state to be robust, there must be an odd number of forward- (backward-)movers. This even/odd effect is the primary reason a QSH insulator is characterized by a  $Z_2$  topological invariant.

If extra probes for measuring voltage are added to the Hall bars in *Figure 3.13*, the QHE chiral edges and the QSHE helical edges will have distinct responses. Chiral modes propagate voltages perfectly until they reach a current source or drain (*Section 3.2.1*). In a helical mode, though, at every probe, there are left- and right-moving states entering and exiting at both the left and right sides. When an electron enters the probe, it loses its spin coherence, and therefore has an equal probability of leaving in either direction. The result is two-way conduction between each probe. By the probe numbering convention in *Figure 3.2*,  $\tau_{i,i+1} = \tau_{i+1,i} = 1$ , with all other elements zero. The new voltage relations in the QSHE are

$$V_2 = \frac{2}{3}V_1 \text{ and } V_3 = \frac{1}{3}V_1. \quad (3.53)$$

In the QSHE, voltage drops by in equal amounts as we move along the probes in the direction of current flow. Generalizing to any device geometry, the resistance between adjacent leads is always  $\frac{h}{e^2}$ .

### ***3.5 Two-Dimensional Topological Insulators***

Though the search for QSHE systems, or 2D topological insulators (2DTIs) began with graphene,<sup>114</sup> the spin-orbit coupling of carbon orbitals is too weak to produce a significant gap.<sup>134</sup> Thus began the search for QSHE candidate systems among heavier elements, which typically have stronger spin-orbit coupling. In 2006, Bernevig, Hughes, and Zhang (BHZ) proposed the first generalized mechanism for finding topological insulators, and specifically suggested mercury telluride quantum wells (QWs), nanoscale layers sandwiched between thicker layers of other materials, would exhibit QSH states.<sup>120</sup> They argued that HgTe QWs are topological insulators above a critical thickness  $d_c$ , and the general mechanism for the QSHE in these

systems is band inversion, in which the usual valence band lies above the usual conduction band due to spin-orbit coupling.<sup>135</sup>

BHZ proposed sandwiching the HgTe QWs between layers of CdTe (*Figure 3.15a*). *Figure 3.15b* shows the band structure of both compounds. In most semiconductors, including CdTe (*Figure 3.15b right*), the conduction band above the Fermi energy is composed of electrons in  $s$  orbitals (the  $\Gamma_6$  band), and the valence band is composed of electrons in  $p$  orbitals (the  $\Gamma_7$  and  $\Gamma_8$  bands). HgTe, however, has an inverted band structure (*Figure 3.15b left*) due to large spin-orbit coupling, and the roles of the two bands are reversed. When HgTe is sandwiched between CdTe, the band positions are shifted, pushing the  $s$ -type  $\Gamma_6$  up in energy and pushing  $p$ -type  $\Gamma_8$  down. If the QWs are thin, this will force the bands back into a conventional alignment, like in CdTe (*Figure 3.15c left*). If the QW thickness is above the critical thickness,  $d_c$ , however, the band structure will stay inverted (*Figure 3.15c right*).

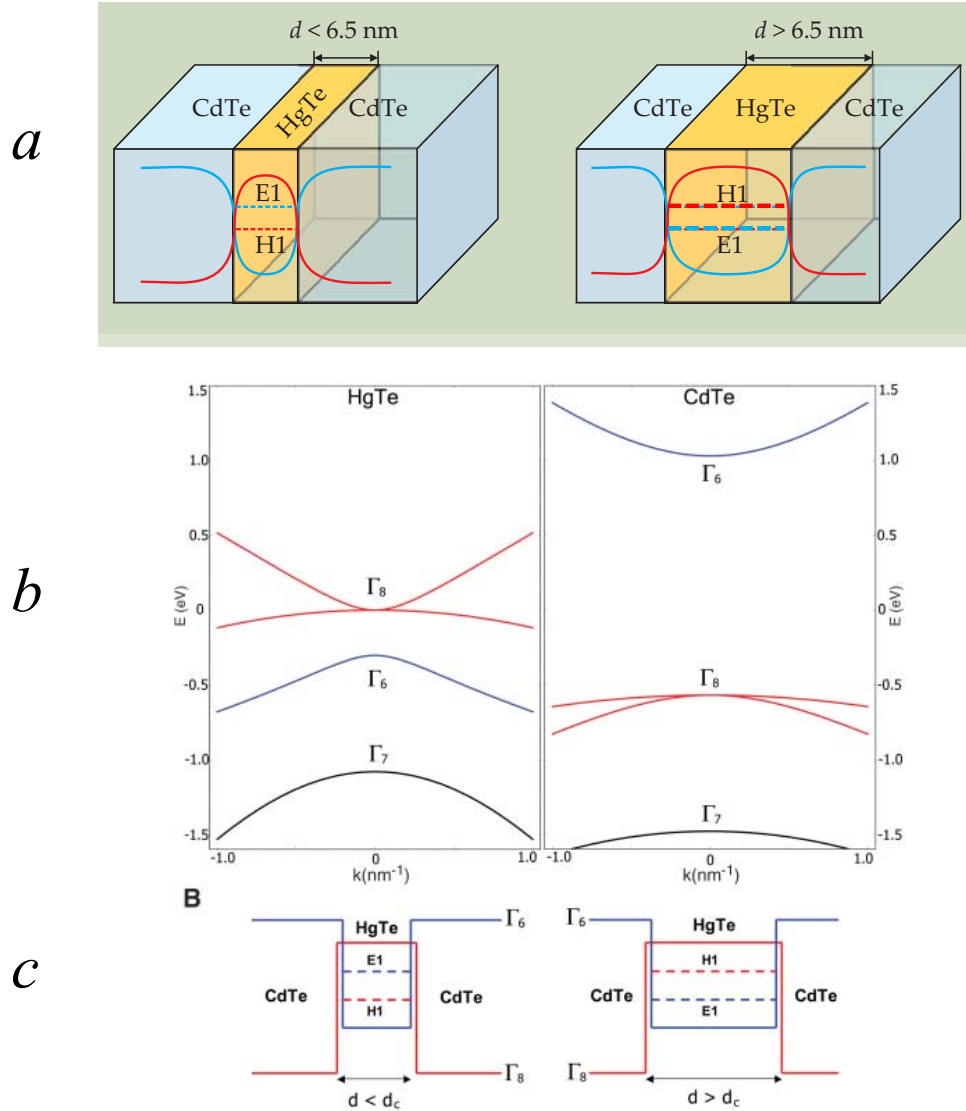


Figure 3.15: a) [Adapted from Ref. 86] HgTe QWs sandwiched between CdTe. b) [Adapted from Ref. 120] Relevant bands for HgTe and CdTe. CdTe (right) has a conventional semiconductor band structure, with p-type valence band  $\Gamma_8$  and s-type conduction band  $\Gamma_6$ . In HgTe (right) the band positions are inverted c) Band positions in the quantum well structure. When the QW is thin (left) the band positions are forced back into conventional ordering by confinement. For a sufficiently thick well (right), the inverted gap structure is restored.

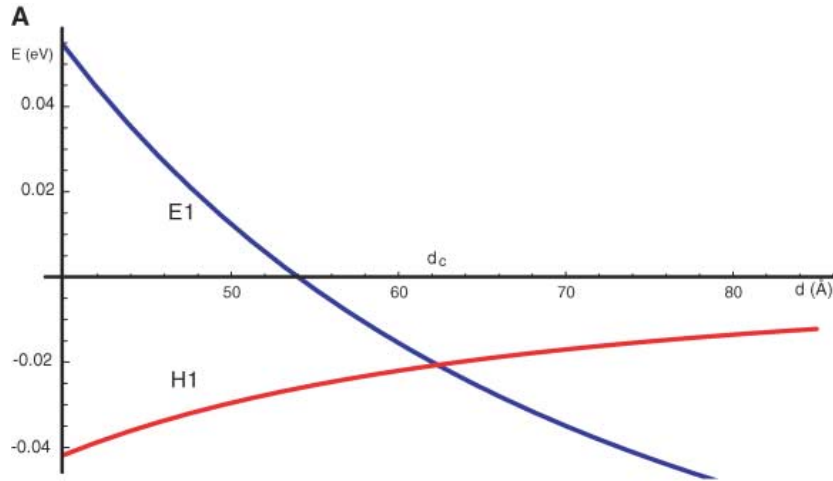


Figure 3.16: [Adapted from Ref. 120] Energy of the E1 and H1 bands as a function of well thickness. The point at which the bands cross is considered a topological phase transition from the trivial to the nontrivial  $Z_2$  topology.

Figure 3.16 plots the energy positions of the bands ( $\Gamma_8 = H_1$ ,  $\Gamma_6 = E_1$ ) as a function of well thickness. BHZ interpreted the point at which the electron and hole bands cross (about 6.3 nm in Figure 3.16) as the topological phase transition from trivial and nontrivial  $Z_2$  topology.<sup>113,136</sup> The inverted gap must close before reopening again at the sample boundary. The QSH state in HgTe can be understood by a simple model of the electron-like Kramers pair of  $E_1$  states and the hole-like Kramers pair of  $H_1$  states.<sup>120</sup> Using the ordered set of four six-component (for six bands) basis vectors  $\{|E_{1+}\rangle, |H_{1+}\rangle, |E_{1-}\rangle, |H_{1-}\rangle\}$ , the BHZ Hamiltonian is given by<sup>120</sup>

$$\hat{H} = \begin{pmatrix} h(\mathbf{k}) & 0 \\ 0 & h^*(-\mathbf{k}) \end{pmatrix} \quad (3.54)$$

with

$$h(\mathbf{k}) = \epsilon(\mathbf{k}) + \mathbf{d} \cdot \boldsymbol{\sigma}, \quad (3.55)$$

where  $\sigma$  are the Pauli matrices, and

$$\epsilon(\mathbf{k}) = C - Dk^2, \mathbf{d} = A\mathbf{k} + M(\mathbf{k})\hat{z}, \text{ and } M(\mathbf{k}) = M - Bk^2. \quad (3.56)$$

$A, B, C$ , and  $D$  are materials parameters and  $2M$  is the positive or negative energy gap between  $E_I$  and  $H_I$ .  $A$  incorporates the interband coupling to lowest order, and  $B$ , which is typically negative, describes the curvature of the bands. For  $M/B < 0$ , the eigenstates of the Hamiltonian describe a normal insulator. For a thick QW with inverted bands, however,  $M$  is negative, and the eigenstates reveal the helical edge states of a QSH insulator (QSHI). For a semi-infinite geometry with a boundary at  $x=0$ , the effective Hamiltonian at the edge, to leading order in  $k_y$ , is given by<sup>86,120</sup>

$$\hat{H}_{edge} = Ak_y\sigma_z = \begin{pmatrix} Ak_y & 0 \\ 0 & -Ak_y \end{pmatrix}, \quad (3.57)$$

which is a helical, linear spectrum with an odd number of Kramers partners.

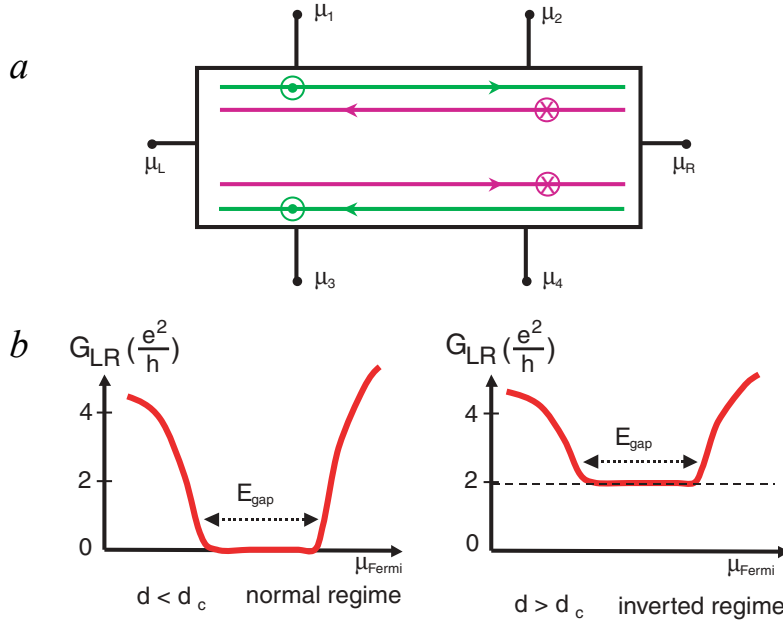


Figure 3.17: [Adapted from Ref. 120] a) Proposed setup for a QSHE experiment b) Predictions for conductance as a function of gate voltage for a thin well (left) and thick well (right).



This model readily forms predictions for transport measurements. *Figure 3.17a* is a typical device geometry, a Hall bar with gate electrode to control the position of the chemical potential. Heterostructures with thin QWs should exhibit the behavior of a typical semiconductor, namely as an increasingly negative voltage is applied to the gate, the conduction band becomes depleted and transport in the band gap is eliminated, then restored once the valence band is populated with holes (*Figure 3.17b left*). For thicknesses  $d > d_c$ , however, the four-terminal gap conductance will be quantized to  $\frac{2e^2}{h}$  (*Figure 3.17b right*).

In less than a year after the BHZ theoretical prediction, König *et al.* observed the QSHE in HgTe quantum wells. BHZ had predicted a  $d_c$  of around 6.3 nm for HgTe. König *et al.* were able to measure an insulating gap for a QW thickness of 5.5 nm (*Figure 3.18 black line*), but for QW thickness of 7.3 nm, the group observed quantized conductance close to  $\frac{2e^2}{h}$  in Hall bars with edge lengths of 1.0 and 0.5  $\mu\text{m}$  (*Figure 3.18 red and green lines*). When the spacing between probes was increased to 20  $\mu\text{m}$ , though, they found the resistance increased (*Figure 3.18 blue line*). Thus, unlike the QHE that is dissipationless on a macroscopic level, the inelastic mean free path for the helical edge states of the QSHE is on the order of only a few micrometers.<sup>86,137</sup> Potential sources of scattering are discussed further in *Section 3.6*.

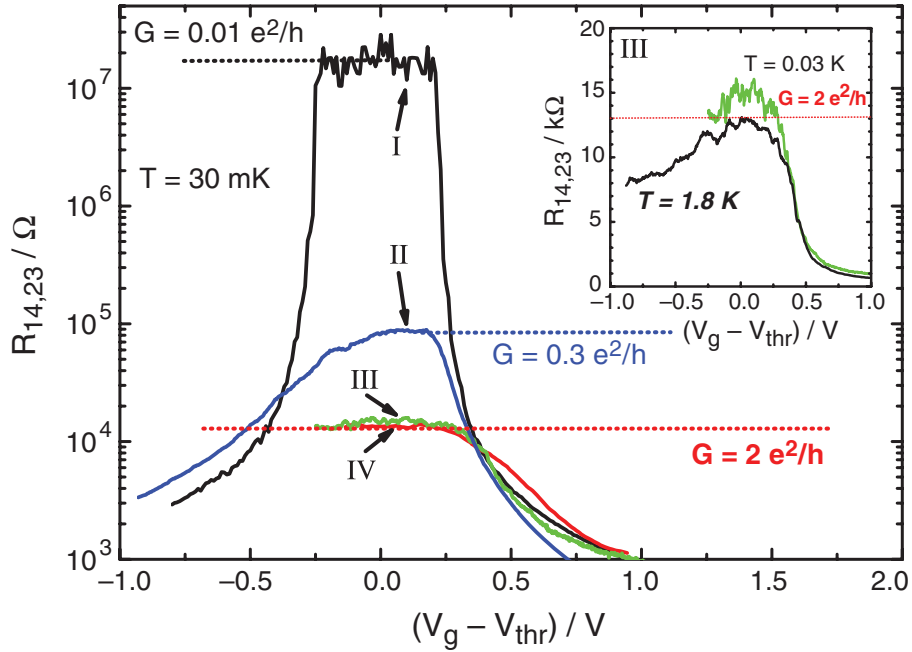


Figure 3.18: [Adapted from Ref. 121] First experimental evidence of the QSHE in HgTe quantum wells. Four-terminal conductance as a function of gate voltage for well with 5.5 nm thickness (black), 7.3 nm thickness with 1 and 0.5  $\mu\text{m}$  length between leads (red and green lines), and 7.3 nm thickness with 20  $\mu\text{m}$  length between leads (blue line). Inset) Comparison of inverted-gap devices with 1  $\mu\text{m}$  length at 30 mK and 1.8K.

Further measurements of inverted-gap heterostructures confirmed edge transport by measuring non-local geometries in transport,<sup>138</sup> and by imaging devices with scanning magnetometry.<sup>139</sup> QSH helical edges were next predicted in QWs of the type-II semiconductor InAs/GaSb,<sup>140</sup> and in 2011 InAs/GaSb QWs were confirmed to exhibit transport properties similar to the HgTe.<sup>141</sup> Semiconductors continue to be at the forefront of the search for other potential topological insulators. More recently, in 2016, Li *et al.*, using scanning tunneling microscopy, observed a large bulk band gap with topological edge states in ZrTe<sub>5</sub>.<sup>142</sup> Also in

2017, Fei *et al.* observed helical edge conduction in the type-II Weyl semimetal monolayer  $\text{WTe}_2$ ,<sup>143</sup> the first example of the QSHE in a monolayer material.

### ***3.6 Scattering Mechanisms in Two-Dimensional Topological Insulators***

While the successes of measuring helical edges in different materials continues to grow, the mechanisms that limit the ballistic transport to the few-microns scale and below are still relatively unknown. Relevant scattering processes are those capable of flipping the spin of forward-moving electrons such that they can scatter into the backward-moving state with some finite probability on the length scale of a few micrometers.<sup>121,144</sup> Further evidence of the edge states being prone to some sort of interactions, Du *et al.* found the Fermi velocity in QWs of InAs/GaSb was at least an order of magnitude smaller than that of a GaAs 2D electron gas.<sup>144</sup> The leading theory suggests inelastic scattering occurs between the helical edge modes and puddles of charge density in the bulk.

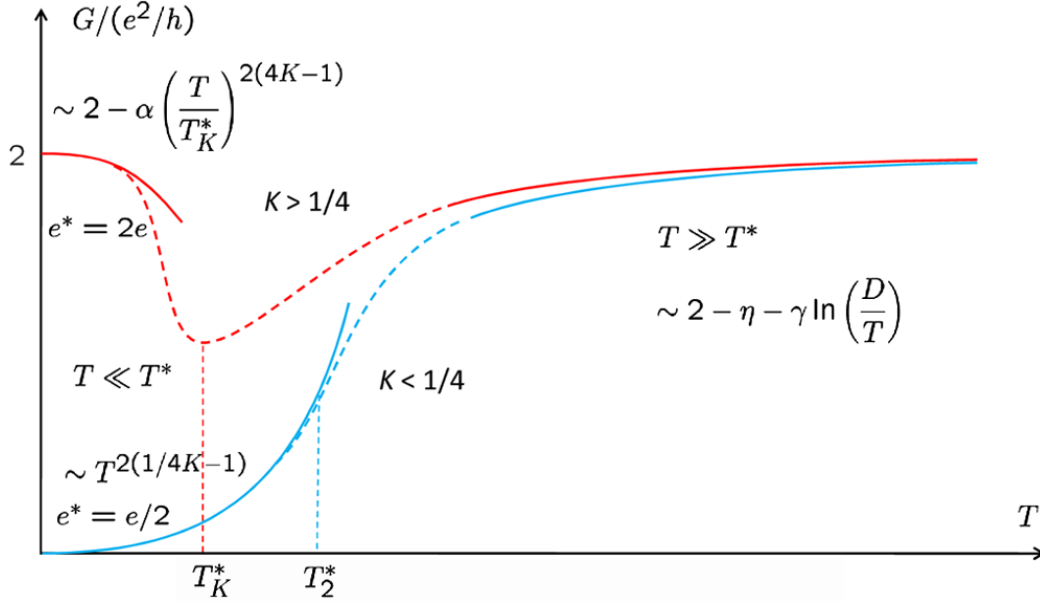
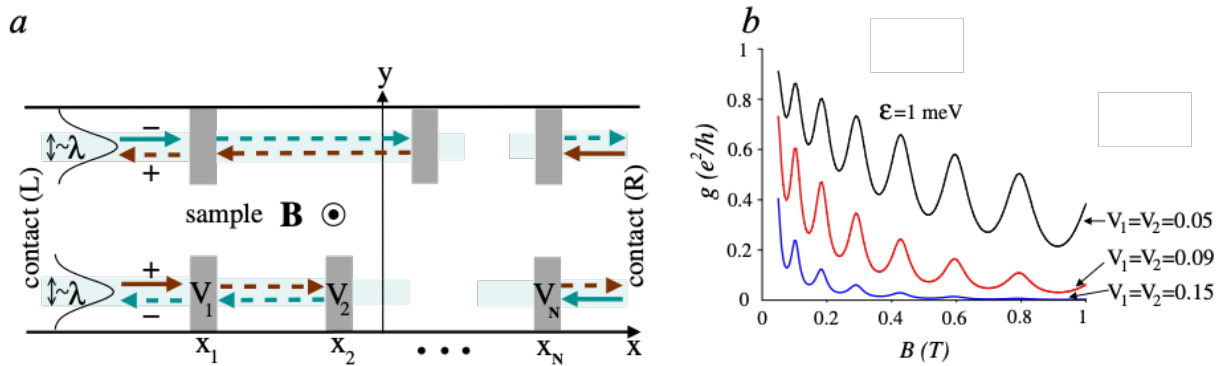


Figure 3.19: [Adapted from Ref. 145] Model of backscattering in the QSHE due to the Kondo effect. Conductance of a helical edge mode coupled to a locally doped region with an odd number of electrons for two regimes of Coulomb interactions, characterized by the Luttinger parameter  $K$ . For weak interactions (red line) scattering is maximized at a Kondo temperature  $T_K^*$  and eliminated at  $T=0$  due to the formation of a Kondo singlet. For strong interactions (blue line) the system is a low temperature insulator.

Maciejko *et al.* proposed one of the first models of the edge conduction of a QSHI in the presence of a magnetic impurity, in the form of an odd number of electrons in a locally doped region, as a function of temperature (Figure 3.19).<sup>145</sup> At high temperatures, they found the main mechanism at work was two-particle scattering and/or the Kondo effect,<sup>146</sup> a process by which conduction electrons in a metal are scattered by a magnetic impurity, which leads to a term in the resistivity that increases logarithmically with temperature. At low temperatures, they showed for weak Coulomb interactions, normal helical transport is restored, but if Coulomb interactions are

strong, transport occurs by tunneling through the impurity, and only half an electron charge is transferred per tunneling event. The strength of the Coulomb interaction is characterized by the Luttinger parameter  $K$ . For  $K > \frac{1}{4}$  (weak interactions, red line in *Figure 3.19*), scattering is maximized at a finite Kondo temperature  $T_K^*$  and eliminated at zero temperature due to the formation of a Kondo singlet. For  $K < \frac{1}{4}$  (strong interactions, blue line in *Figure 3.19*), two-particle backscattering occurs at all temperatures below some characteristic  $T_2^*$ , and the system is a normal insulator. This theory has been difficult to observe experimentally, however, due to the very low temperatures required, but the expressions for the weak coupling regime with  $K \approx 1$  seem to agree with the experimental thesis work of M. König on HgTe QWs.<sup>147</sup> Maciejko *et al.* suggest QWs with stronger interaction effects would be better suited for verifying their model.



*Figure 3.20: [Adapted from Ref. 148] a) Model of QSHE backscattering due to strong potential scatterers  $V_N$  on the edge in the presence of a magnetic field. b) Proposed conductance of edge modes as a function of magnetic field for different values of scattering site potentials.*

Another model by Tkachov *et al.*<sup>148</sup> focuses on strong potential scattering at specific points along the edge (*Figure 3.20*). They found if the Fermi level is shifted from the band gap

into the Landau-quantized conductance or valence band, the system will transition from the quantum spin Hall state into the quantum Hall state. Near the transition point as a function of field and chemical potential, the conductance of the helical edge modes is greatly suppressed due to a combination of the nonlinearity of the spectrum and enhanced backscattering. Edge channels are localized within the magnetic length, separated by the backscattering centers  $V_N$  (Figure 3.20a) that could include sample inhomogeneities where electronic trap states interact with the edge modes and randomize their propagation directions.<sup>138</sup> At zero magnetic field, this effect is believed to be weak, but near the QSH-QH transition point, they argue backscattering is greatly enhanced due to the reduction of the group velocities of the coupled QSH modes. The conductance has an oscillatory response and decays with magnetic field  $B$  as  $B^{-2N}$ , where  $N$  is the number of backscatterers on the edge (Figure 3.20b).

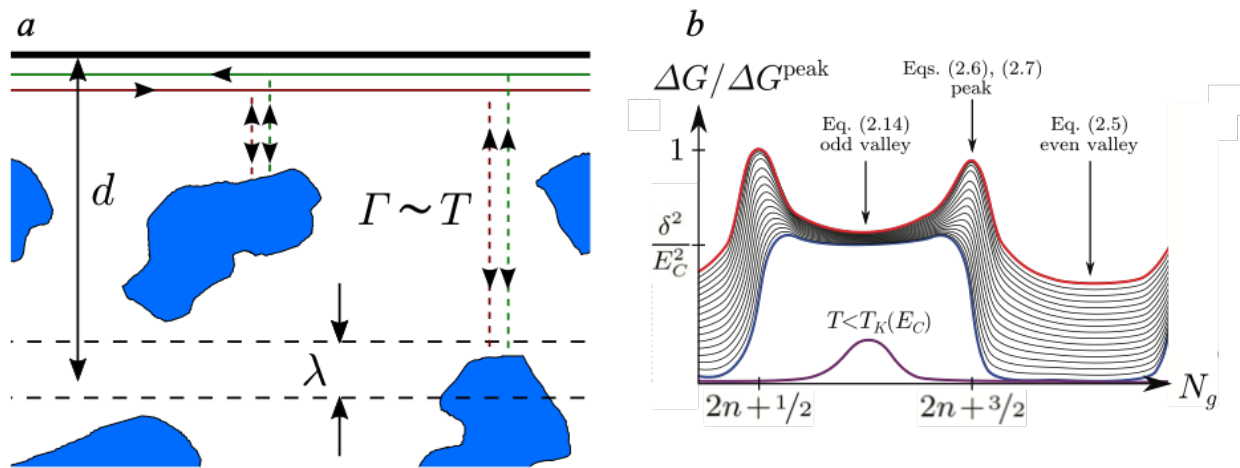


Figure 3.21: [Adapted from Ref. 149] a) Electrons moving in a helical edge tunnel into and out of puddles formed by inhomogeneous charge distribution within the bulk. Electrons may undergo inelastic backscattering within the puddles. b) Proposed edge conduction changes due to inelastic backscattering. For an odd number of electrons in the puddle, conductance change saturates with temperature, for an even number of electrons, it is proportional to  $T^4$ .

A relatively widely accepted model was proposed by Väyrynen, Goldstein, and Glazman.<sup>149,150</sup> Their theory looks into the zero-field interaction between helical edge modes and charge puddles formed by doping a 2DTI (*Figure 3.21*). They model a single puddle as a quantum dot with a well-defined number of electrons coupled to the helical edge and find it could lead to significant backscattering in the edge due to the long electron dwelling time within the quantum dot. With some finite probability, electrons can tunnel laterally from the helical edge to the dot, spend enough time in the dot to lose phase coherence, then reverse-tunnel to the backward-moving edge channel (*Figure 3.21a*). The main result is a stark difference in the temperature dependence of even versus odd number of occupied dots (*Figure 3.21b*). For an even number of occupied quantum dots,  $\Delta G \equiv \frac{e^2}{h} - G \propto T^4$ . In the case of an odd number of occupied quantum dots, ground state degeneracy initially causes weak temperature dependence,  $\Delta G \propto \ln^2\left(\frac{T}{T_K}\right)$ , where  $T_K$  is the Kondo temperature, leading to a plateau in  $\Delta G$  as a function of gate voltage. Once temperature is lowered below the Kondo temperature for the bottom of the odd valley, the odd valley contribution begins to decrease as the curve begins to take on the  $T^4$  dependence. In other words the change in conduction for an odd occupation saturates at a Kondo temperature. This model is best suited for experimental tests at low temperatures with small samples, such that the number of contributing puddles is relatively few.

While some experimental studies have been fairly consistent with aspects of these models, none of the theories have been thoroughly confirmed. Furthermore, some observed effects seem inconsistent with the proposed backscattering mechanisms. For example, though most theories predict some temperature dependence, Spanton *et al.* observed an edge resistance in InAs/GaSb independent of temperature up to 30K.<sup>234</sup>

### ***3.7 Predictions of Noise in Two-Dimensional Topological Insulators and Initial Experiments***

As discussed in Chapter 2, electronic noise can often provide answers that normal transport measurements cannot, particularly in the transmission rates of 1D channels. Several predictions for noise in 2DTIs exist, with many proposing noise as a method for distinguishing the backscattering mechanism in the helical edge modes. Theories of noise in the helical edges of 2DTIs generally model backscattering as either a constriction or quantum point contact (QPC) that allows tunneling between two edges<sup>151–154</sup> or magnetic impurity coupled to a single edge.<sup>155–</sup>

158



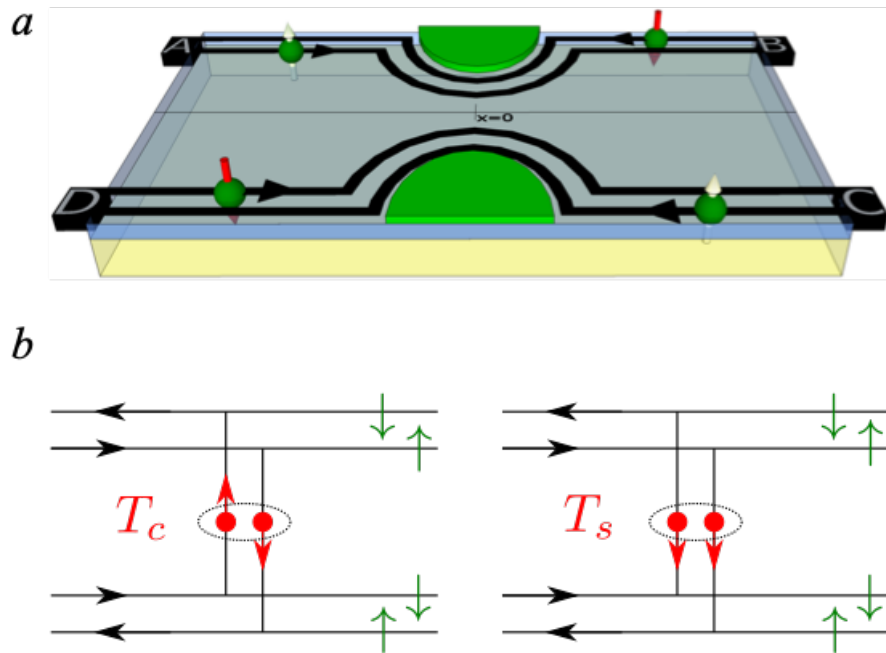


Figure 3.22: [Adapted from Ref. 151] a) Proposed experimental setup for a 2DTI with two helical edge modes. Each corner is connected to an electron reservoir held at a set bias voltage. The upper and lower channels can be coupled locally by a top gate, allowing tunneling between the two edges. b) Charged-pair (left) and spinful-pair (right) tunneling. Due to the helicity, a single  $T_c$  event turns two right-movers into two left-movers (or vice versa) so that the charge transferred from left to right changes by two.  $T_s$  affects the spin sector in an analogous way. Both processes lead to opposite cross correlations between spin-up and spin-down tunneling currents.

Figure 3.22a is an example of a proposed experimental setup for measuring current noise and current cross correlations due to interedge tunneling by Thomas Schmidt from 2011.<sup>151</sup> Using perturbation theory to derive the cumulant generating function for the edge-to-edge current, Schmidt found different possible transport channels lead to distinct signatures in the

current noise. He considers single-particle tunneling, charged-pair tunneling, and spinful-pair tunneling. The contributions to the noise by each type of tunneling couple differently to the applied voltages. Since no noise is produced along the edges themselves, the tunnel current noise is the noise measured at any of the four contacts. According to Schmidt, a clearer picture can be obtained by looking at the cross-correlations of the currents along opposite edges. The cross-correlation noise does not include any contribution from single particle tunneling because this interaction does not couple spin-up and spin-down electrons. Otherwise, the correlations are positive for charged-pair transport and negative for spinful-pair transport. In the limit of repulsive interactions with applied voltage at a single contact, the two types of pair tunneling have different voltage-dependencies. The cross-correlation noise is then a non-monotonic function of voltage and could potentially be used to distinguish between different transport regimes.

Theoretical approaches to coupling of the helical edge states with magnetic impurities span a variety of approaches, including deriving dependence of noise on magnetic field,<sup>154</sup> frequency,<sup>157,159</sup> and anisotropy of the impurity.<sup>159</sup> Motivated by experimental evidence of backscattering in HgTe QWs,<sup>138,160,161</sup> Aseev and Nagaev<sup>156</sup> calculated the edge resistance and nonequilibrium current noise that arise from the tunnel coupling between edge states and puddles of charge within the bulk believed to arise from inhomogeneous distribution of doping impurities (*Figure 3.23*).<sup>149</sup> Under the assumptions that the impurity puddles have a continuous energy spectrum and that the motion of electrons within the puddles is two-dimensional, the impurity scattering combined with spin-orbit coupling may result in temperature-independent spin relaxation. The result is increased resistance of the edge states and finite shot noise. The authors explore the limiting cases of either absent or very strong energy relaxation of electrons in the

puddles and compare the magnitude of the shot noise with the increase in resistance to ascertain the model's relevance to experimental systems.

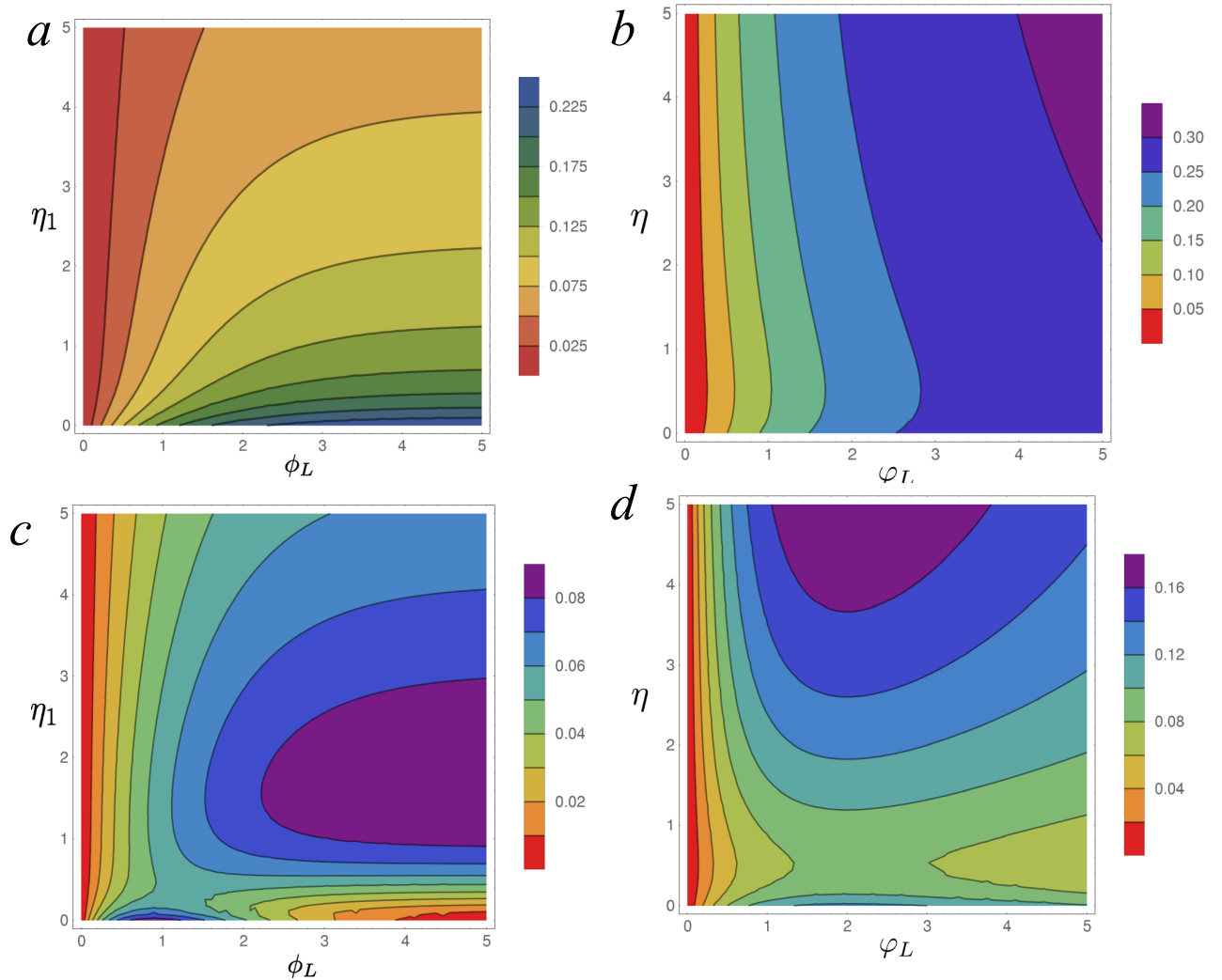


Figure 3.23: [Adapted from Ref. 156] Contour plots of Fano factor as a function of coupling strength  $\phi_L$  and normalized spin flip time  $\eta$  for a) one puddle in the absence of energy relaxation, b) continuous distribution of puddles in the absence of energy relaxation, c) one puddle with strong energy relaxation, and d) continuous distribution of puddles with strong energy relaxation.

Their general setup is a pair of helical edge states with linear dispersion  $\varepsilon_p = |p|v$  that connect electron reservoirs kept at constant voltages  $\pm \frac{V}{2}$ . Both forward- and backward-moving electron momentums are locked to a definite spin projection,  $\sigma = \pm 1$ . The edge states are able to exchange electrons with conducting puddles in the 2D bulk formed by large-scale potential fluctuations. The puddles are assumed to be large enough to have continuous spectra, and electrons inside the puddles are subject to spin relaxation due to spin-orbit processes and to energy relaxation. In the first version of their model, they consider a single puddle with arbitrary coupling to the edge states, and in the second version, they repeat the calculations for a continuum of puddles with weak coupling. For a single puddle with no energy relaxation (elastic scattering only), conductance decreases with increasing coupling strength and relaxation rate, from  $\frac{e^2}{h}$  to  $\frac{e^2}{2h}$ , while the Fano factor increases from 0 to  $\frac{1}{4}$  (*Figure 3.23a*). In the limit of a continuum of puddles, still with no energy relaxation, conductance tends to zero with increasing coupling and spin-relaxation rates, and the Fano factor increases from zero to  $\frac{1}{3}$  (*Figure 3.23b*), as in diffusive metals. It follows that in the most realistic case of several puddles strongly coupled to the edge,  $\frac{1}{4} < F < \frac{1}{3}$ .

Including strong energy relaxation does not affect the conductance, but does significantly alter the noise. In this limit, the maximum Fano factors are smaller than those in the purely elastic limit and correspond to the intermediate values of the conductance. Furthermore, the Fano factor becomes a nonmonotonic function of both the coupling strength and the spin-flip rate (*Figure 3.23c,d*), and it vanishes completely when conductance reaches zero.

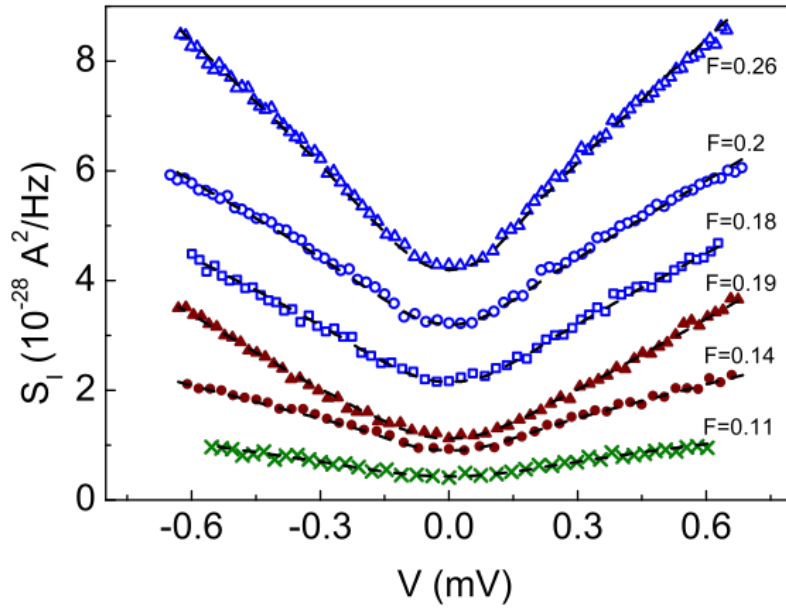
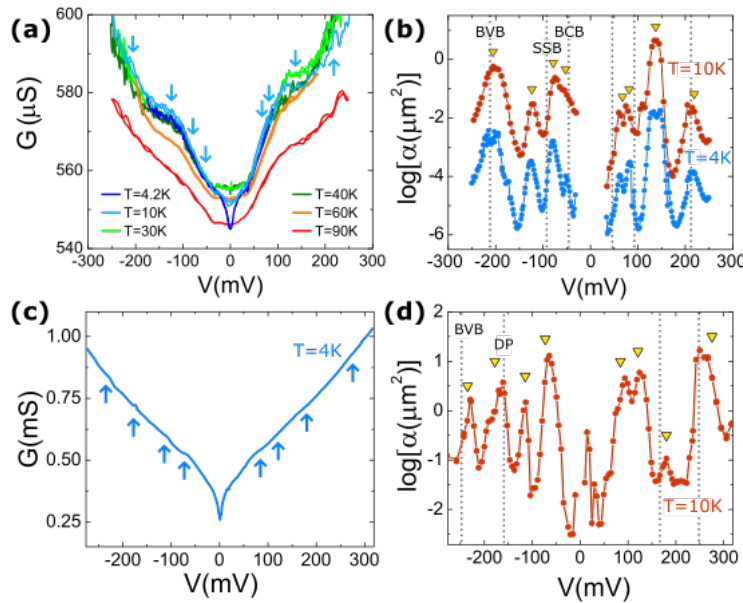


Figure 3.24: [Adapted from Ref. 162] Shot noise in HgTe QWs. Fano factor is within the range  $0.1 < F < 0.3$ , consistent with disordered multi-mode transport.

An experimental report on the shot noise in HgTe-based QWs by Tikhonov *et al.*<sup>162</sup> found Fano factors in the range  $0.1 < F < 0.3$  for noise in the regime of disordered edge transport (Figure 3.24). As their devices were shorter than even the shortest estimate of the ballistic dephasing length, the authors argue they must be in the limit of disordered multi-mode transport. They found the Fano factor was dependent on gate voltage and exhibited sample-sample variation, which, along with the range found for  $F$ , may be evidence in support for Aseev and Nagaev's model. This would seem to argue against purely topologically-protected helical edge modes as the dominant edge conduction in these samples.

Another experimental study by Aliev *et al.*<sup>163</sup> fabricated TI/Al<sub>2</sub>O<sub>3</sub>/Co tunnel junctions, with bottom TI electrodes of either Bi<sub>2</sub>Te<sub>3</sub> or Bi<sub>2</sub>Se<sub>3</sub>. They found features related to the band structure appeared in both the tunneling conductance and the low frequency noise spectrum. The

bias dependence of the  $1/f$  noise showed peaks at specific energies corresponding to features in the band structure of the TI (*Figure 3.25*). This study could be useful to repeat in the search for new TI materials and attempts to manipulate their spin-polarized properties.



*Figure 3.25: [Adapted from Ref. 163] Conductance (a,c) and  $1/f$  noise (b,d) for TI/ $Al_2O_3$ /Co tunnel junctions, with bottom TI of  $Bi_2Te_3$  (a,b) and  $Bi_2Se_3$  (c,d). Triangles in  $1/f$  noise plots indicate points of inflection in the conductance and dashed lines indicate features in the band structure.*

The most recent theoretical approach to zero-frequency shot noise due to a single magnetic impurity at the edge of a 2DTI was proposed by Kurilovich *et al.* in 2019.<sup>158</sup> Beginning with an impurity of spin  $S$ , they find for  $S > \frac{1}{2}$ , the Fano factor can be arbitrarily large, due to the bunching of large groups of electrons (*Figure 3.26b*). If  $S = \frac{1}{2}$ , however,  $1 < F < 2$ . The authors were motivated by a 2017 theoretical proposal by Väyrynen and Glazman<sup>157</sup> in

which they calculated the shot noise due to backscattering of helical edge electrons by anisotropic exchange with a local spin  $S = \frac{1}{2}$ . They quantified noise in terms of the backscattering Fano factor,  $F_{bs}$ , the ratio of the zero-frequency noise of the backscattering current and the absolute value of the average backscattering current in the high bias limit. Väyrynen and Glazman found that  $F_{bs}$  is bounded between 1 and 2, with  $F_{bs} = 1$  corresponding to independent backscattering of single electrons and  $F_{bs} = 2$  corresponding to bunched scattering of pairs of electrons. Their results were limited to cases of almost isotropic exchange interaction, which is reasonable for cases of puddles with spin- $\frac{1}{2}$ , but fails for magnetic impurities with larger spins. For example,  $S = \frac{5}{2}$  for a  $\text{Mn}^{2+}$  ion in a HgTe/CdTe QW, and moreover realistic magnetic impurities can be expected to have strongly anisotropic exchange interactions.<sup>164,165</sup> Kurilovich *et al.* generalized this approach by making  $S$  arbitrary and using a general exchange interaction matrix. The fact that they found no upper limit to  $F$  implies a dynamical magnetic moment with  $S > \frac{1}{2}$  can bunch helical electrons together. In this model, electrons are found to backscatter at a rate  $\propto S_z^2$ , where  $S_z$  is the impurity spin projection in  $z$ , which itself undergoes slow changes. The result is a modulation of the backscattering events into long, correlated pulses (*Figure 3.26a*) when  $S > \frac{1}{2}$ . This effect is absent for  $S = \frac{1}{2}$ , allowing the authors to construct an exact expression for  $F_{bs}$ , setting it precisely between 1 and 2.

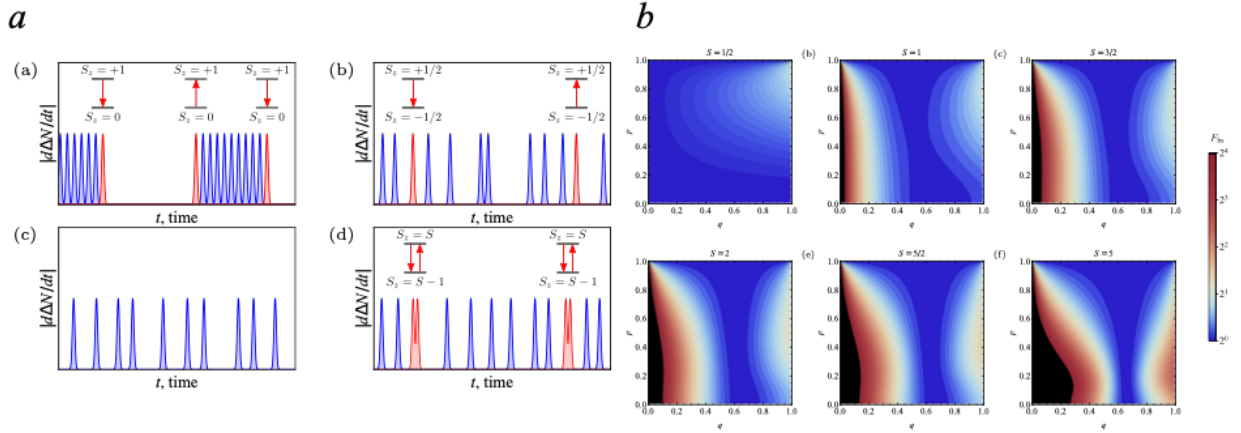


Figure 3.26: [Adapted from Ref. 158] a) Sketches of backscattering current as a function of time for (a)  $q \ll 1, S = 1$ ; (b)  $q \ll 1, S = \frac{1}{2}$ ; (c)  $p = 1$ ; (d)  $1 - p \ll 1$ , where  $p$  is a dimensionless parameter indicating the polarity of the impurity. Red and blue peaks indicate backscattering processes with and without impurity flips, respectively. Transitions between impurity levels are shown above each spin-flip process. b) Backscattering Fano factor as a function of  $p$  and  $q$  for different spin values. For  $S = \frac{1}{2}$ , the Fano factor is bound between 1 and 2. For  $S > \frac{1}{2}$ ,  $F$  diverges as  $q \rightarrow 0$ , except for  $p=1$ , for which  $F=1$ .

While theoretical and experimental studies of the shot noise in the helical edges of 2DTIs are already few, analyses of the noise arising from the insulating bulk and contact effects are essentially nonexistent. Chapter 6 of this thesis will describe an initial attempt to explore noise properties of the bulk and contact resistances using Corbino disks of InAs/GaSb QWs.



# Chapter 4

## Experimental Methods

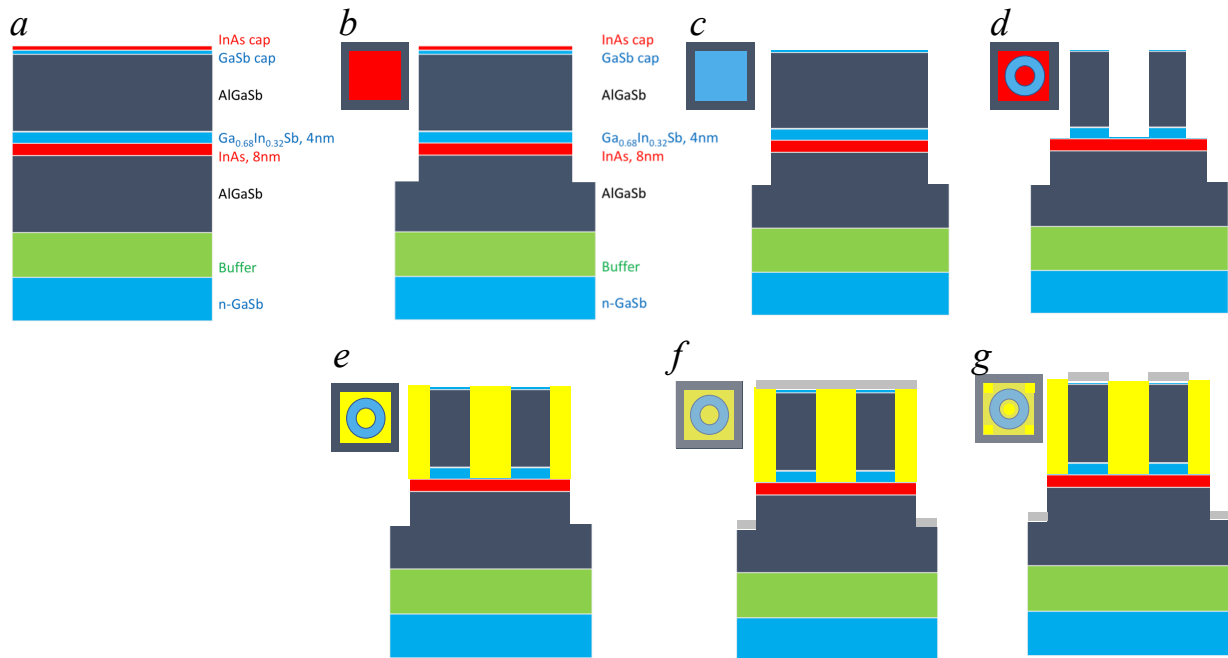
### 4.1 Fabrication of InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb Quantum Well Structures

*Figure 4.1a* shows the composition of the InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb wafers grown by Gerard Sullivan using molecular beam epitaxy (MBE) at Teledyne Scientific and Imaging. At the top are two caps, of InAs and GaSb, each 3 nm thick. Then, there is a buffer of 50 nm Al<sub>0.7</sub>Ga<sub>0.3</sub>Sb, followed by the two QWs, 4 nm Ga<sub>0.68</sub>In<sub>0.32</sub>Sb and 8 nm InAs. The bottom layer of *n*-doped GaSb serves as a global back gate. The full wafer structure and layer widths is given in Appendix 1.

For fabricating Corbino structures, the first step (*Figure 4.1b*) is to define the total mesa square by photolithography. The area outside the mesa is etched away using a “cocktail etch” solution of H<sub>3</sub>PO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub>:citric acid:H<sub>2</sub>O with corresponding ratio 3:5:55:220 mL. The citric acid component is a solution of 1 g monohydrate citric to 1 mL water. Once the mesa has been defined, the photoresist is removed, and an etch solution of 10:5 mL citric acid solution:H<sub>2</sub>O<sub>2</sub> is used to remove the InAs cap from the entire wafer (*Figure 4.1c*). Photolithography is then used again to define the contacts, and an etchant of 1:8 mL NH<sub>4</sub>OH:H<sub>2</sub>O removes the layers down to the InAs QW in the contact areas (*Figure 4.1d*).

Next, Ti (10 nm) and Au (50-100 nm) are deposited using an electron-beam evaporator (*Figure 4.1e*). As a protective coating, a 20-50 nm layer of Al<sub>2</sub>O<sub>3</sub> is deposited over the entire chip using an atomic layer deposition system (ALDS) (*Figure 4.1f*). The final etching step is to create windows to the contacts via photolithography and an Al<sub>2</sub>O<sub>3</sub> etchant known as Transetch-N (*Figure 4.1g*).

Top gates may be added over the mesa area by additional photolithography and evaporation steps. Also, if it seems the contacts have not reached the QWs, the device may be annealed in  $\text{H}_2\text{N}_2$  forming gas at  $250^\circ\text{C}$  for 5 minutes. More detailed instructions are covered in Appendix 2.



*Figure 4.1: Fabrication procedure for InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb Corbino devices. a) Initial wafer structure b) Defining the mesa c) Removing the InAs cap layer d) Defining the contacts e) Depositing Ti/Au for the contacts f) Adding a protective layer of Al<sub>2</sub>O<sub>3</sub> g) Opening windows to the contacts*

## 4.2 Cryostat System

Both conductance and noise measurements were primarily performed in a Janis Research Co. cryostat. In the Janis cryostat, either liquid nitrogen (for measurements at or above 77 K) or liquid helium (for measurements between 1.7 and 50 K) is kept in a reservoir surrounding the

central annular space where the sample is held. Samples are mounted on a custom probe with both coaxial and twisted pair wiring, and may be positioned to be parallel or perpendicular to external magnetic fields, supplied by a superconducting coil magnet. The annular space is continually pumped on, while a needle valve can be opened or closed to allow exchange gas from the cryogen reservoir to cool the sample. Temperature control is tuned by two heaters, one in the annular space, and a second directly on the probe.

Detailed instructions for filling liquid nitrogen and helium are given in Appendix 3.

### ***4.3 Conductance Measurements***

Differential conductance ( $dI/dV$ ) measurements are performed by using a summing amplifier to superimpose a small sinusoidal voltage signal (0.1-0.5 mV) with relatively low frequency ( $\sim 70$ -700 Hz) onto a dc bias applied across the sample. Both bias components are supplied by a lock-in amplifier that is synchronized to the frequency of the sinusoidal ac component. This lock-in is set to measure the first harmonic of the  $I$ - $V$  curve, while a second, also tuned to the same frequency, picks up the second harmonic component. The current signal from the junction is amplified by a current pre-amplifier before being fed to the two lock-ins as well as directly to a data acquisition module (DAQ).

### ***4.4 RF Shot Noise Measurements***

RF noise measurements are useful in that they are less sensitive to  $1/f$  resistance fluctuations and low-frequency background noise sources, such as that from the amplifiers. Following the work of Wu *et al.*<sup>166</sup> and Reznikov *et al.*,<sup>167</sup> the experimental setup for RF noise measurements is outlined in *Figure 4.2*.<sup>168</sup> A square wave is sourced from a function generator

to the junction, toggling between 0 V and some  $V_{max}$  at a frequency on the order of 70 Hz - 1 kHz. The measurements have two main components: low-frequency current response and high frequency (250-600 MHz) measurements of the change in integrated RF noise power. The two components are separated out by bias tees.

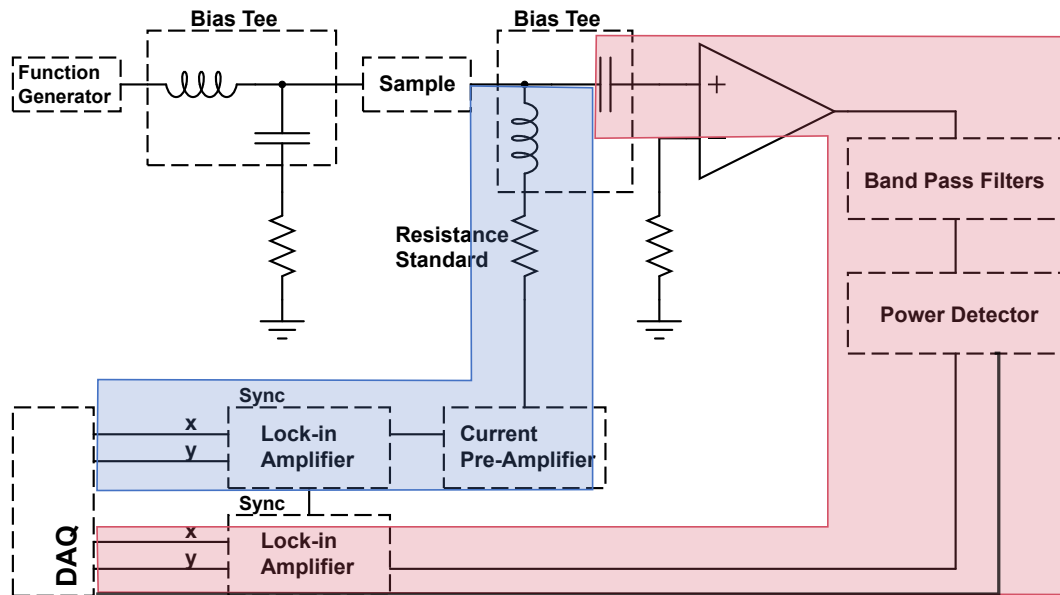


Figure 4.2: [Adapted from Ref. 168] Setup diagram for RF noise measurements. Highlighted in blue is the low-frequency signal path for monitoring current response. Highlighted in red is the high-frequency noise measurement scheme, consisting of amplifier chains, bandpass filters, and a logarithmic power detector, which feeds to a lock-in as well as directly to the DAQ.

On the low-frequency side (highlighted in blue in Figure 4.2), a current pre-amplifier follows the junction, and then a lock-in amplifier, synchronized to the square-wave, measures the current response of the device, which is then digitized by a DAQ. For devices with low resistance (or widely changing resistance), a current-limiting resistance standard is put in place to avoid overloading the current amplifier.

The high-frequency side (red in *Figure 4.2*) measures the RF noise characteristics of the junction. At zero bias, only the Johnson-Nyquist is present, but at  $V_{max}$ , there is both JN noise and shot noise. The noise signal is transmitted via coaxial cable through a chain of amplifiers, then bandwidth limited to approximately 250-600 MHz. The signal is then fed to a logarithmic power detector whose output is transmitted to both a second synchronized lock-in and directly to the DAQ. The recorded noise signal is then the change in integrated noise power, and using the gain-bandwidth product, the signal is converted into the average current noise power spectral density,  $S_I = S_I(V_{max}) - S_I(0)$ , with the direct measurement from the power detector to the DAQ as a basis for the absolute noise power. Appendix 4 describes how to translate the lock-in reading of the noise to units of  $A^2/Hz$ .

#### **4.4.1 Reflection Coefficient Considerations: Equivalent Circuit**

All coaxial cables used in the RF measurements have  $50 \Omega$  characteristic impedance to minimize RF loss and capacitive filtering of the high bias signal. All electrical components also have  $50 \Omega$  characteristic impedance. *Figure 4.3a* is the equivalent circuit deduced by P. Wheeler for the RF noise setup to determine the correct value of the current noise spectral density from the measured power signal.<sup>169</sup> One can consider the device being measured as an ideal current source of the mean square current fluctuations  $i_s^2$  in parallel with some frequency-dependent impedance  $Z_s$ . In reality, the noise source is a two-terminal device at RF, terminated at one end by  $Z_0 = 50\Omega$ . The current  $i_g$  and characteristic impedance  $Z_g$  of the single-port noise generator are then

$$i_g = i_s \frac{Z_s}{Z_s + Z_0} \quad (4.1)$$

$$\text{and } Z_g = Z_s + Z_0. \quad (4.2)$$

If then a transmission line and the RF amplifier chain, both also with  $50 \Omega$  characteristic impedance, are added to the circuit, the current delivered to the  $Z_0$  is

$$i_0 = i_g \frac{Z_g}{Z_g + Z_0}. \quad (4.3)$$

The power transferred to the  $Z_0$  load is given by

$$|i_0|^2 Z_0 = i_g^2 Z_0 = Z_0 i_s^2 \left| \frac{Z_s}{Z_s + Z_0} \right|^2. \quad (4.4)$$

*Figure 4.3b* is the diagram for a measurement of the reflectance looking from a  $50 \Omega$  line to the noise generator. The reflectance is

$$\Gamma = \frac{Z_g - Z_0}{Z_g + Z_0} = \frac{Z_s}{Z_s + Z_0}. \quad (4.5)$$

Therefore the power transferred to the amplifier is

$$P = Z_0 i_s^2 |\Gamma|^2, \quad (4.6)$$

and the current spectral density is

$$\langle i^2 \rangle = \left\langle \frac{P}{Z_0 |\Gamma|^2 G \Delta f} \right\rangle. \quad (4.7)$$

$G\Delta f$  is the gain-bandwidth product, a sample-independent quantity, found by using a spectrum analyzer. For example, the gain-bandwidth is about 162 for the setup used in *Chapter 5*.

In the limit of  $Z_s \gg Z_0$ ,  $\Gamma \rightarrow 1$ . Most samples measured are at least  $\sim 1 \text{ k}\Omega$ , and thus fall in this limit.

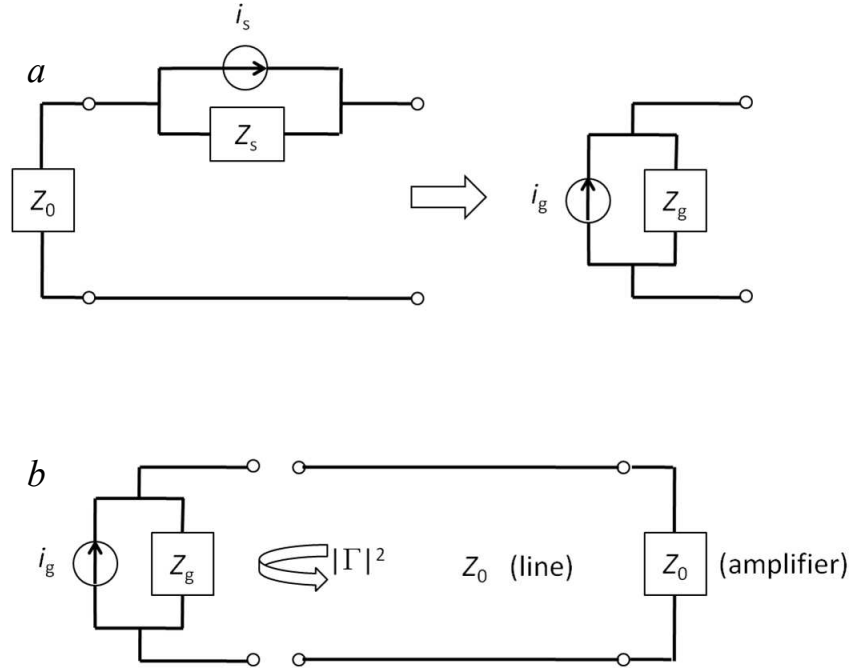


Figure 4.3: [Adapted from Ref. 169] a) Equivalent circuitry for a noise source. The sample has some root mean square current fluctuations  $i_s$  and an effective impedance  $Z_s$ . If the device is terminated on one side with impedance  $Z_0$ , it is possible to describe an equivalent circuit with a single-port noise generator  $i_g$  with effective impedance  $Z_g$ . b) Reflectance measurement for determining the actual power delivered from the sample to the amplifier chain.

In the latest iterations of the RF setup, the standard room-temperature bias tees have been replaced by individual inductors and capacitors placed on a PC board with the sample inside the cryostat. Furthermore, in an effort to precisely calibrate the high frequency noise data, the Johnson-Nyquist noise spectra for a resistor over a span of temperatures from 10 K to 300 K were measured using a spectrum analyzer. The amplified power is expected to be of the form  $P = AR_0 \frac{4k_B TR}{(R+R_0)^2} + B$ , where  $A$  is the gain of the power amplifier,  $R$  is the sample resistance,  $R_0$  is the power amplifier input impedance, and  $B$  is the background of the amplifier. By plotting  $P$  as

a function of temperature, the aim is to establish the coefficients  $A$  and  $B$ .

#### ***4.5 Low-Frequency Noise Spectroscopy***

Noise spectroscopy measurements were also performed using a low-frequency (0-100 kHz) method (*Figure 4.4*). The benefit of the low-frequency technique is the ability to clarify the absolute magnitude of the noise by confirming the accurate Johnson-Nyquist noise at zero bias,  $S_V = 4k_BTR$ , which should be white in frequency aside from parasitic capacitive effects. The downside of course is that at nonzero bias these measurements are susceptible to  $1/f$  resistance fluctuations.

The low-frequency spectroscopy is based on a cross-correlation method similar to that of Hashiaka *et al.*<sup>170</sup> and Zhou *et al.*<sup>171</sup> A tunable DC bias is applied to the device by a DAQ (National Instruments DAQ6215) through strong LC filters and symmetrically placed 150 k $\Omega$  current-limiting resistors to ensure a very clean voltage signal. Additionally, extra shielding from external noise is put in place around both the sample and the wiring. On each side of the sample is a low-noise amplifier chain that individually amplifies the voltage across the sample and its fluctuations. The total gain of each chain is  $10^4$ . The signals are then cross-correlated by a Stanford Research Systems SR785 spectrum analyzer to find the voltage noise power spectral density  $S_V$ . Amplifier noise is suppressed by the cross-correlation because only noise from the device will be correlated across both chains.



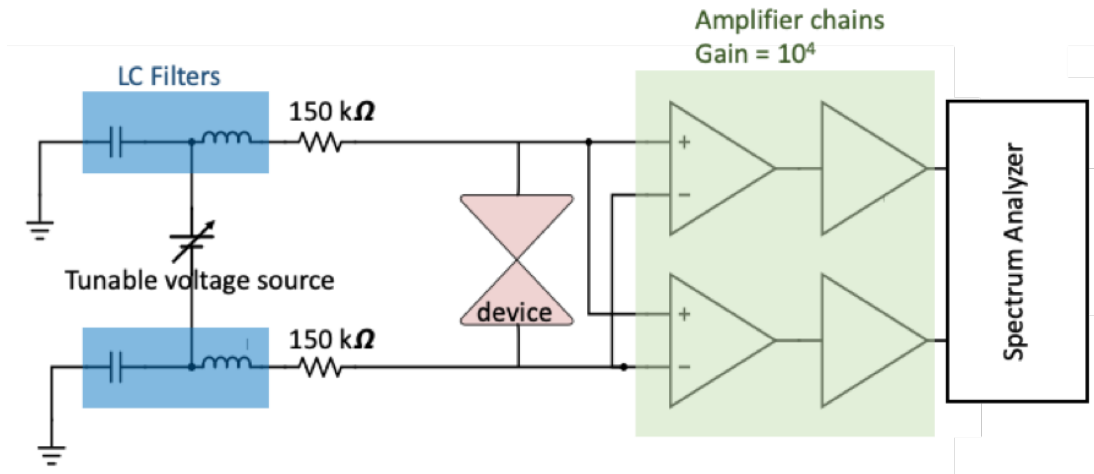
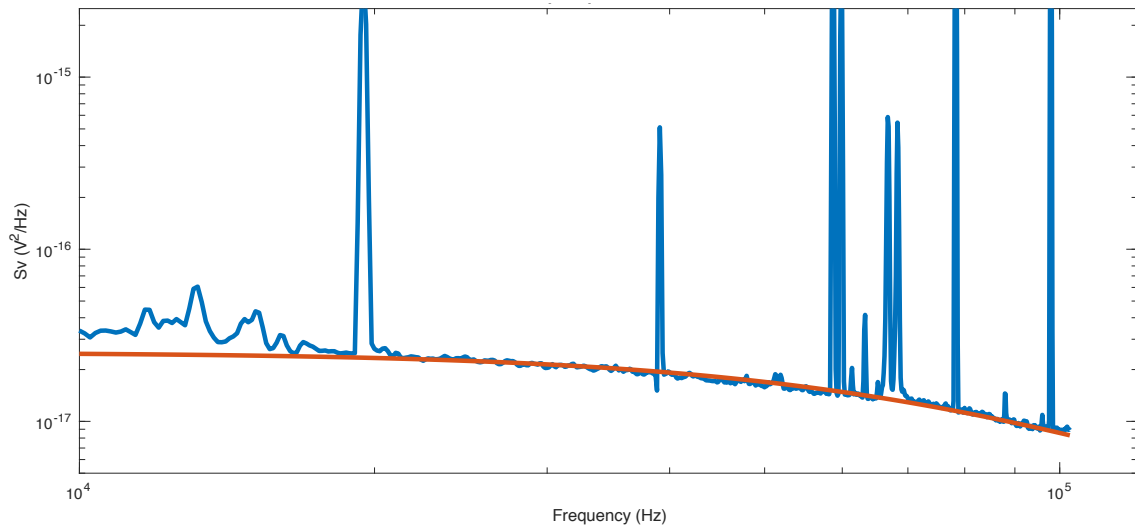


Figure 4.4: Low-frequency noise measurement setup. Noise from each side of the device is individually amplified, then cross-correlated to suppress contributions from the amplifiers.

Voltage noise spectra are fit, giving consideration to resistive and capacitive parasitic contributions, to the equation

$$S_{V,meas} = \frac{gS_V}{(1+(R_S C\omega)^2)}, \quad (4.1)$$

where  $S_{V,meas}$  is the total measured voltage noise,  $g$  is the total amplifier gain ( $10^4$ ),  $S_V = S_I^2 R_S^2$  is the intrinsic voltage noise power spectral density from the sample,  $R_S$  is the differential resistance at the applied bias,  $S_I$  is the current noise and the  $R_S C$  factor denotes the decay of the measured spectrum due to the parasitic resistive and capacitive factors. An example spectrum with fit is show in *Figure 4.5*. The voltage noise can then be converted to current noise by  $S_I = \frac{S_V}{R_S^2}$ , using  $R_S$  found by the differential conductance measurements described above.



*Figure 4.5: Example of a finite bias spectrum fit to Equation 4.1. Spectrum includes contributions from both Johnson-Nyquist and shot noise, which are both white in frequency except for a capacitive roll off toward higher frequencies.*

## Chapter 5

# Current Noise Enhancement via Nonequilibrium

## Phonon Backaction in Atomic-Scale Au

### Junctions

#### 5.1 Introduction

Measurements, by Ruoyu Chen in our group, of ensembles of atomic-scale gold junctions found a discrepancy between low bias and high bias measurements of current noise.<sup>172</sup> At low bias, measurements were consistent with the finite-temperature Landauer-Büttiker picture (Section 2.6), but at high bias, the ensemble-averaged excess noise was enhanced above the value of  $S_I$  extrapolated from the low-bias data (Figure 5.1). For all data sets, the nonlinearity of the excess noise is concave upward and independent of the order in which biases were applied, implying that the nonlinearity is an intrinsic function of the bias and not due to irreversible changes made within the junction. In exploring possible origins for the superlinear noise, quantitative analysis ruled out flicker noise<sup>39</sup> and bulk heating of the electrode.<sup>173</sup> Other candidate mechanisms for enhanced noise consistent with the observed bias dependence included modification of the Fano factor due to electron-phonon interactions,<sup>174,175</sup> particularly quasi-equilibrated phonon populations, or local electronic heating from Fermi liquid shear viscosity effects.<sup>176,177</sup> The observed power law of the noise increase seemed to imply that the electrons did not significantly pump local phonon modes. If the increased noise was modeled instead as local electronic heating, however, the enhancement implies an elevated electron temperature of

up to 330 K at a bias of 300 mV. The measured noise was roughly consistent with a model accounting for effects of electron viscosity.

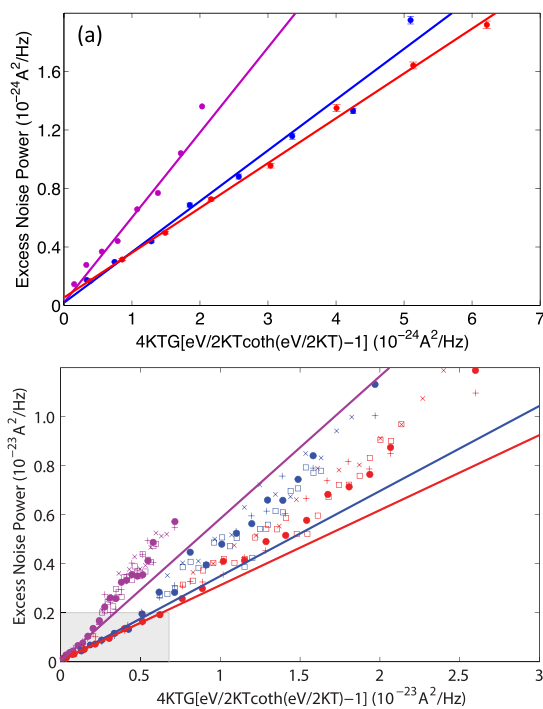


Figure 5.1: [Adapted from Ref. 172] Excess noise power as a function of scaled bias for low bias (top) and up to high bias (bottom) for three different conductance values (purple =  $0.8G_0$ , blue =  $3G_0$ , and red =  $5G_0$ .) The shaded region in the bottom plot corresponds to the top plot. At low biases, the excess noise power is linear with scaled bias, but as bias increases, significant nonlinearities are present.

To further clarify the origin of the enhanced noise, we studied the bias dependence of STM-style Au break junctions (Section 1.5) at 77 K.<sup>168</sup> The large number of successive breaking and reforming cycles of an atomic-scale contact in the STM-style experimental setup provides a clear picture of the average system behavior. During a single breaking cycle, each particular junction geometry lasts only a short period of time. For instance, at low bias, a typical

conductance plateau lasts for around 100 ms. By using this method of rapid data collection over short time durations with many repetitions, measurements are less susceptible to the junction instabilities at high bias seen in mechanical break junction techniques.<sup>55</sup>

77 K was chosen as an appropriate temperature for the measurements, as it is low enough to suppress the equilibrium population of the gold optical phonons ( $\sim 17$  meV or 200 K), but it is high enough to avoid any large contributions of bulk heating of the electrodes. Compared to Au STM junctions measured at room temperature, those at 77 K also exhibit better temporal stability, drastically reducing contributions from transient states during atomic rearrangements and allowing for more detailed quantitative analysis.

## ***5.2 Experimental Setup***

Current noise measurements were performed using the RF technique described in *Section 4.3* and previously reported by our group.<sup>169,172,178</sup> Measurements were performed in the Janis cryostat with attocube piezo positioners and minicoaxial wiring for RF signals. The environmental temperature of the gold substrate and gold tip were set at 77 K by exchange gas from liquid nitrogen in the cryostat reservoir. Extra care was taken to isolate the system from external RF signals and mechanical vibrations, including keeping the reservoir very full and performing measurements overnight to avoid vibrations from building construction. During measurements, a thick gold wire, carved at one end into a sharp tip is continuously moved into and out of contact with a 300 nm gold film evaporated onto a silicon substrate. The tip movement is controlled by one of the piezo positioners. Over a period of 4.2 seconds, atomic-scale junctions are repeatedly formed and broken between the two gold contacts. Simultaneously, noise and conductance data are collected at a rate of 200k samples/second.

To collect ensemble averages of the current noise at a given conductance for a range of biases, around 1000-1500 breaking traces were acquired at 20 different biases. Conductance histograms were recorded, and the presence of the expected peaks near 1, 2, and 3  $G_0$  with low surrounding background was used as the figure of merit for each data set.

After collection, the raw data of conductance traces were processed with an automated, multistep procedure to remove transition points, leaving only noise and conductance data for conductance plateaus, when the junction is in a stable configuration. The first step in the process removes points in the data adjacent to transients in the logarithmic power detector output that result from the times when the resistance changes abruptly during breaking. Next, readings for which the background noise level exceeds two standard deviations of the open circuit background are removed, as well as the following sixty points, to ensure all spurious data surrounding a transient are caught. After that procedure, points with a conductance not within 1% of the 20 preceding readings are removed. Lastly, a final check ensures all remaining data points are within 3% of a neighboring point, so that all remaining data is within a stable conductance plateau. This procedure minimizes any artifacts of stepwise changes in the conductance that can cause speciously high standard deviations of the noise when the junction is undergoing an atomic reconfiguration, despite only a small amount of the data lying in these regions.<sup>179</sup>

Excess current noise ( $S_{I,meas} = S_I(V_{max}) - S_I(V = 0)$ ) is averaged over conductance bins of  $0.01G_0$ , chosen to be small to avoid over-averaging. The background level that was subtracted out for each applied bias was calculated by taking the mean of the measured noise at the 10k points with the lowest conductance values at that bias. These points were chosen as the

closest approximation to the broken state ( $G = 0$ ). This background was approximately  $1 \times 10^{-24}$  A<sup>2</sup>/Hz for every bias.

At finite temperature, the excess current noise is expected to be given by

$$S_I = 4k_B T G_0 \sum_{i=1}^n \tau_i^2 + 2eV G_0 \coth\left(\frac{eV}{2k_B T}\right) \sum_{i=1}^n \tau_i(1 - \tau_i), \quad (5.1)$$

which should be linearly dependent on a scaled bias

$$X = 4k_B T G \left[ \frac{eV}{2k_B T} \coth\left(\frac{eV}{2k_B T}\right) - 1 \right]. \quad (5.2)$$

The ensemble-averaged Fano factor for each conductance bin is then taken to be

$$F = \frac{S_I(V_{max}) - S_I(V=0)}{X}. \quad (5.3)$$

The calculated Fano factor is plotted as a function of conductance and compared to the minimum Fano factor line, the theoretical plot of the Fano factor in the ideal Landauer-Büttiker regime of successive channel openings with full suppressions at points of fully-transmitting channels. Using the lowest bias dataset as a reference point, it is then possible to establish an overall normalization of the  $S_I(V_{max}) - S_I(V = 0)$  signal.

### 5.3 Results and Analysis

#### 5.3.1 Low Bias Data

At low biases, current noise and conductance data agreed well with the Landauer-Büttiker picture. *Figure 5.2* shows the data and analysis for an example bias of 160 mV, applied across both the junction and the series-limiting resistor. The conductance histogram (*Figure 5.2a*) displays the large peak at  $1G_0$ , with smaller peaks at 2 and  $3G_0$ , as expected for Au junctions.<sup>17,169,179</sup> *Figure 2b* is the calculated ensemble-averaged Fano factor (Equation 5.3) as a function of conductance, with the standard error. Standard error was calculated as  $\frac{\sigma}{\sqrt{N}}$ , where  $\sigma$  is

the standard deviation of Equation 5.3 for a given conductance bin, and  $N$  is the number of data points in the bin. In the range of roughly  $0.8-3G_0$ , the data closely follows the theoretical line for the case of individual transmission channels opening sequentially. Relatively clean successive channel addition has also been observed in previous studies.<sup>51,53,55</sup> At higher conductances, more channel mixing is expected as the number of possible atomic configurations grows rapidly with  $G$ .



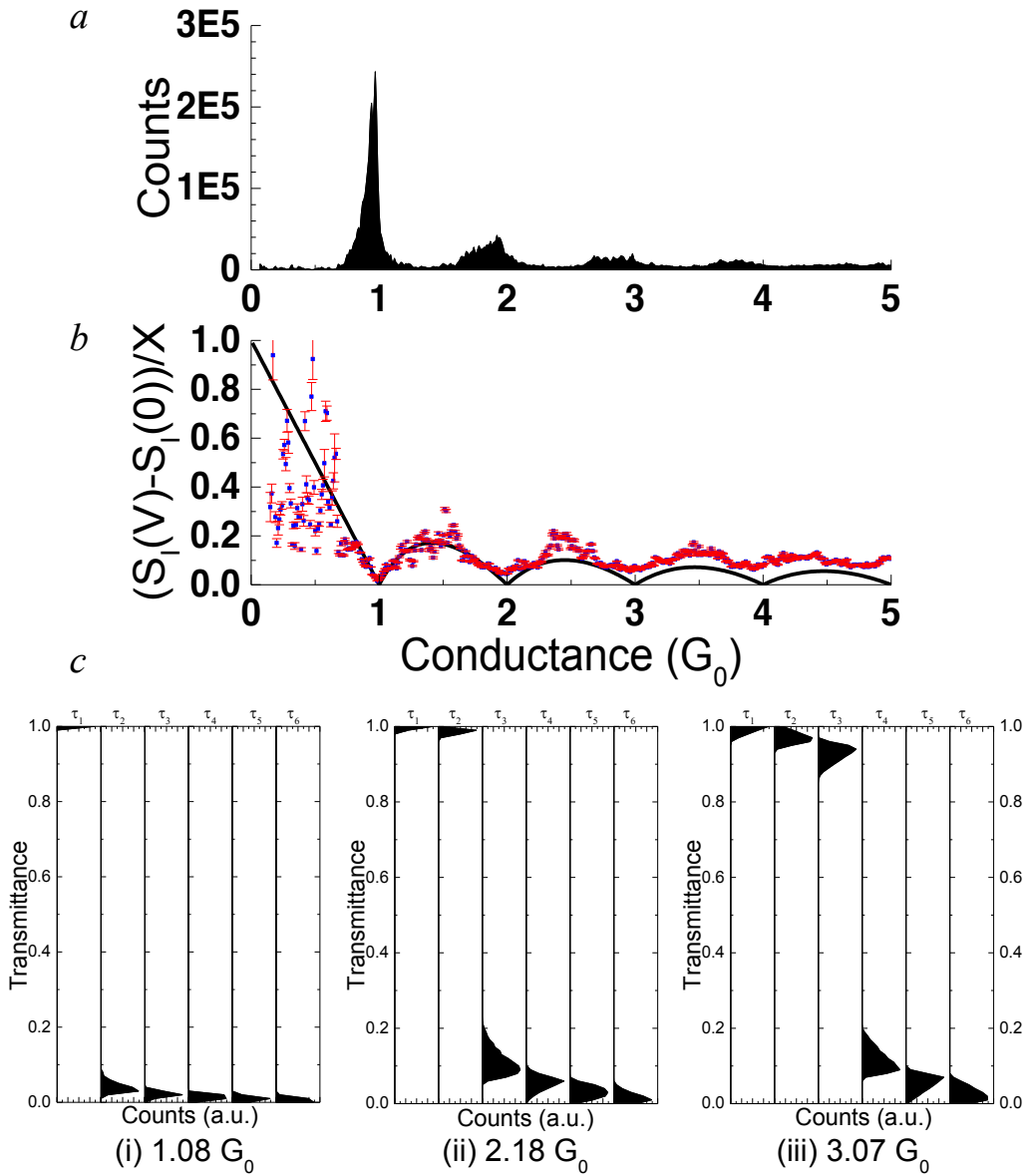


Figure 5.2: [Adapted from Ref. 168] Example data and analysis for low bias measurements (160 mV). a) Conductance histogram with typical peaks at integer values of  $G_0$ . b) Ensemble averaged Fano factor as a function of conductance with the theoretical minimum Fano factor line shown in black. The Fano factor is suppressed at integer values of  $G_0$ , consistent with Landauer-Büttiker formalism. c) Transmittance histograms of the first six channels for  $G = 1.08, 2.18,$  and  $3.07G_0$ , with respective Fano factors of 0.07, 0.09, and 0.09, calculated using the inequality in Equation 5.4. The values seem consistent with a successive channel opening picture.

In *Figure 5.2b*, we find that points below  $\sim 0.75G_0$  are less consistent, with many points lying in the “forbidden” regime below the theoretical  $F(G)$  line. During the course of a breaking cycle, comparatively little data is collected in this conductance range because junction configurations are considerably less stable for  $G < 1G_0$ , and therefore most data points are removed by the transient-removal analysis program. The small number of counts for this range is apparent in the conductance histogram. Consequently, the noise data for conductances below  $\sim 0.75G_0$  were much sparser and more broadly distributed than the high conductance data. Additionally, the RF detection efficiency seems to worsen for low conductance junctions. Therefore, we attribute the data within the forbidden regime to be due to the sparseness and subsequently inadequate averaging in this range, as well as the degraded RF detection. We acknowledge that Fano factors below the theoretical line have been observed by Vardimon *et al.* in molecular junctions between ferromagnetic electrodes due to spin-polarized conduction,<sup>63</sup> but it is very unlikely to be applicable to this system. Thus, we exclude this low-conductance data from our analysis.

To examine the role of channel transmittances in the conductance and noise of the junctions, we followed the example of Vardimon *et al.* to calculate the  $\tau_i$  values based on the measured conductance and Fano factor.<sup>53</sup> As described in *Section 2.6*, if the junctions are treated within the Landauer-Büttiker approach as coherent quantum conductors with transverse dimensions on the order of the electron Fermi wavelength, the conductance of the junction can be considered as the sum of the conductances of individual channels. Experimentally, the Fano factor was found by Equation 5.3 from the shot noise dependence on applied bias, and the

conductance was read by the lock-in (*Section 4.3*). The general assumption is that the derived transmission coefficients must satisfy the following inequalities

$$G - \Delta G \leq G_0 \sum_{i=1}^N \tau_i \leq G + \Delta G \text{ and } F - \Delta F \leq \frac{\sum_{i=1}^N \tau_i(1-\tau_i)}{\sum_{i=1}^N \tau_i} \leq F + \Delta F, \quad (5.4)$$

where  $\Delta G = 0.01 G_0$ , based on the bin size, and  $\Delta F = \frac{\sigma}{\sqrt{N}}$ , the standard error of the ensemble averaged Fano factor at each conductance. The process also operates under an additional constraint,  $\tau_i \geq \tau_{i+1}$ , enforcing sequential channel opening. The bin size for the transmission coefficients is set to  $\Delta\tau = 0.005$  to ensure a finite number of solutions. This model is valid in the limit that the transmission of the last channel goes to zero. Thus, we chose to use  $N = 6$  for conductances up to  $3G_0$ , and at low bias the last three channels all tend toward zero (*Figure 5.2c*). For gold, six channels is a reasonable limit for the low-conductance regime, since, as discussed in *Section 1.5*, the number of channels for a single atom contact is limited by the number of atomic valence orbitals.<sup>23</sup> Gold is a monovalent metal, implying single atom conduction is carried out by one transmission channel, namely the  $s$  valence orbital, such that each gold atom should contribute only one channel per  $1G_0$  of conductance.<sup>23,53</sup> For  $G \leq 3G_0$  then, there should ideally only be three conduction channels involved in transport through the junction.

Given the described criteria, a Monte Carlo simulation produces an array of possible sets of transmission coefficients for six channels for a given conductance and Fano factor. Normalized histograms are then created from these arrays to portray the most likely values for each transmission coefficient. *Figure 5.2c* shows example histograms for a bias of 160 mV at conductances of 1.08, 2.18, and 3.07  $G_0$ , with respective Fano factors 0.07, 0.09, and 0.09. The results are quantitatively consistent with those of Vardimon *et al.* in Au junctions at 4.2 K.<sup>53</sup> Furthermore, similar to Bürki *et al.* and Vardimon, we observe single channel saturation up to

$1G_0$ , but as conductance increases above this, contributions from partially opened channels also increase.<sup>53,180,181</sup> We therefore limit our analysis and comparisons between low and high bias transmittance distributions to the few channels regime of  $G < 3G_0$ .

### ***5.3.2 Bias Dependence***

Similar to the aforementioned room temperature studies, we observe a superlinear dependence of measured noise on the scaled bias at high applied biases. In this high bias regime, the Fano factor rises considerably above the theoretical minimum Fano factor line. *Figure 5.3* shows the Fano factor as a function of conductance for five example biases, demonstrating the progression of the Fano factor rising with increasing bias. The enhancement at high biases was independent of the order in which the biases were applied, suggesting it is an inherent bias dependence and not due to irreversible changes to the junction during data acquisition. Additionally, the tip-to-substrate junction was periodically annealed at currents of 2-5 mA to minimize effects of work-hardening or contaminants.<sup>25</sup>

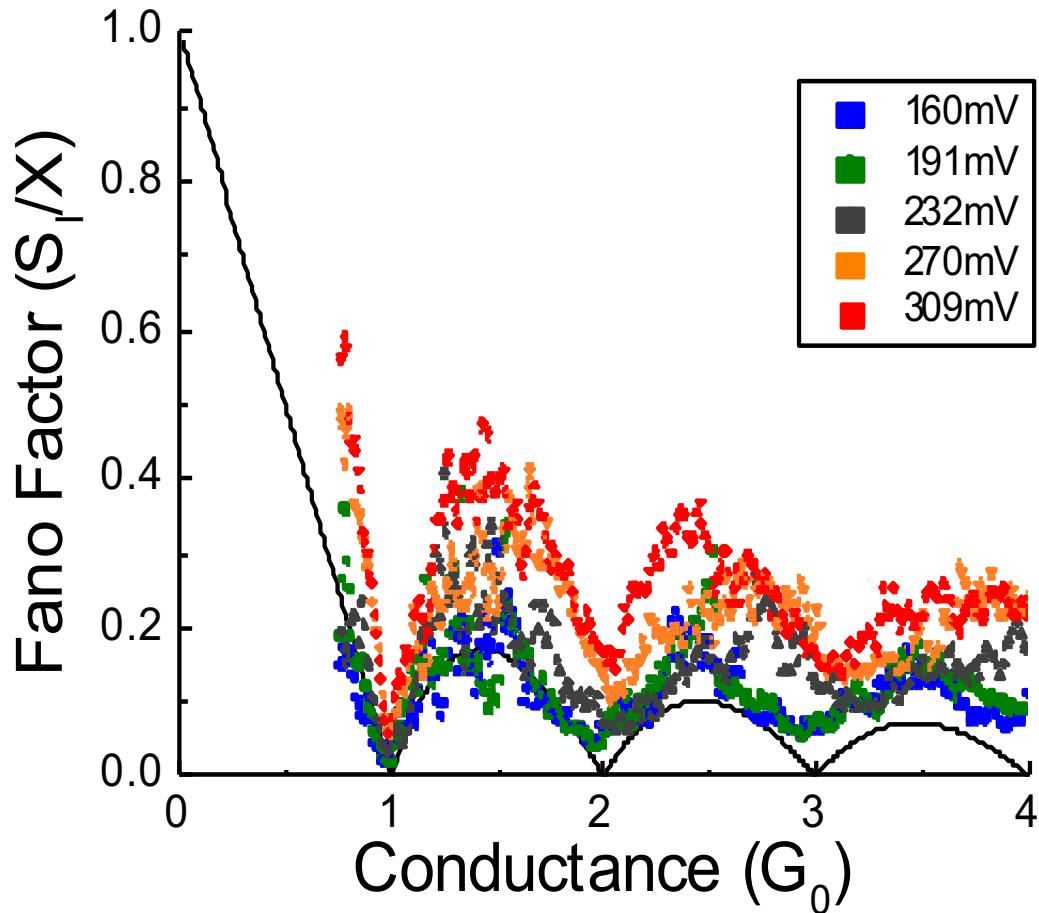
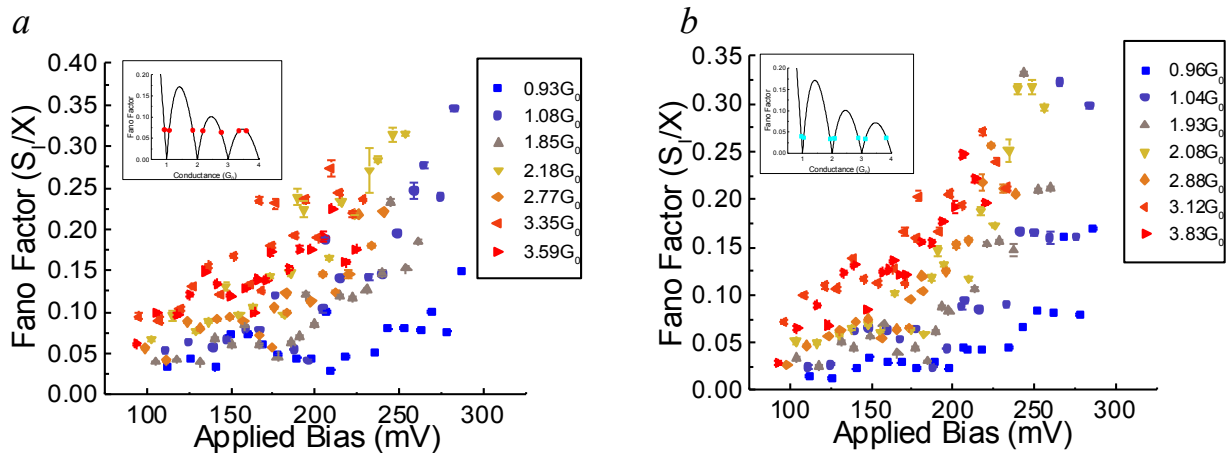


Figure 5.3: [Adapted from Ref. 168] Ensemble averaged Fano factor as a function of conductance for a sample of increasing biases, minimum Fano factor line shown in black. As bias is increased, the Fano factor rises above the values expected for the theoretical successive channel opening case. Note applied bias shown is bias across the junction,  $V_{\text{appt}} =$

$V \frac{12906}{12906 + (R_{\text{stand}}G)}$ , where  $V$  is the total bias applied by the function generator,  $R_{\text{stand}}$  is the resistance standard in series with the junction, and  $G$  is given in units of  $G_0$ .

Figures 5.4a and b are the ensemble averaged Fano factors as a function of bias voltage across the junction for several conductance values, chosen for having similar Fano factors on the

theoretical sequential channel opening line (*Inset*). In general, the trend appears to be conductance-dependent, with the spread in data increasing with increasing bias, but for all conductances, there is some trend of enhancement. Furthermore, the enhancement factor appears to be even larger at 77 K than that observed in the room temperature measurements. For example, at  $1.08 G_0$  (*Figure 5.4a*) we observe a roughly four-fold enhancement of the Fano factor, compared to a less than twofold enhancement around the same conductance over a similar bias range at room temperature.



*Figure 5.4: [Adapted from Ref. 168] Ensemble averaged Fano factor as a function of bias across the junction for conductances with similar values on the minimum Fano factor line (Insets).*

*While the trend seems conductance dependent, there is some enhancement at every conductance.*

### 5.3.3 Candidate Mechanisms for Enhanced Noise

As mentioned in the room temperature study, a number of mechanisms could be responsible for the enhancement of noise at high bias, but extrinsic effects are quantitatively unlikely to be responsible for our observations.<sup>172</sup> Flicker ( $1/f$ ) noise has been previously ruled out by the observed current noise dependence on bias and by repeating measurements over a higher frequency bandwidth with no appreciable change in results. Likewise, we rule out trivial

resistive heating of the bulk electrodes. While the heating of the bulk electrodes when the junction is biased can increase the level of Johnson-Nyquist noise, and that perceived increase would appear in the measured lock-in signal, this effect is set by the magnitude of the dissipated power. The thermal conductivities of the tip and metal film were estimated for our experimental conditions to be much too small to explain the large magnitude of noise increase.

In considering possible mechanisms of noise enhancement, we first considered the models proposed by the room temperature work: 1) local heating of the electronic distribution and 2) electron-phonon interactions. Unlike interactions with equilibrated phonon populations, which should freeze out at decreased temperatures, intrinsic electronic heating due to Fermi liquid viscosity effect would be expected to have a comparatively weak temperature dependence, since the Fermi temperature of gold ( $\sim 64,000$  K) is always much larger than experimentally accessible temperatures. Attempts to model the 77 K data as a rise in the local effective electron temperature from the measured noise and Fano factor using the method described by Chen *et al.*<sup>172</sup> were inconsistent with the room temperature results, yielding either negative temperature changes or temperature changes that were orders of magnitude larger than the room temperature values for similar conductances. The very simple model for the room temperature behavior assumed local heating proportional to  $IV$  and dominant thermal transport derived from the electronic thermal conductivity based on the Wiedemann-Franz law. Therefore the locally-limited increase in the electronic temperature, in other words the widening of the electronic distribution functions going into the junction, would be approximately independent of conductance. The 77 K data, however, clearly shows some dependence of the enhancement on  $G$  and the enhancement also seems to be larger at lower temperature, so we abandon local heating as a candidate mechanism and turn to other possible explanations for the noise enhancement.

Of remaining possible options for the increase in observed Fano factor, we first consider that the distribution of conductance channel transmittances,  $\tau_i$ , are changing with bias. In other words, higher biases could alter the combination of  $\tau_i$  that typically correspond to each conductance according to the minimal channel mixing picture. Changes in the preferred stable junction geometry could arise from the large current densities in the high bias regime. Indeed, for the largest biases applied in this study, the nominal current density is comparable to those employed in electromigration techniques, a process by which atoms of a material shift due to momentum transfer from conduction electrons.<sup>182</sup> It is conceivable then, that gold atoms in the junction may shift in response to the large current densities and acquire different stable configurations that may be unrelated to those that favor fully transmitting channels.

If this were the dominant effect, one would expect to observe a strong bias dependence in the conductance histograms, such as broadened peak widths or the appearance of new or different conductance maxima for data acquired at high biases, reflecting a preference for junction geometries that favor increased mixing. Some signs of this are seen in *Figure 5.5*; the overall number of counts decreases with increasing bias, despite the data being drawn from the same number of breaking traces because the traditionally stable junction configurations are less so at high bias. For example, at low bias, a typical  $1G_0$  conductance plateau during breaking lasted on the order of 100 ms, while at high bias, this lifetime dropped to 10 ms. Like in electromigration processes, at sufficiently large current densities, momentum may be transferred from conduction electrons to atoms at surfaces or grain boundaries, causing them to be displaced. In addition to large current densities, the high electric fields due to large potential differences over short distances across the junction may also contribute to mechanical instability. As a result, configurations that are very stable at low biases are significantly less so at high biases,



which is reflected by smaller peaks in the histograms. When the histograms are normalized to their  $1G_0$  peak, we find the histogram maxima are also asymmetric at high bias, developing shoulders toward lower conductance values (*Figure 5.5 Inset*), and we observe several additional peaks of smaller intensity, indicating a larger number of relatively stable configurations at non-integer conductance values. Compared to the integer  $G_0$  peaks, however, these additional features are less frequent and show no regular trend with bias. Overall, the basic shape of the histograms remains qualitatively similar. It remains an open question, then, whether dramatic changes in the transmittances can occur while keeping the general shape of the histogram consistent. Alternatively, large current densities could also contribute to inelastic processes in the junction, which may lead to weakening the stability of junction configurations, thus decreasing peak heights.

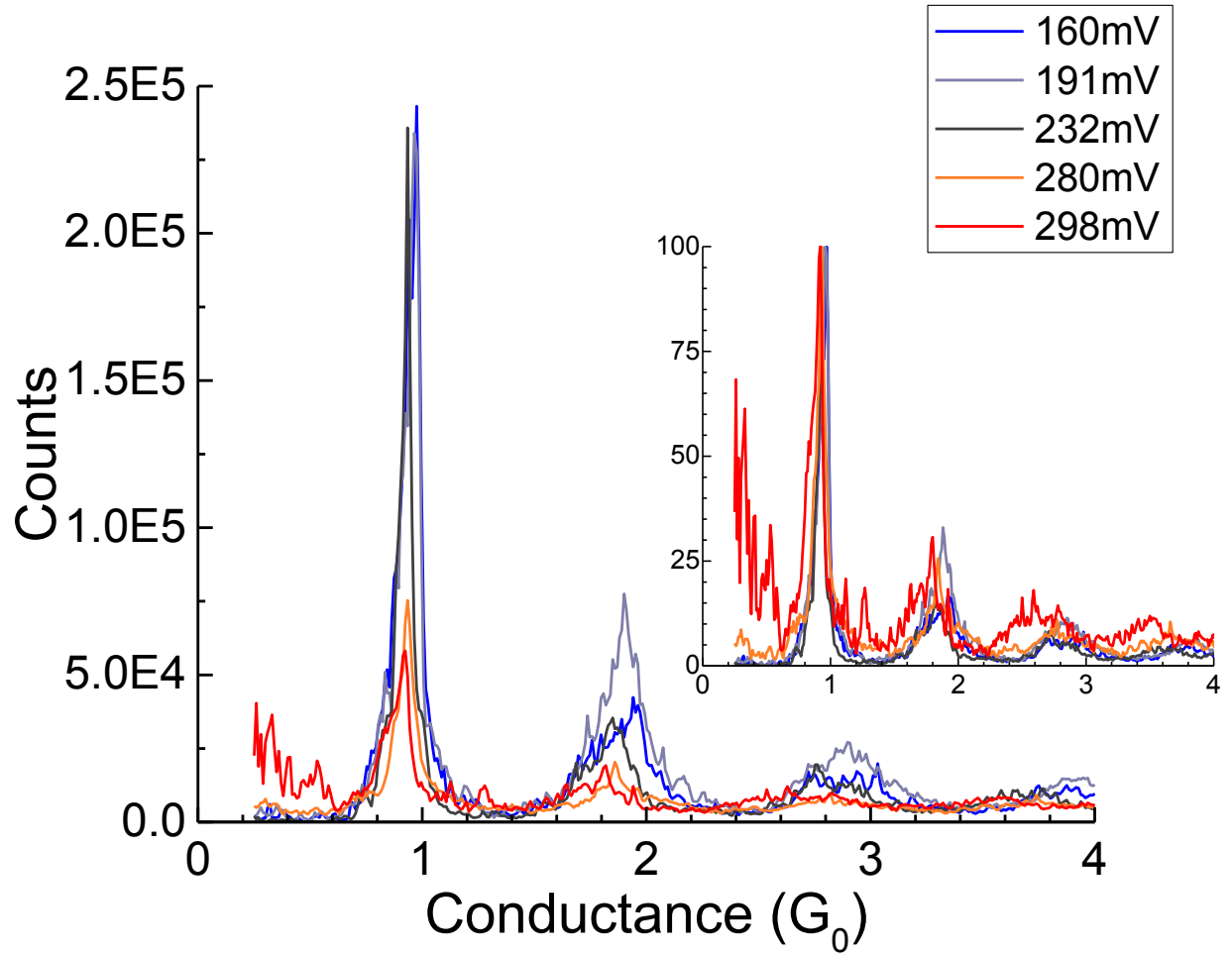


Figure 5.5: [Adapted from Ref. 168] Conductance histograms for increasing bias applied bias. The main plot illustrates the decreasing number of stable junction configurations achieved at high biases. The histograms in the inset were normalized to the  $1G_0$  peak to emphasize the changes in overall shape as bias is increased. At high biases, we note a widening of the  $nG_0$  peaks, as well as new peaks at non-integer conductances.

If bias-dependent transmittances are responsible for the elevated Fano factors, the transmission coefficients would need to change dramatically as channel mixing increases with bias. Figure 5.6 depicts the inferred transmission histograms for  $G = 0.8$ ,  $1.08$ , and  $2.18G_0$  for a range of applied bias voltages. We observe a universal lowering of the primary transmission

channel, accompanied by a broadening of all channel contributions as bias increases. According to Equation 5.1, at finite temperature, noise should be maximized for equal contributions from all involved channels. It is therefore not surprising that an enhanced Fano factor can be explained by greater contributions from more channels. While this seems reasonable from a mathematical standpoint, it is unclear what physical processes would cause a change in a channel's transmittance, particularly a decrease in the contribution by the primary channel from a fully transmitting state.

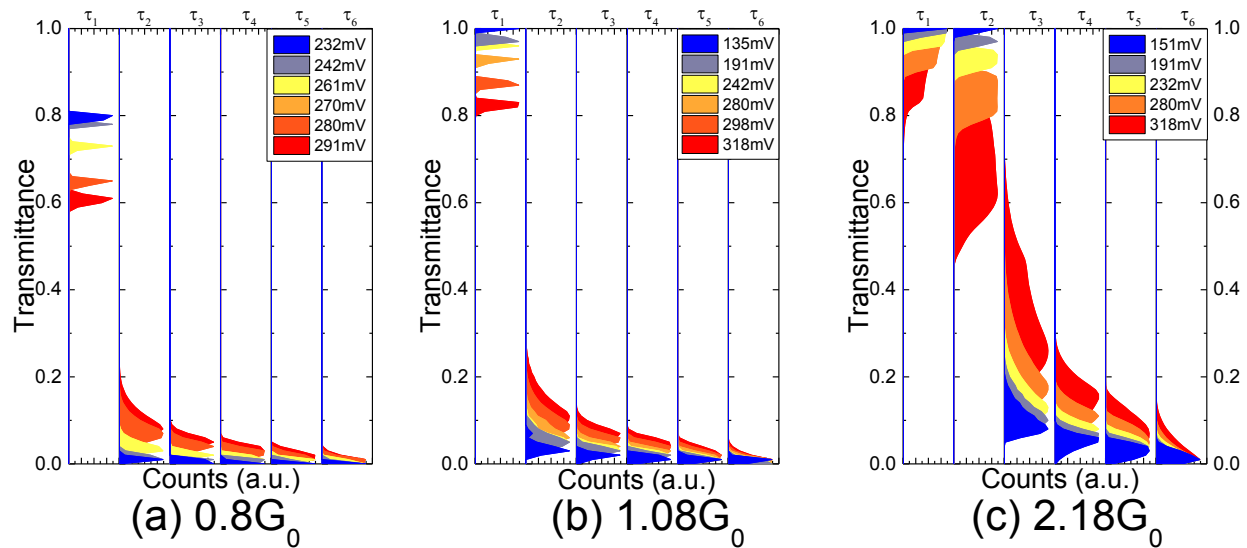


Figure 5.6: [Adapted from Ref. 168] Transmission coefficient histograms for a)  $0.8G_0$ , b)  $1.08G_0$ , and c)  $2.18G_0$  to demonstrate how bias-dependent channel mixing could account for the enhancement of the Fano factor.

Another candidate mechanism considered is the possibility of electrons coupling to a local vibrational mode. There have been several theoretical approaches to this general situation,<sup>174,175,182–187</sup> with the precise predictions for the noise depending on several parameters, such as the magnitude of the electron-vibrational coupling; multiple energy scales including the

bias  $eV$ , the vibrational quantum  $\hbar\omega$ , and the broadening of the electronic states  $\Gamma$ ; and the coupling of the local vibrational mode to the bulk phonons. The majority of theoretical predictions have been performed under the assumption of a single electronic channel and without ensemble averaging considerations. Two limiting cases are of particular relevance to this work: 1) thermalized phonon populations, assuming relaxation between the local vibrational mode and the bulk phonons is fast, so that the population of the local vibrational mode is assumed to be thermally distributed at the temperature of the substrate; and 2) strongly nonequilibrium vibrational populations, such that the local vibrational mode's population is strongly athermal and predominantly driven by coupling to conduction electrons with energies sufficient to excite the mode.

Generally, the Fano factor is expected to change when the applied bias across a system exceeds the characteristic energy of the phonon mode.<sup>174,175,182–187</sup> The particular modification of  $F$  depends on both the total transmittance of the junction<sup>174,175,184,185</sup> and on the coupling of the conductance channels to the phonon degree of freedom. In previous measurements of individual, single-channel Au point contacts at cryogenic temperatures and low biases,<sup>55</sup> discrete changes in  $F$  were observed at biases near the Au optical phonon energy (17 meV), with either increases or decreases in  $F$  with increasing bias depending on the transmittance of the single dominant channel. Likewise, in individual electromigrated Au junctions with multiple channels, qualitatively similar features have been reported at higher threshold energies.<sup>188</sup>

The environmental temperature of 77 K ( $k_B T \sim 6.6$  meV) was chosen to fulfill the condition  $k_B T < \hbar\omega$  for the Au optical phonons. Thus, thermal phonon populations should be greatly reduced relative to the room temperature case ( $k_B T \sim 26$  meV). If the previous room temperature bias-dependence of  $F$  was due to *thermalized* phonon populations, then it would be

reasonable to expect the enhancement to be considerably less at 77 K. The greater enhancement at 77 K over 300 K, then, implies that coupling to thermal phonon populations is unlikely to be the mechanism behind the increase in  $F$  with bias.

In addition to effects due to thermalized phonon populations, theories also suggest that a *nonequilibrium* phonon distribution can enhance the Fano factor in the case of a single channel. Models suggest that the large current densities arising from high biases passing through atomic-scale junctions could excite athermal phonon populations. In theories that assume strong pumping of the local phonon modes by the electrons predict contributions to the noise with voltage dependences of  $V^3$  and  $V^4$ .<sup>174,175,187</sup> In the single transmission channel model of Novotný *et al.*,<sup>174,175</sup> the  $k$ th moment of the inelastic current depends on voltage as  $V^{2k}$ , so that the current noise (the second moment of the current) should increase as  $V^4$  when the primary inelastic process of the conduction electrons is interactions with nonequilibrium phonons. No precise calculation exists for the multi-channel or ensemble-averaged case, though.

Furthermore, this model contradicts the expected bias dependence for coupling to a phonon mode with a fixed average population (such as the thermally populated optical phonons) versus backaction of current-driven fluctuations of the phononic populations. The dependence of current noise on bias voltage in the presence of a thermally equilibrated local phonon population is expected to be at most  $V^2$ . In the limit of  $eV > \hbar\omega$  and nonequilibrated phonons, the dependence of noise is expected to be  $V^4$ . *Figure 5.7* shows attempts to pull out the dominant power law of  $S_I$  as a function of  $V$ . Though the variation within the ensemble average is large enough to make constraining polynomial coefficients difficult, it is possible to plot the excess current noise as a function  $V^n$  for  $n = 1-4$  to find the best fit power law. Doing so, the data tends toward linearity as  $n$  approaches 4, suggesting  $S_I \sim V^4$ .

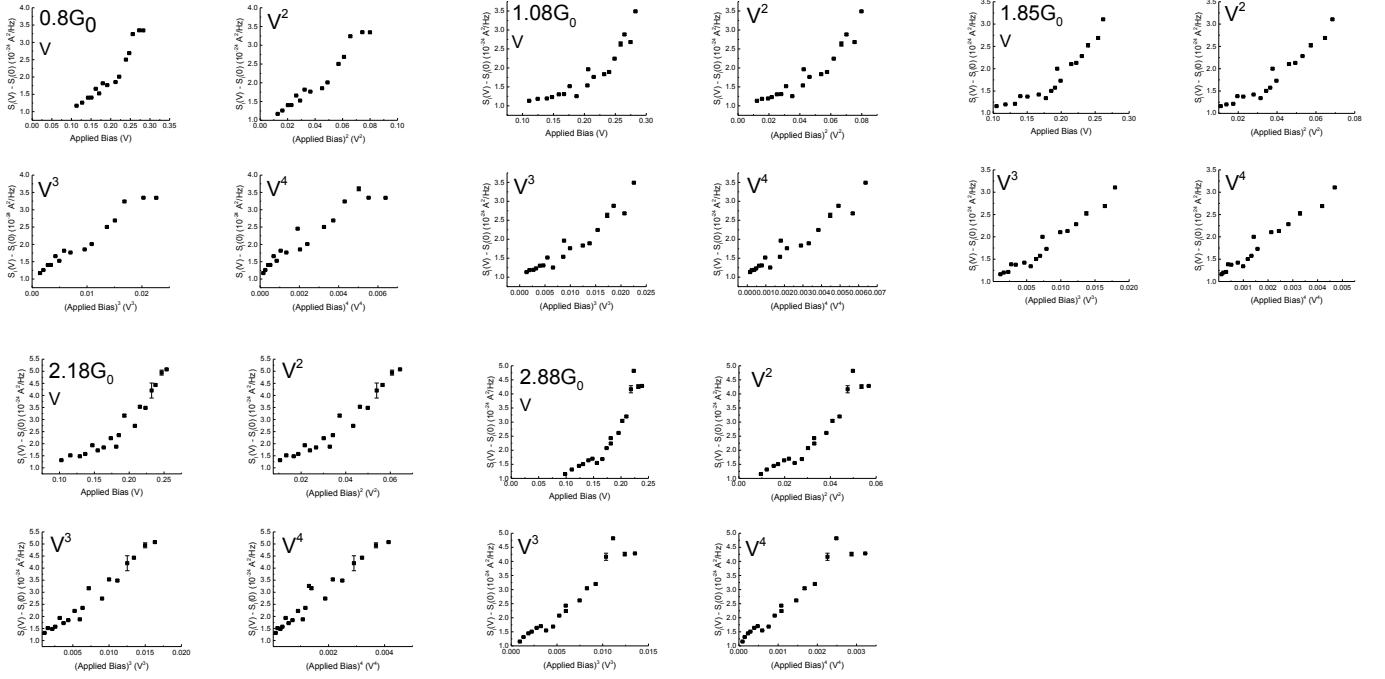


Figure 5.7: Excess noise as a function of  $V^n$  as an attempt to pull out the dominant power law of the bias dependence of the noise. While it is difficult to constrain fits to the function form of a four-term polynomial, we can plot the excess noise versus  $V^n$  and determine the dominant power law by seeing when  $S_I(V^n)$  appears most linear. We find the data tends toward linearity as  $n \rightarrow 4$ .

Based on models that consider strong pumping of local phonon modes,<sup>174,175,182–187</sup> at higher temperatures, one would expect the equilibrium contribution to dominate, while at cryogenic temperatures, the non-equilibrium, fluctuating populations should be the primary mechanism. Though at both 77 K and 300 K, the shot noise was larger than expected, the degree of Fano factor enhancement at room temperature were much smaller than those observed at 77 K. For example, at room temperature, at a conductance of  $3G_0$  the Fano factor increased from 0.35 at low bias to about 0.45 at approximately 225 mV, a roughly 30% increase. In our observations at 77 K for a conductance of  $3.12G_0$ , the inferred Fano factor rose from about 0.08

at 100 mV to 0.28 around 225 mV, a 250% enhancement. The stronger bias dependence at 77 K versus 300 K may agree with coupling to nonequilibrated phonons as a good candidate mechanism for the superlinear dependence of the Fano factor at high bias.

While this model seems to describe the data relatively well, it must still be noted that it was originally constructed assuming a single conduction channel, which is distinct from and simpler than our case of multiple channels and ensemble averaging over many junction configurations. Though it has been argued that the multi-channel case should display the same asymptotic behavior as the single-channel case,<sup>174</sup> the theoretical treatment of ensemble averaging is more complicated and has not been considered in full.

Another relevant aspect is the limit of the weak electron-phonon coupling theory, under which this model is developed. Namely, one must discern what circumstances produce strong coupling. In previous studies of atomic-scale gold junctions, the  $I$ - $V$  characteristics appeared Ohmic even at high bias, suggesting any inelastic contributions to the current are small. Fitting our current noise data for  $0.8G_0$  (the best approximation for single-channel transport) to the model by Novotný *et al.*,<sup>175</sup> we calculate the dimensionless coupling constant  $\gamma_{e-ph} \approx 0.002$ . Typically electron-phonon coupling is considered weak for  $\gamma_{e-ph} < 0.1$ , so the system appears to be well in the limit of weak coupling. The bias dependence of noise in the strong coupling limit has not yet been thoroughly calculated.

Experimentally, studies of flexible mechanical break junctions where it is easier to hold a junction at a particular conductance would be beneficial to measure both current and noise as bias is swept into the high bias limit. This would allow a direct comparison of the bias dependence in the first two moments of the current for a single junction, thereby giving a clearer picture of whether enhancement of the noise follows the predicted bias dependence for phonon

contributions. The conductance dependence may also be clarified by comparing junctions of different configurations to determine if noise increases for all configurations or if as seen in the work by Kumar *et al.*<sup>55</sup>, it decreases for certain transmittances. Studies of this nature are essential for confirming the nature of the observed noise enhancement in the high bias limit.

#### ***5.4 Conclusions and Further Progress***

In this chapter we examined the bias dependence of the ensemble averaged Fano factor in STM-style gold junctions at 77 K. At low biases, the Fano factor exhibits the behavior predicted by the Landauer-Büttiker formalism, but at high biases, the Fano factor is enhanced above the theoretical minimum line. In considering possible explanations for the enhancement, we ruled out mechanisms including flicker noise, bulk heating of the electrodes, and the model of electronic heating proposed by the work at room temperature. Our leading potential processes for the source of the increased noise are: 1) increased channel mixing at high bias; and 2) back-action by current driven phonon populations. Current densities on the order of those used in electromigration could also be responsible for rearrangement of the atoms in the junction such that different geometries become stable at high biases. A reordering of the junction atoms would result in more channels being involved in transport than is expected for a given conductance, leading to an increase in the Fano factor. While this explanation works well computationally to derive an increase in the Fano factor, the conductance histograms still point to a preference of fully transmitting channels. Additionally, the large current densities at high biases in atomic-scale junctions could excite nonequilibrium phonon populations, which in turn inelastically scatter conduction electrons. In the single-channel, single-junction picture, inelastic backaction of these fluctuating phonon populations is expected to increase the noise and to have a stronger



dependence on bias than interactions with equilibrium phonon populations. While this agrees with the higher level of enhancement observed at a temperature below the gold optical phonon energy, there is still no full theoretical analysis of this effect including multi-channel transport and ensemble averaging effects. Experimentally, studies using mechanical break junctions would be useful for probing the effects of inelastic interactions by simultaneously measuring the first two moments of the current in a single junction as a function of increasing bias.

## Chapter 6

# Noise Measurements in InAs/GaSb Quantum

## Well Structures

As introduced in Chapter 3, inverted InAs/GaSb composite quantum wells were theorized in 2008 by Liu *et al.*<sup>140</sup> and in 2011 were experimentally shown by Knez *et al.*<sup>141</sup> to demonstrate the gapped 2D bulk and 1D helical edge modes characteristic of a two-dimensional topological insulator. Surprisingly, the helical edge modes persisted even with sizable bulk conduction and displayed only a weak dependence on magnetic field, a consequence of the hybridization gap opening away from the zone center.<sup>141</sup> These structures offer numerous benefits over the initial HgTe 2DTIs. Firstly, the mercury content of those devices imposes stringent fabrication restrictions, and in general, the processing and micro- and nanofabrication techniques for III-V semiconductors, including InAs and GaSb, are very well developed compared to II-VI systems such as HgTe. Additionally, the QSH phase in HgTe QWs can only be tuned by the width of the well. The band structure of InAs/GaSb, however, can also be continuously tuned<sup>189</sup> to the topological insulator state by variable gate voltages, allowing for quantitative comparison of the normal and topologically insulating states within a single sample. InAs/GaSb also has low Schottky barriers to most metals compared to HgTe/CdTe.

Since their initial discovery, further studies have sought to confirm the helical, nontrivial nature of the edge states by suppressing the bulk conductivity through doping dilute silicon between the InAs and GaSb QWs<sup>144</sup> and by replacing the GaSb quantum well with a Ga(In)Sb alloy.<sup>190</sup> This chapter will describe two noise studies in InAs/GaSb quantum well devices: the

first in Hall bar structures of Si-doped InAs/GaSb, and the second in Corbino geometries of InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb.

### ***6.1 Band Structure of InAs/GaSb Quantum Wells***

The three semiconductors InAs, GaSb, and AlSb form a lattice-matched materials set called the 6.1Å Family for the approximate value of their lattice constants.<sup>191</sup> The 6.1Å Family are Type-II semiconductors, meaning they exhibit a staggered gap band structure. The room temperature energy gaps of the family range from 0.36 eV in InAs to 0.78 eV in GaSb to 1.61 in AlSb. *Figure 6.1* shows the band lineups for the 6.1Å Family. The most interesting feature is the broken band gap alignment of InAs/GaSb heterostructures, first observed by Sakaki *et al.* in 1977.<sup>192</sup> The bottom of the conduction band of InAs is about 150 meV below the top of the valence band of GaSb,<sup>191</sup> resulting in charge transfer between GaSb and InAs layers and an inherent electric field at the interface, which allows for the tuning of the band structure with external electric fields.<sup>189</sup> The valence band of AlSb is about 0.4 eV lower than that of GaSb, while its conduction band is about 0.4 eV higher. This leads to a very large conduction band offset between InAs and AlSb of about 1.35 eV,<sup>193</sup> making AlSb a very good quantum well barrier to confine both electrons in InAs and holes in GaSb layers.

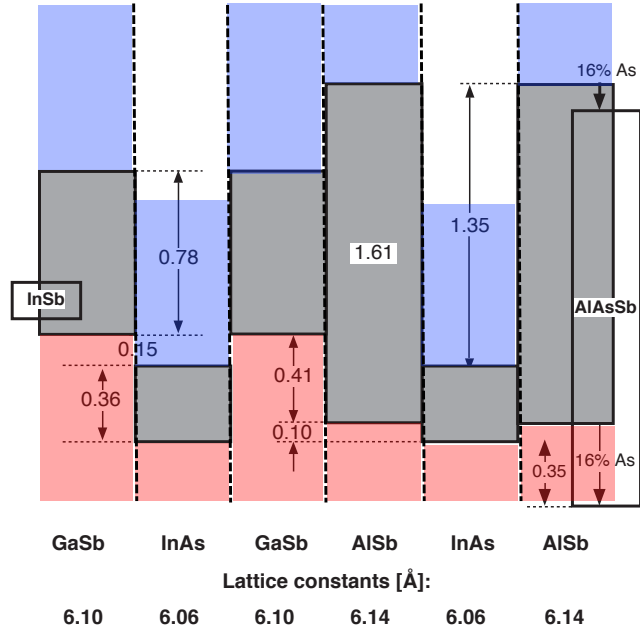


Figure 6.1: [Adapted from Ref. 191] Band lineups for the 6.1 Å family of semiconductors. Dark shaded areas represent energy gaps.

Figure 6.2a is an example of an InAs/GaSb composite quantum well (CQW) structure with AlSb as the well barriers. As a consequence of the unique band alignment of InAs/GaSb/AlSb and the negative effective mass of holes, the electron subband (InAs) and the hole subband (GaSb) are localized to two different quantum wells in immediate proximity. The broken band gap alignment should imply equal densities of both electron and hole gases in the inverted CQW. Due to a high density of surface states, however, the Fermi level is typically pinned at around 130 meV above the bottom of the bulk conduction band of InAs.<sup>194</sup> Therefore, metal-InAs interfaces do not form significant electron-blocking Schottky barriers.<sup>191</sup> Also, depending on the CQW width parameters, gating is required to induce holes into the GaSb layer.

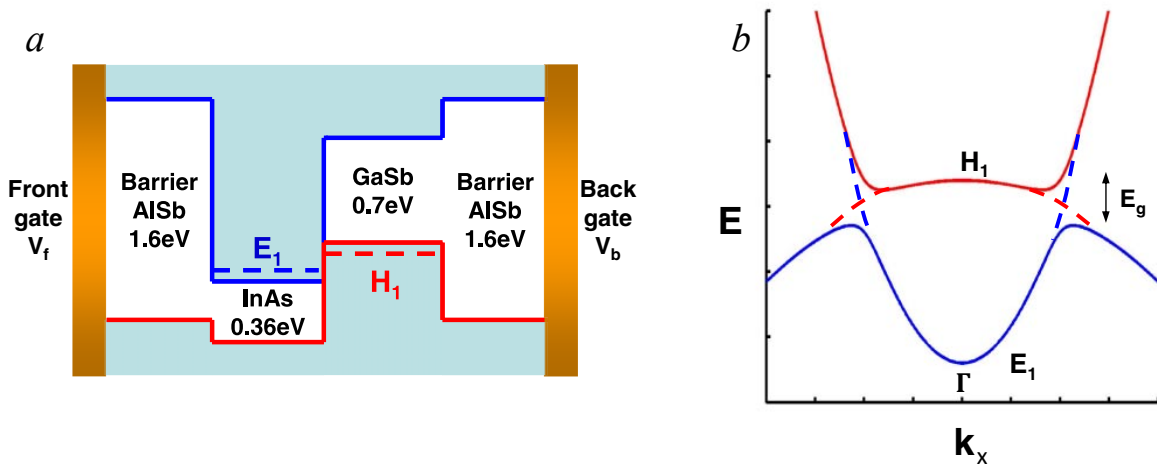


Figure 6.2: [Adapted from Ref. 140] a) Composite quantum well of InAs/GaSb with AlSb barriers. b) Schematic band structure diagram of InAs/GaSb. The dashed line indicates the crossing of  $E_1$  and  $H_1$  in the inverted regime at  $k_{cross}$ . The gap opened due to the hybridization of the bands is shown as  $E_g$ .

Much like in the previously discussed HgTe QWs, the width of the InAs/GaSb CQW defines two different regimes for the system: the normal regime, where the valence band lies below the conduction band, and the inverted regime, where the band order with respect to energy is reversed. In the infinite quantum well approximation, the energy difference from the bottom of the quantum well to bottom of the first electron or heavy-hole subband ( $E_1$  and  $H_1$ , respectively) is inversely proportional to square of the quantum well width and the electron or hole effective mass. It follows, then that as the CQW becomes narrower, the  $E_1$  subband is lifted up and the  $H_1$  subband is pushed down. Thus, when the QWs are below some critical thickness, the InAs electron subband ( $E_1$ ) is pushed above the GaSb heavy-hole subband ( $H_1$ ), resulting in a normal narrow-gap semiconductor structure. If the CQW thickness is increased, however, the  $E_1$  band will be lowered while  $H_1$  band rises. After some critical thickness,  $E_1$  will sink below

$HI$ , and the system will be in the inverted state. Since  $EI$  disperses upward and  $HI$  disperses downward, this naturally leads to an intersection point  $k_{cross}$  between the two bands (*Figure 6.2b*).

The inverted regime is especially interesting for use in quantum spin Hall insulator studies due to the coexistence of electron and hole quantum wells. Though the inverted state was initially thought to be semimetallic, M. Altarelli theoretically demonstrated that due to the mixing of  $EI$  and  $HI$ , a small hybridization gap opens, resulting in bulk insulating behavior characteristic of QSHIs.<sup>195</sup> Experimental evidence of this hybridization gap was demonstrated by Yang *et al.*<sup>196</sup> and Lakrimi *et al.*<sup>197</sup> Due to the proximity of the QWs, a real nonzero coupling constant allows for penetration of the electronic wavefunction into the hole layer. Effective mass theory<sup>198</sup> implies this penetration is small, such that when the in-plane momentum and carrier energy in the two wells are nearly equal, they will hybridize. The electron and hole states, then, are indistinguishable. The hybridization lifts the degeneracy at the point in  $k$ -space where the dispersion relations of the electrons and holes cross,  $k_{cross}$ , and an anti-crossing of the two branches results in a “mini-gap”  $\Delta$  on the order of 2-5 meV.<sup>189,195,196,199–202</sup> The size of the gap depends on the strength of the coupling between the electrons and holes.

### 6.1.1 Electron-Hole Hybridization

The simplest model of the electron-hole hybridization involves a two-band Hamiltonian with uncoupled electron,  $|\psi_e\rangle$ , and hole states,  $|\psi_h\rangle$ , as the basis:

$$H = \begin{pmatrix} E_e & V(k) \\ V(k)^* & E_h \end{pmatrix}, \quad (6.1)$$

where  $E_e = \frac{\hbar^2 k^2}{2m_e}$  and  $E_h = E_{g0} - \frac{\hbar^2 k^2}{2m_h}$  are the uncoupled electron and hole energies, respectively, and  $E_{g0}$  is the relative separation between the *EL* and *HL* subbands. In this simplified approach, we ignore the effects of the self-consistent potential, which would only shift the band edges of the quantum well, and therefore to lowest order only affect the value of  $E_{g0}$ .<sup>203</sup>

Coupling between the wells is incorporated via the off-diagonal term  $V(k)$ . Tunneling is only allowed between states with the same angular momentum; in other words, the states must have the same symmetry. Therefore the heavy-hole state ( $J = \frac{3}{2}, m_J = \pm \frac{3}{2}$ ) at the center of the subband cannot hybridize with the electron states ( $J = \frac{1}{2}, m_J = \pm \frac{1}{2}$ ). Away from the center of the band at some finite in-plane wavevector however, the heavy-holes are mixed with the light-holes component ( $J = \frac{3}{2}, m_J = \pm \frac{1}{2}$ ), making hybridization with electrons possible.<sup>204,205</sup> The coupling is further restricted by parity conservation. The parity operator returns the spatial mirror image of a variable; variables that are unchanged under this transformation are said to have even parity, and those that flip sign are said to have odd parity. The electron states come from *s*-orbitals (even), while the hole states are from spin-orbit coupled *p*-orbitals (odd) with  $p_x + ip_y$  rotational symmetry about the growth axis of the wells. Since the states have opposite parity, the states can only be coupled through an operator that is odd under space inversion. This implies, to second order, the off-diagonal coupling terms must be linear in  $k$ . It was shown by Altarelli and Laikhtman *et al.* that these conditions require the off-diagonal terms to be of the form  $V(k) = w(k_x + ik_x)$ , where  $w$  is a constant that characterizes the coupling between the two layers.<sup>195,204</sup> The final two-band Hamiltonian is then

$$H = \begin{pmatrix} \frac{\hbar^2 k^2}{2m_e} & w(k_x + ik_x) \\ w(k_x - ik_x) & E_{g0} - \frac{\hbar^2 k^2}{2m_h} \end{pmatrix}, \quad (6.2)$$

with energies<sup>206</sup>

$$E_{1,2}(k) = \frac{E_e + E_h}{2} \pm \sqrt{\left(\frac{E_e - E_h}{2}\right)^2 + w^2 k^2} \quad (6.3)$$

and eigenstates

$$|\psi_{1,2}\rangle = \cos \theta |\psi_{e,h}\rangle + i \sin \theta |\psi_{h,e}\rangle, \quad (6.4)$$

where  $\sin 2\theta = \frac{wk}{\sqrt{(E_e - E_h)^2 + w^2 k^2}}$ .

In the limit that the difference in the electron and hole energies is much larger than the strength of the coupling ( $E_e - E_h \gg wk$ ), the eigenstates represent two uncoupled quantum wells. In other words, for in-plane momentum values far from the crossing point ( $k_{cross}$ ), tunneling does not affect the electron or hole bands. In this case, the longitudinal conductivity can be described by the classical Drude model,  $\sigma_{xx} = ne\mu_e + pe\mu_h$ , where  $n(p)$  is the free electron(hole) density, and  $\mu_e(\mu_h)$  is the electron(hole) mobility. When the energies of carriers in both wells are matched ( $E_e = E_h$ ), however, the two bands anticross at the in-plane momentum value  $k_{cross} = \sqrt{\frac{E_{g0} m^*}{2\hbar^2}}$ , where  $m^* = \frac{m_e m_h}{m_e + m_h}$  is the reduced mass, and  $E_{g0}$ , the overlap between the bands, is positive. In the absence of any disorder, the band structure is then gapped out due to hybridization, and the bulk mini-gap conductivity will vanish in the low temperature limit.

The energy dispersion from Equation 6.3 can be used to find the density of states of the system,  $n_{2d} = \frac{k}{\pi} \left(\frac{\partial E}{\partial k}\right)^{-1}$  (Figure 6.3a). Within the hybridization gap, the density of states vanishes, so the expected bulk conductivity of the system in the zero-temperature limit would be fully suppressed to zero. At the edges of the mini-gap, though, due to the non-monotonic dispersion of the system (Figure 6.2, Figure 6.3b), the hybridized energy bands have local extrema at finite wavevector values. This leads to van Hove singularities, points at which  $\frac{\partial n_d}{\partial E}$



diverges, at the mini-gap edges. It is worth noting that this model was derived for an idealized case without disorder. Disorder effects can lead to a reconstruction of the density of states singularities and to a finite population of states within the gap.

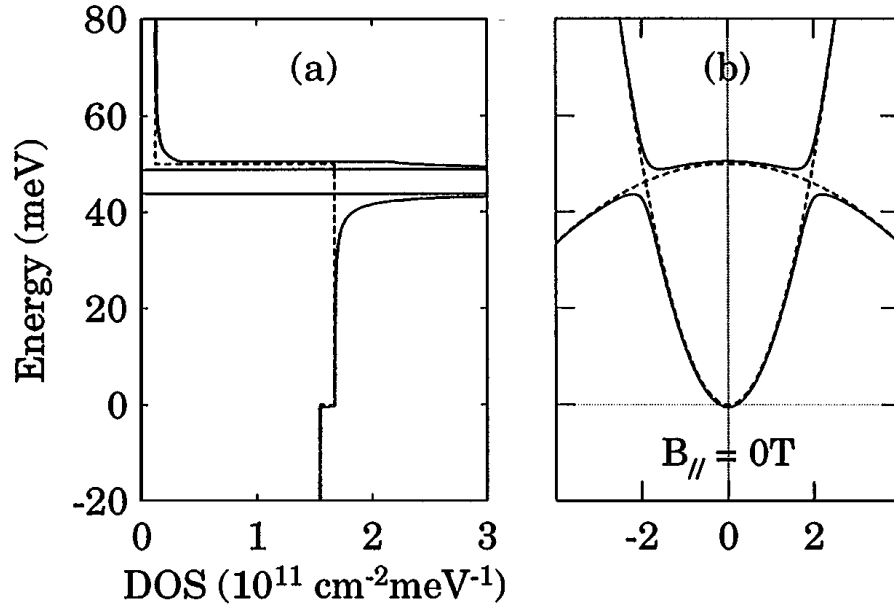


Figure 6.3: [Adapted from Ref. 201] Density of states (a) and in-plane dispersion relations (b) of electrons in InAs and holes in GaSb with (solid lines) and without (dashed lines) consideration of the hybridization effect. Density of states vanishes within the hybridization gap and exhibits van Hove singularities at the gap edges.

### 6.1.2 Tuning the Band Structure in InAs/GaSb Composite Quantum Wells

As mentioned above, one of the appealing features of InAs/GaSb CQWs is the tunability of the band structure by perpendicular electric fields.<sup>189,207</sup> Electric fields are typically generated by applying gate voltages to the top and/or bottom of the sample. The band bending due to the applied electric field follows  $\mathbf{F} = \frac{1}{e}\nabla E$ , where  $E$  is the band energy. Consider applying an electric field  $\mathbf{F}$  in the  $+z$  direction in the system of Figure 6.2. The  $E_I$  band will shift

downward, while the *HI* band will shift upward, increasing the separation between the subbands  $E_{g0}$  and pushing the anticrossing point  $k_{cross}$  to higher values of  $k$ . Applying an electric field in the  $-z$  direction will have the opposite effect.

For small electric fields, the addition of the field can be treated as a perturbation, with the first order energy correction  $\langle \psi | V(z) | \psi \rangle$ , where  $V(z) = eFz$  for electrons and  $V(z) = -eFz$  for holes. Thus the electron and hole subbands will shift in opposite directions by  $\Delta E = \pm eF \langle z_{e,h} \rangle$ , respectively. *Figure 6.4* shows the dispersion of *EI* and *HI* for different perpendicular electric fields in a theoretical 17 nm InAs/ 4.85 nm GaSb double-gated CQW proposed by Naveh and Laikhtman.<sup>189</sup> This demonstrates that, in principle, it is possible to switch between the normal and inverted band regime. Furthermore, the calculated fields required to achieve switching are readily experimentally accessible.<sup>189</sup> For example, the switching in *Figure 6.4* occurs for fields on the order of 100 kV/cm, or 1 V/kÅ. This means for a reasonable dielectric thickness on the order of 1 kÅ, switching can be achieved for applied gate biases on the order of 1 V.

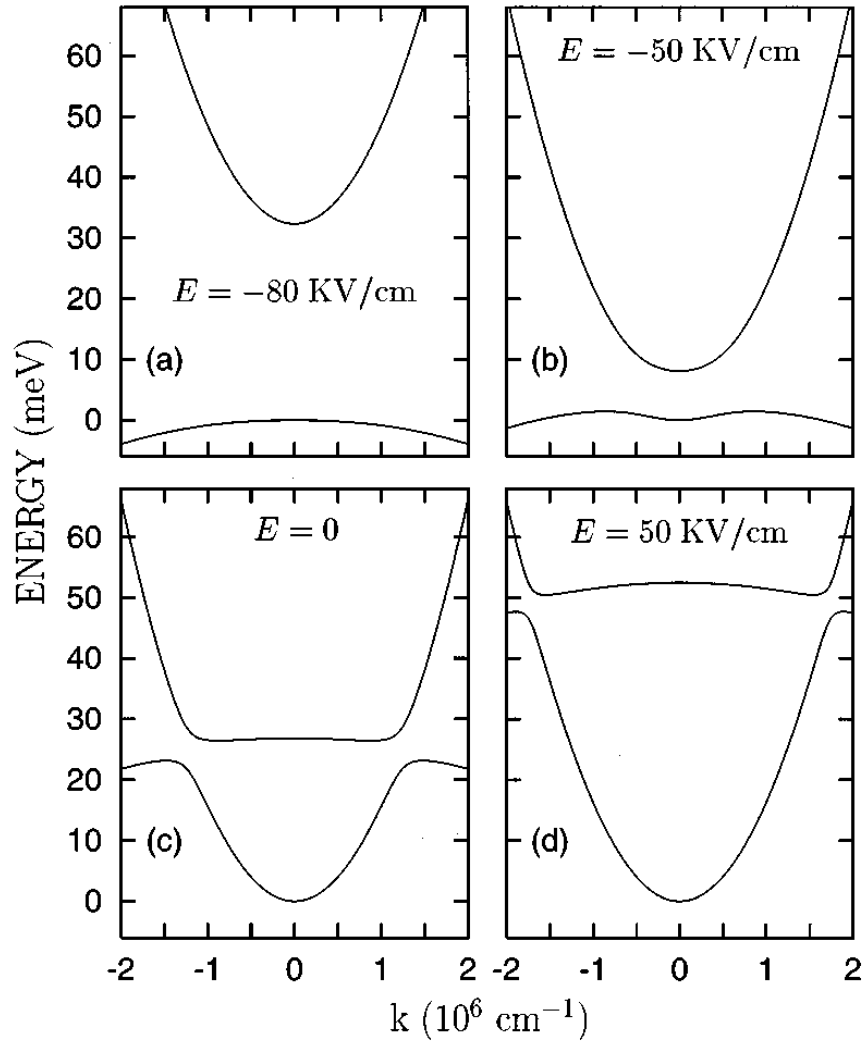


Figure 6.4: [Adapted from Ref. 189] In-plane band structure of a double-gated InAs/GaSb composite quantum well under different external electric fields. The system can be tuned between the normal and inverted regime by applying different gate voltages.

In addition to tuning the band structure, applied electric fields can also be used to alter the Fermi level by changing the carrier density in the wells. Using the model of charging a capacitor, the carrier density will change as  $\Delta n = \frac{\varepsilon}{ed} \Delta V$ , where  $\varepsilon$  is the dielectric constant,  $d$  is

the dielectric layer thickness, and  $\Delta V$  is the applied bias. The Fermi level is usually pinned above the *HI* band, but with gating, it can be shifted to sit between the *HI* and *EI* subbands to create a regime of strongly coupled electron-hole systems. When  $E_F$  is between *HI* and *EI*, electrons and holes coexist in their respective layers, resulting in two-carrier transport.<sup>196</sup> When  $E_F$  is above *HI* or below *EI*, single-carrier transport occurs, electron-like or hole-like, respectively.<sup>208</sup>

To tune both the band structure and the Fermi level, a double-gated system is necessary. The absolute value of the gate biases relative to the potential of source/drain contacts sets the position of the Fermi level relative to that of the contacts, while the relative difference between the biases sets the positions of the subbands. Double-gated CQWs of InAs/GaSb were first demonstrated experimentally by Cooper *et al.* who observed peaks in the longitudinal resistance as a function of gate voltage when the derived carrier densities were nearly equal,  $n = p = \frac{k_{cross}^2}{2\pi}$ , corresponding to the resonant condition of equal energy and in-plane momentum in the two wells, and therefore indicating the presence of a hybridization gap.<sup>199</sup> Using the temperature dependence of the resistance peak, the group deduced the hybridization gap in their structures was approximately 2 meV. Alternatively, the size of the gap can be determined from the relative dependence of the resonant resistance peak with front gate for a constant back gate voltage (*Figure 6.6* for example).<sup>141</sup> The resistance maximum corresponds to the middle of the gap, while the resistance dip corresponds to the van Hove singularity at the gap edge. The gap size can then be found by

$$\Delta = 2(V_{peak} - V_{dip}) \frac{\Delta n}{\Delta V} \frac{1}{DOS}, \quad (6.5)$$

where  $\frac{\Delta n}{\Delta V}$  is the rate of carrier change with  $V_{front}$ , and  $DOS = \frac{(m_e + m_h)}{\pi \hbar^2}$  is the density states.

## 6.2 Theoretical Prediction of the Quantum Spin Hall Effect in Inverted Type-II Semiconductors

Liu *et al.* were the first to model the quantum Spin Hall effect in Type-II semiconductor quantum wells, such as InAs/GaSb.<sup>140</sup> With the inverted band structure and consequent hybridization gap, one should expect the inverted regime in InAs/GaSb QWs to be a topologically nontrivial QSHI phase protected by the bulk gap. This assumption is complicated by the fact that the electron and hole subbands lie in two separate layers. Liu *et al.* derive several consequences of this. First, the coupling strength between  $E1$  and  $H1$  is reduced relative to the HgTe case, but this can be incorporated as a quantitative correction. Additionally, since there is no inversion symmetry in the QW growth direction, structural inversion asymmetry (SIA) may be large enough to compete with the hybridization. Both SIA and bulk inversion asymmetry (BIA) must be included for InAs/GaSb system calculations, whereas for HgTe QWs, these terms may be neglected because BIA corrections are small relative to the gap, and the QW is symmetric, minimizing SIA. Lastly, the spatial separation of the electron and hole layers results in an automatic charge transfer between the layers, yielding a coexistence of  $p$ -type and  $n$ -type carriers. This requires a self-consistent treatment of Coulomb energy.

These systems are well described by the 8-band Kane model,<sup>209</sup> and Liu *et al.* use a 4-band model effectively equivalent to the Bernevig-Hughes-Zhang (BHZ) approach discussed in Section 3.5, though with the addition of terms for the SIA and BIA. The Hamiltonian is then

$$H = H_0 + H_{BIA} + H_{SIA}. \quad (6.6)$$

In the basis  $\{|E1 +\rangle, |H1 +\rangle, |E1 -\rangle, |H1 -\rangle\}$ , up to quadratic terms in  $\mathbf{k}$ ,

$$H_0 = \epsilon(k)\mathbf{I}_{4 \times 4} + \begin{pmatrix} \mathcal{M}(k) & Ak_+ & 0 & 0 \\ Ak_- & -\mathcal{M}(k) & 0 & 0 \\ 0 & 0 & \mathcal{M}(k) & -Ak_- \\ 0 & 0 & -Ak_+ & -\mathcal{M}(k) \end{pmatrix}, \quad (6.7)$$

where  $\mathbf{I}_{4 \times 4}$  is the  $4 \times 4$  identity matrix, and  $\epsilon(k)$  describes band bending. The elements,  $\mathcal{M}(k)$ , are defined as

$$\mathcal{M}(k) = M_0 + M_2 k^2, \quad (6.8)$$

where  $M_0 = -\frac{E_{g0}}{2}$  corresponds to the gap between bands and is negative in the inverted regime, and  $M_2$  relates to the curvature of the bands and corresponds to the inverse of the effective mass. The  $A$  factor incorporates interband coupling to the lowest order, and  $k_{\pm} = k_x \pm ik_y$ .

Without the  $H_{BIA}$  and  $H_{SIA}$  terms, the Hamiltonian is essentially an extension of the two-band model of Equation 6.2 to include spin.  $H_0$  is block diagonal, and each block is exactly a massive Dirac Hamiltonian in  $(2+1)$  dimensions. The rest mass lies on the diagonal, with opposite mass for electrons and holes, and the off-diagonal elements are linear in momentum and couple electrons and holes of the same spin. This then matches the BHZ model, and we should therefore expect a topological phase transition at the gap-closing condition  $M_0 = 0$ . While the addition of  $H_{BIA}$  and  $H_{SIA}$  complicates the picture, we know that the QSH state is a topological phase protected by the band gap. Therefore, if we begin with the Hamiltonian  $H_0$  in the QSH phase and adiabatically turn on  $H_{BIA}$  and  $H_{SIA}$ , then the system will remain in the QSH phase as long as the energy gap between  $EL$  and  $HL$  does not close.

In contrast to HgTe QWs where the carriers are in the same layer and the gap lies at the zone center, in InAs/GaSb CQWs, the energy gap opens at some finite  $k_{cross}$  in momentum space where the ground electron subband in the InAs well and the ground hole subband in the GaSb well are matched in energy and in-plane momentum. This hybridization gap is the bulk gap of

the quantum spin Hall insulator. The charge carrier density at the charge neutral point (CNP) in the inverted regime is defined as  $n_{cross} = \frac{k_{cross}^2}{2\pi}$ , and can be used to characterize the degree of band inversion, which in turn indicates the nature of the bulk states. For  $n_{cross} \gtrsim 2 \times 10^{11} \text{ cm}^{-2}$ , the system is considered to be deeply inverted, and there will be considerable residual states within the hybridization gap. The result is that the bulk of the QW is not truly insulating, yielding the system an ineffective QSHI.<sup>141,208,210,211</sup> When  $n_{cross} \lesssim 1 \times 10^{11} \text{ cm}^{-2}$ , however, the system is in the shallowly inverted regime, and the bulk is highly insulating, allowing for a true QSHI state.<sup>144,212</sup>

To confirm the quantum spin Hall state in the InAs/GaSb CQWs, Liu *et al.* studied the edge state spectrum. Using a tight-binding model with cylindrical geometry, they confirmed there was one Kramers pair of edge states with opposite spin per edge in the QSHI phase and no edge states in the normal insulator phase. The group then numerically solved the full 8-band Kane model, taking into account the built-in electric field between the two QW layers, and deduced gap behavior as a function of QW thickness and applied front and back gate voltages. They found that as a function of front and back gate voltages, there were six parameter regions, as seen in *Figure 6.5*. In regions IV, V, and VI (blue), the system is in the normal regime. In regions I, II, and III (red), the system has an inverted band structure, but is only in the QSH state in region II, when the Fermi level sits within the bulk gap. Region I(III) has roughly the same spectrum as region II, but with finite  $p(n)$ -doping. Therefore, it is possible to continuously tune the system into and out of the QSHI phase using applied gate voltages. Compared to other proposals of gate-induced phase transitions in asymmetric HgTe/CdTe QWs, InAs/GaSb CQWs are much more sensitive to the gate voltage due to the fact that the electron and hole

wavefunctions are centered in separate layers, making the effect of the gate voltage highly asymmetric.

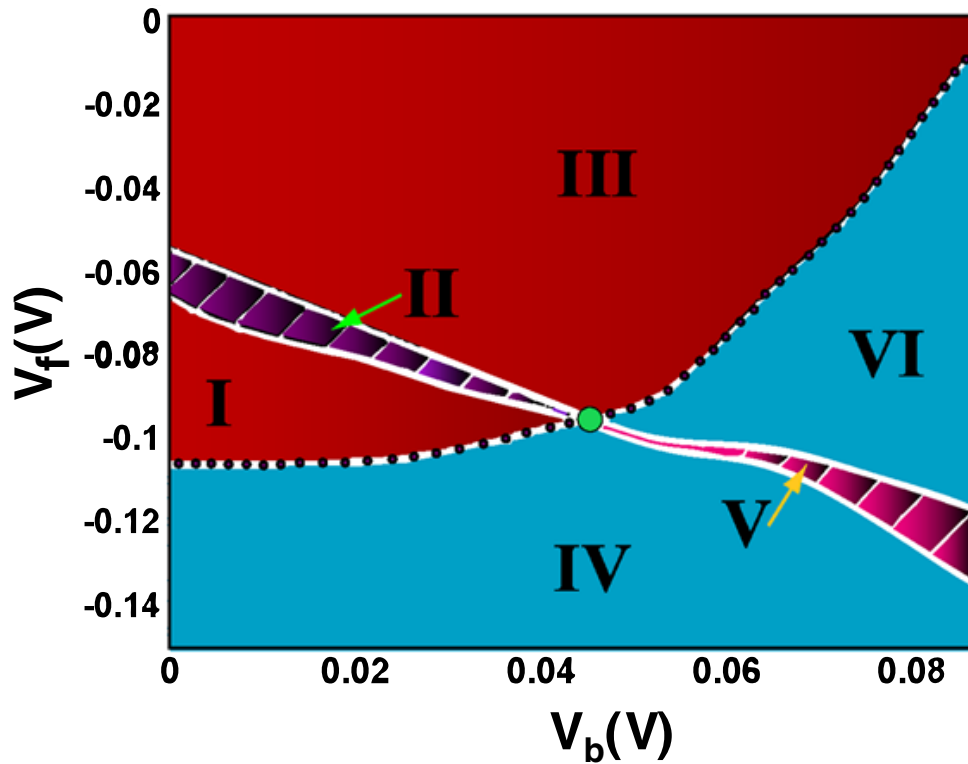


Figure 6.5: [Adapted from Ref. 140] Phase diagram for InAs/GaSb CQW as a function of front ( $V_f$ ) and back ( $V_b$ ) gates. Regions I, II, and III are in the inverted regime. The purple striped area in region II is the quantum spin Hall phase with the Fermi-level in the hybridized gap. Region I(III) is the p(n)-doped inverted system. Regions IV, V, and VI are in the normal regime. The purple striped area in region V is the normal insulator phase with the Fermi level in the bulk gap, and region IV(VI) is the p(n)-doped normal semiconductor. The black dotted line indicates the phase boundary between the inverted and normal band structure, and the green circle denotes the quantum critical point between the NI and QSH phases.



### 6.3 Experimental Evidence of the Quantum Spin Hall State in InAs/GaSb

Prior to the formulation of the QSH phase in InAs/GaSb CQWs, the first experimental evidence of the mini-gap due to electron-hole hybridization was demonstrated by Yang *et al.* via capacitance and transport measurements.<sup>196</sup> Further evidence of the hybridization gap was seen in far-infrared measurements by Kono *et al.*<sup>200</sup> and later by Yang *et al.*<sup>201</sup> Additionally, Cooper *et al.*<sup>199</sup> observed strong peaks in the longitudinal resistance of double-gated structures corresponding to the hybridization gap. In both the capacitance studies by Yang and the transport measurements by Cooper, the groups observed signs of localized states within the hybridization gap. Naveh and Laikhtman considered the finite gap conductivity a bulk effect related to a small amount of dissipative tunneling between the wells.<sup>202</sup> Previously, Caldeira and Leggett,<sup>213</sup> in studying the influence of dissipation on quantum mechanical tunneling in macroscopic systems, found that if a particle is coupled to external degrees of freedom, its tunneling probability will be reduced. Therefore, dissipative tunneling yields some number of non-hybridized electronic states, resulting in a nonzero density of states in the gap and finite gap conductivity. Even a slight amount of broadening of the electron and hole bands can lead to significant changes in transport behavior, yielding finite bulk conductivity even in the zero temperature limit and in the limit of level broadening much smaller than the gap size.<sup>202</sup>

In 2010, Knez *et al.* experimentally studied the origin of the residual conductivity at low temperatures in dual-gated InAs/GaSb CQWs.<sup>208</sup> Their findings were consistent with the model of Naveh and Laikhtman; namely that the residual conductivity could be explained by contributions from both free and hybridized carriers in the presence of impurities. As a function of front gate bias, they observed clear peaks in the longitudinal resistance (*Figure 6.6*) at the point where the electron and hole densities are matched, indicating the existence of a

hybridization gap. The temperature dependence of the resonance peak was consistent with a mini-gap value of 3-4 meV. Even at  $T = 20$  mK, however, the largest measured resistance was only 5 k $\Omega$ , less than half of the  $\frac{h}{2e^2}$  seen in HgTe QWs.<sup>121</sup> The group therefore attributes the resonance peaks to a bulk effect, with residual conductivities on the order of  $\frac{10e^2}{h}$ , a few times larger than the expected contribution from the edge. Naveh and Laikhtman predicted the residual gap conductivity should go as  $\sigma(T = 0) \sim \frac{e^2 E_{g0}}{h \Delta}$ , for the condition  $\Gamma \ll \Delta \ll E_{g0}$ , where  $\Gamma$  is the level broadening,  $\Delta$  is the hybridization gap, and  $E_{g0} = n \frac{\pi \hbar^2}{m^*}$  is the relative separation of the two bands.<sup>202</sup> Using this equation, Knez *et al.* derives a residual conductivity of  $\sim \frac{7e^2}{h}$ , consistent with their measurements. They also found hybridized states will contribute to residual gap conductivity on the order of  $\frac{e^2 E_{g0}}{h \Delta}$  due to level-broadening-induced charge of hybridized states.

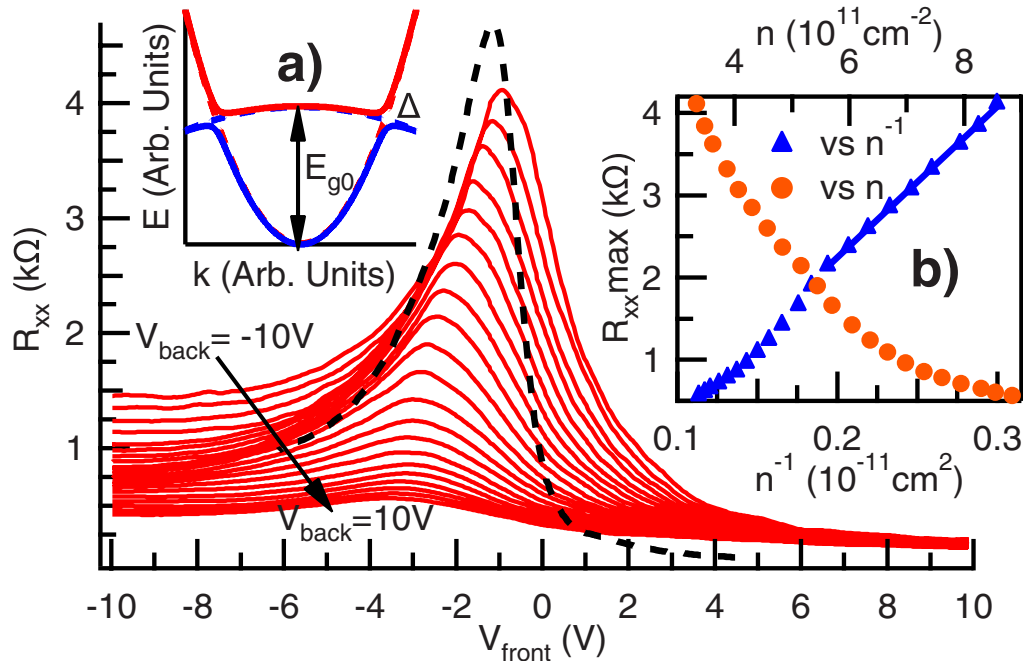


Figure 6.6: [Adapted from Ref. 208] Peaks observed in the longitudinal resistance  $R_{xx}$  as a function of  $V_{front}$  in InAs/GaSb CQW with  $V_{back}$  between -10 V and 10 V at  $B = 0$  T and  $T = 0.3$  K. Inset (a) dispersion of the hybridized bands. At the anticrossing point, where the electron and hole densities are approximately equal, a hybridization gap is opened. Inset (b) The resonance peak heights as a function of the inverse carrier density. For  $n \lesssim 5 \times 10^{11} cm^{-2}$ , the peaks vary linearly with  $n^{-1}$ .

A second study by Knez *et al.* in 2011<sup>141</sup> demonstrated evidence for the quantum spin Hall state in InAs/GaSb CQWs in direct agreement with the predictions of Liu *et al.*<sup>140</sup> They found that due to the gap opening away from the zone center, edge and bulk states are decoupled, leading to a large disparity in their Fermi wavevectors and the ability for edge transport to persist even in the limit of a conductive bulk. The group used a four-terminal device structure, so that

the expected edge conduction is  $\frac{4e^2}{h}$ , with device lengths between 2 and 100  $\mu\text{m}$ . In longer devices, the four-terminal conductance, including dephasing effects,<sup>8,73</sup> is predicted to be

$$G_{14,23} = \frac{2e^2}{h} \left[ \frac{l_\phi}{L} + \left( \frac{l_\phi}{L} \right)^2 \right], \quad (6.9)$$

where  $l_\phi$  is the phase coherence length and  $L$  is the device length. *Figure 6.7a* shows the resistance peaks at  $L = 100, 10, 4,$  and  $2 \mu\text{m}$  at 300 mK. The resistance peak of the 100  $\mu\text{m}$  device is used to estimate the bulk gap resistance,  $R_{bulk} \sim 10.2 \text{ k}\Omega$ , and using a parallel combination of  $R_{bulk}$  with the expected edge resistance,  $\frac{h}{4e^2}$ , gives a resistance  $R_{bulk} \parallel \frac{h}{4e^2} \sim 3.95 \text{ k}\Omega$ , which is very close to the peak resistance of  $3.75 \text{ k}\Omega$  for  $L = 2 \mu\text{m}$ . Additionally, fitting  $G\left(\frac{1}{L}\right)$  to Equation 6.9 yields  $l_\phi \sim 2 \mu\text{m}$ , providing further evidence for the presence of helical edge states in the shortest samples.

As a function of sample width  $W$  (*Figure 6.7b*) for  $L = 2 \mu\text{m}$ , resistance peaks tend to increase as  $W$  is decreased. *Figure 6.7c* of  $G(W)$  shows a relatively linear relationship, with the intercept  $G_{edge} = (4.08 \pm 0.69) \frac{e^2}{h}$ , supporting the existence of helical edge transport. The slope of the same fit yields bulk conductivity  $g_{bulk} = (5.46 \pm 1.01) \frac{e^2}{h}$ , which is consistent with the expected bulk conductivity<sup>202</sup>  $g_{bulk} \sim \frac{e^2 E_{g0}}{h \Delta} \sim \frac{4e^2}{h}$ , where  $E_{g0}$  is found from the front gate dependence of  $\frac{B}{R_{xy}}$ . The minimum in  $\frac{B}{R_{xy}}(V_{front})$  corresponds to the anticrossing carrier density, and  $E_{g0} = n_{cross} \frac{\pi \hbar^2}{m^*}$ . Therefore, both the length and width dependence point to the existence of helical edge states.

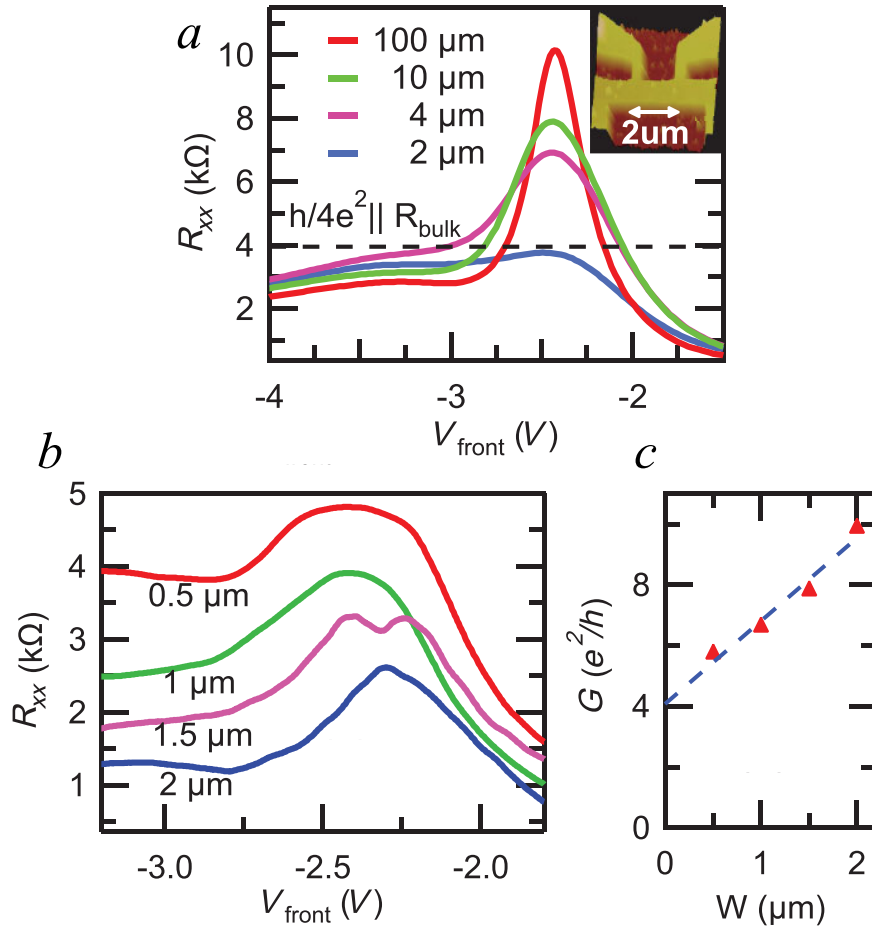


Figure 6.7: [Adapted from Ref. 141] a) Longitudinal resistance  $R_{xx}$  vs. front gate voltage  $V_{front}$  for various device lengths  $L = 100, 10, 4,$  and  $2 \mu\text{m}$ . Width is varied to give a constant ratio  $\frac{L}{W} = 2$ . Resistance peaks decrease as device length shortens and approach the limit  $R_{bulk} \parallel \frac{h}{4e^2}$  (dashed line) for the  $2 \mu\text{m}$  device. b)  $R_{xx}$  vs.  $V_{front}$  for device widths  $W = 0.5, 1, 1.5,$  and  $2 \mu\text{m}$  with  $L = 2 \mu\text{m}$ . Resistance peaks decrease with increasing  $W$ . c) The gap conductance  $G$  shows a linear relationship with the device width. The intercept of the linear fit gives  $G_{edge} = (4.08 \pm 0.69) \frac{e^2}{h}$ , consistent with helical edge transport, and the slope of the fit gives the bulk conductivity  $g_{bulk} = (5.46 \pm 1.01) \frac{e^2}{h}$ .

The data of *Figure 6.8a* ( $g_{bulk} \lesssim 5 \frac{e^2}{h}$ ) may suggest the edge conduction is independent of gap bulk conductivity in the regime of low  $g_{bulk}$ . The degree of bulk conduction can be tuned using a bias cooling technique<sup>208</sup> to push the sample deeper into the inverted regime, increasing the value of  $E_{g0}$ . The edge conduction, estimated as  $\Delta G = G_{2\mu m} - G_{100\mu m}$ , ranges from  $\Delta G \sim 0$  for large  $g_{bulk} \sim 19 \frac{e^2}{h}$  to  $\Delta G \sim 3 \frac{e^2}{h}$  when bulk conductivity has dropped to  $g_{bulk} \lesssim 5 \frac{e^2}{h}$  (*Figure 6.8d*). The resilience of the edge conduction to bulk transport was unexpected, given that naively, one would expect a conductive bulk to enable the tunneling of electrons between edges and thus a reduced edge conductivity.<sup>214,215</sup> A large mismatch in the Fermi wavevectors of the bulk and edge states, however, may greatly reduce the tunneling probability.<sup>141</sup> The bulk gap states follow from the nonhybridized band structure, with Fermi wavevector  $k_{bulk} = k_{cross} \gg 0$ . The edge modes, for  $E_F$  in the middle of the gap, have  $k_{edge} \sim 0$ . It follows, then, that because  $k_{edge} \ll k_{cross}$ , edge modes are completely reflected by the bulk modes. The tunneling probability for edge electrons is proportional to the edge-bulk transmission probability, which scales as  $\frac{k_{edge}}{k_{cross}}$ , in addition to the bulk transmission, which scales as  $g_{bulk} \sim E_{g0} \propto k_{cross}^2$ . Therefore, the overall interedge tunneling should decrease as  $k_{cross}$  is reduced, which is qualitatively consistent with *Figure 6.8d*.

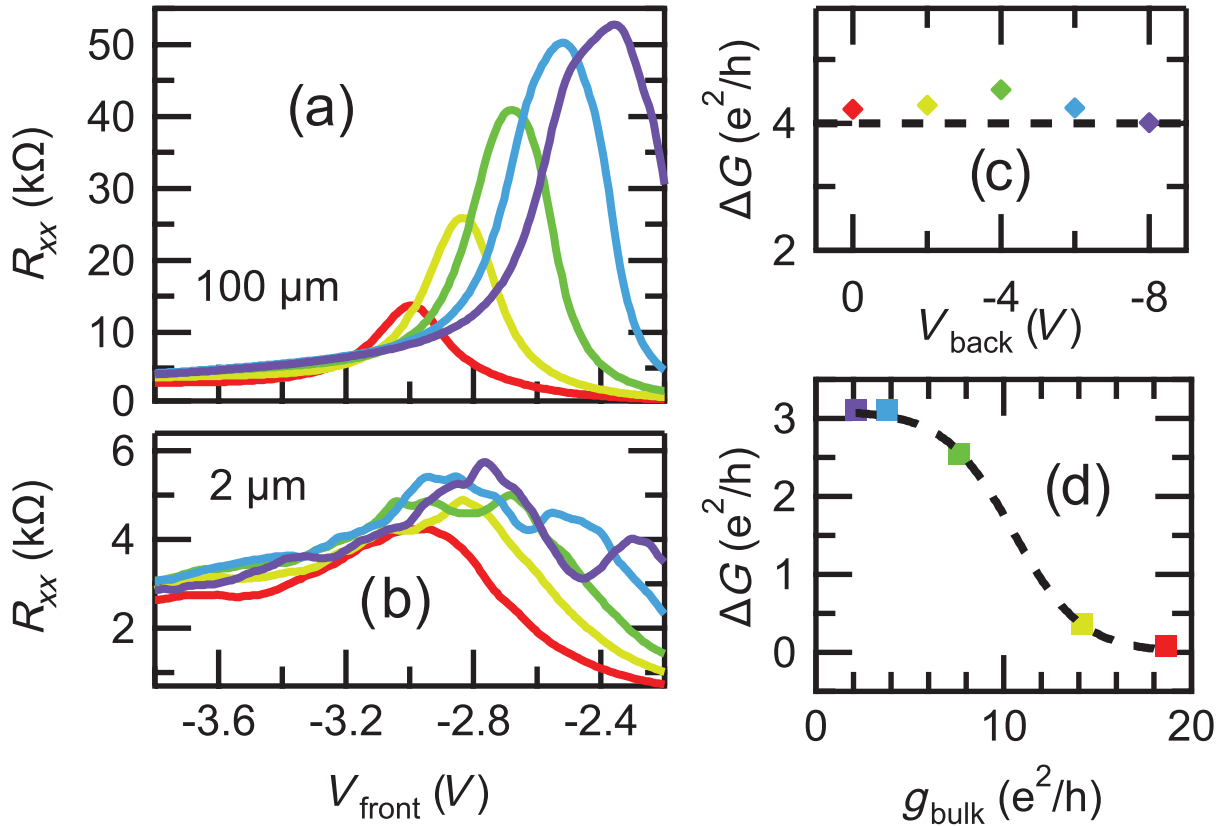


Figure 6.8: [Adapted from Ref. 141]  $R_{xx}$  vs.  $V_{front}$  for  $L = 100 \mu\text{m}$  (a) and  $2 \mu\text{m}$  (b) as  $V_{back}$  is varied in steps of  $2 \text{ V}$  from  $0$  to  $-8 \text{ V}$ , with  $B = 0 \text{ T}$ , and  $T = 20 \text{ mK}$ . As  $V_{back}$  is tuned to more negative values, the gap shifts to smaller values of  $k_{cross}$ , and the resistance peaks increase. The difference in the gap conductance,  $\Delta G$ , is shown as a function of  $V_{back}$  with  $g_{bulk} \lesssim 5 \frac{e^2}{h}$  (c) and bulk conductivity for a bias cooled sample with a larger bulk conduction (d). Edge conduction “activates” for  $g_{bulk} \lesssim 10 \frac{e^2}{h}$ .

Another study by Qu *et al.* explored the full phase diagram put forth by Liu *et al.* (Figure 6.9a,b).<sup>216</sup> They demonstrated the ability to continuously tune between the trivial and topological insulating phases using dual gates in InAs/GaSb CQWs. The group found two

gapped regions, observed as peaks in the longitudinal resistance (*Figure 6.9c*). In one gapped region, the derived electron and hole densities both vanish at the resistance peak, which is typical for a trivial insulator with a normal band gap (*R* in *Figure 6.9*). In the second gapped region, there are finite equal electron and hole densities in the vicinity of the resistance peak, indicative of inverted band alignment and the opening of a hybridization gap at the crossing of the two bands (*L* in *Figure 6.9*). The nature of the two regions was further evidenced by applying an in-plane magnetic field to shift the electron and hole bands relative to each other in momentum space (*Figure 6.9d*). The field, therefore, is expected to reduce the hybridization gap, but not the normal gap. The group observed a decrease in the peak height when the sample was gated into the region expected to be the topological insulator phase, as well as a shift in the peak position. The measured in-plane field dependence is consistent with a relative shift of the two bands in momentum for the inverted regime. In contrast, when the in-plane field was increased for the region expected to be a normal insulator, the resistance remained unchanged up to a field of 9 T, despite the same predicted band shifting as the TI region. This further suggests this region characterizes a normal insulator.



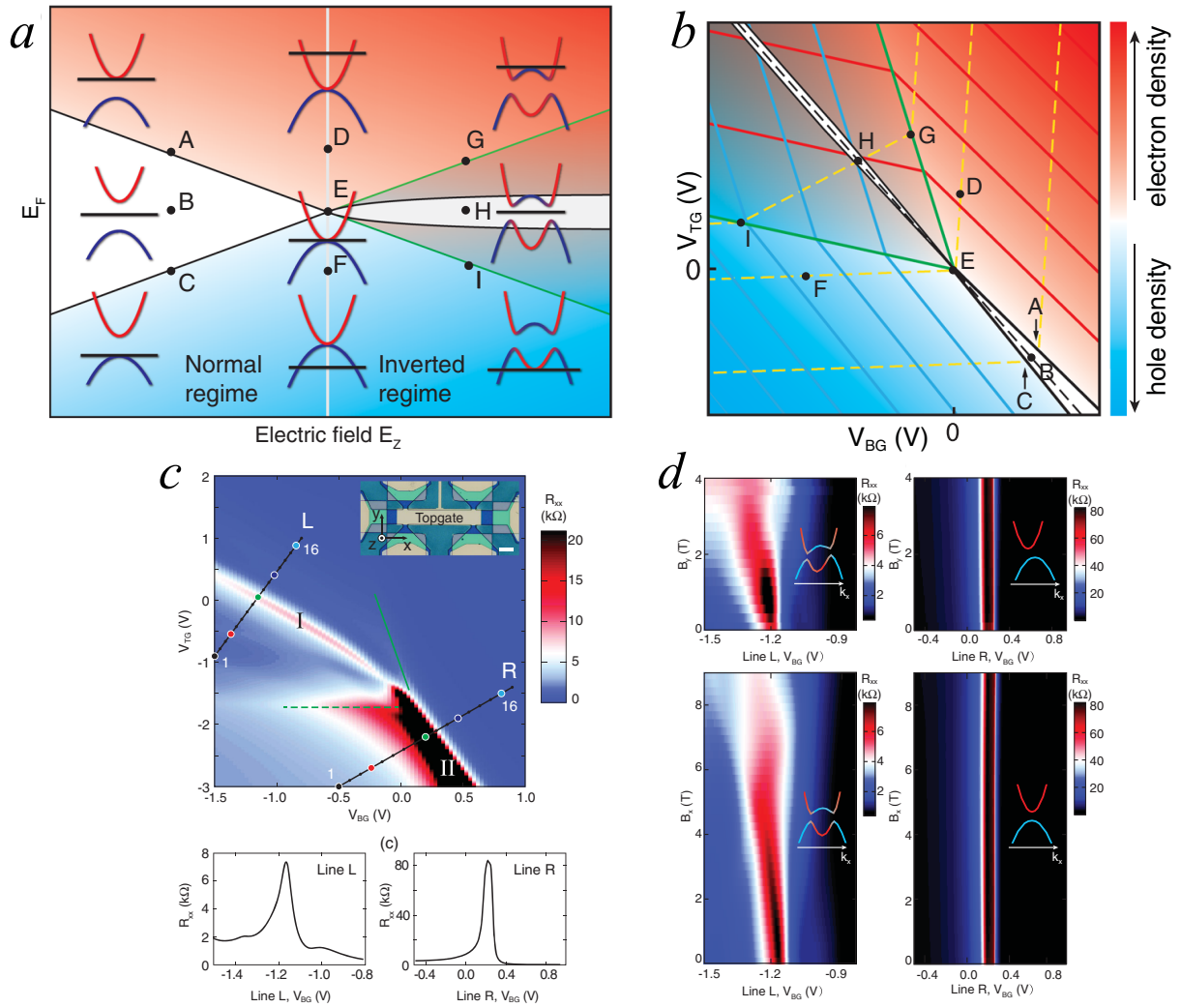


Figure 6.9: [Adapted from Ref. 216] a) Phase diagram of InAs/GaSb double-gated CQWs as a function of the applied electric field  $E_z$  and the Fermi level position  $E_F$ . Both  $E_z$  and  $E_F$  can be tuned by the dual gating. The vertical white line, along which the electron (red) and hole (blue) bands touch, separates the normal and inverted band structure regimes. b) Phase diagram as a function of back gate voltage  $V_{BG}$  and top gate voltage  $V_{TG}$ . The lettered points in (a) are indicated in (b). The red and blue lines mark lines of constant electron and hole densities, respectively. The yellow dashed lines indicate constant band overlap for the inverted regime or constant band separation for the normal regime. The black dashed line denotes charge

neutrality,  $n = p$ . The white regions above and below point  $E$  correspond to the hybridization gap and the normal gap, respectively. c) Top: Four-terminal longitudinal resistance as a function of top and bottom gate at  $T = 300$  mK, revealing the phase diagram of the InAs/GaSb CQW. Lines  $L$  and  $R$  indicate the positions in gate space where longitudinal resistance traces are taken (Bottom). Line  $L$  passes through region I, the QSHI phase, and line  $R$  passes through the normal insulator phase, region II. Inset: Optical image of the Hall bar. Bottom: Resistance along lines  $L$  and  $R$ . d) In-plane magnetic field dependence of the longitudinal resistance along the same lines  $L$  and  $R$ . The resistance of  $L$  decreases with increasing in-plane field, but for  $R$ , resistance is constant with field.

#### 6.4 Initial Noise Measurements of Silicon-Doped InAs/GaSb Devices

Following the initial measurements of the QSH state in InAs/GaSb QWs, Du *et al.* sought to further suppress the observed residual bulk conductivity through the incorporation of a silicon dopant layer (Figure 6.10a).<sup>144</sup> Silicon doping yields a truly insulating bulk by localizing the bulk states, even at finite temperature, leaving only the edge states to carry conduction.<sup>217</sup> A relatively small density ( $10^3$  atoms/ $\mu\text{m}^2$ ) of Si at the interface of InAs and GaSb serve as donors to InAs and acceptors to GaSb, creating a localization gap of  $\Delta_{loc} \sim 26$  K (2.24 meV) in the bulk energy spectrum. The topologically protected edge states, though, are relatively unaffected by the disorder. In  $L=2$   $\mu\text{m}$  samples, this is evidenced by wide plateaus in the longitudinal conductance at  $2 \frac{e^2}{h}$  for a Hall bar and  $4 \frac{e^2}{h}$  for a  $\pi$  bar as a function of front gate voltage, which tunes the Fermi energy into the localization gap (Figure 6.10b). As the length of the Hall bar is increased to macroscopic scales, the longitudinal resistance in the localization gap increases linearly with device length.

The energy scale of the localization gap was determined from conductance measurements of a Corbino disk. In this geometry, edge transport is shunted by concentric electrodes, so that measurements of conductance involve bulk transport exclusively. The energy gap is found by plotting  $\ln\left(\frac{1}{G_{xx}}\right)v_F \frac{1}{T}$ , called an Arrhenius plot.  $G_{xx}$  is fit to  $G_{xx} \propto e^{-\frac{\Delta}{2k_B T}}$ , where  $\Delta$  is the energy required to create an electron-hole pair over the gap. At higher temperatures, they found a gap value  $\Delta_{min} \sim 66$  K, consistent with a hybridization-induced mini-gap. Below  $\sim 10$  K, the conductance continues to drop exponentially, with a different slope, indicating the opening of a localization gap  $\Delta_{loc} \sim 26$  K. It is in this regime that the conductance plateaus are observed (*Figure 6.9c*). The localization gap can be increased to 40 K with a perpendicular magnetic field of 6 T.

The Fermi velocity of the InAs/GaSb edge states  $v_F \sim 1.5 \times 10^4$  m/s is at least an order of magnitude smaller than that of GaAs 2D electron gas or HgTe/CdTe ( $v_F \sim 5.5 \times 10^5$  m/s),<sup>218</sup> due to the gap opening at a finite wavevector  $k_{cross}$ . The edge scattering time, inversely proportional to the Fermi velocity, is thus extremely long, regardless of disorder in the bulk. The robust nature of the topologically-protected edge modes is further demonstrated by the fact that the quantized plateaus of shorter samples and the linear resistance of the larger samples was found to be independent of temperature between 20 mK and 4 K (*Figure 6.10 inset*).

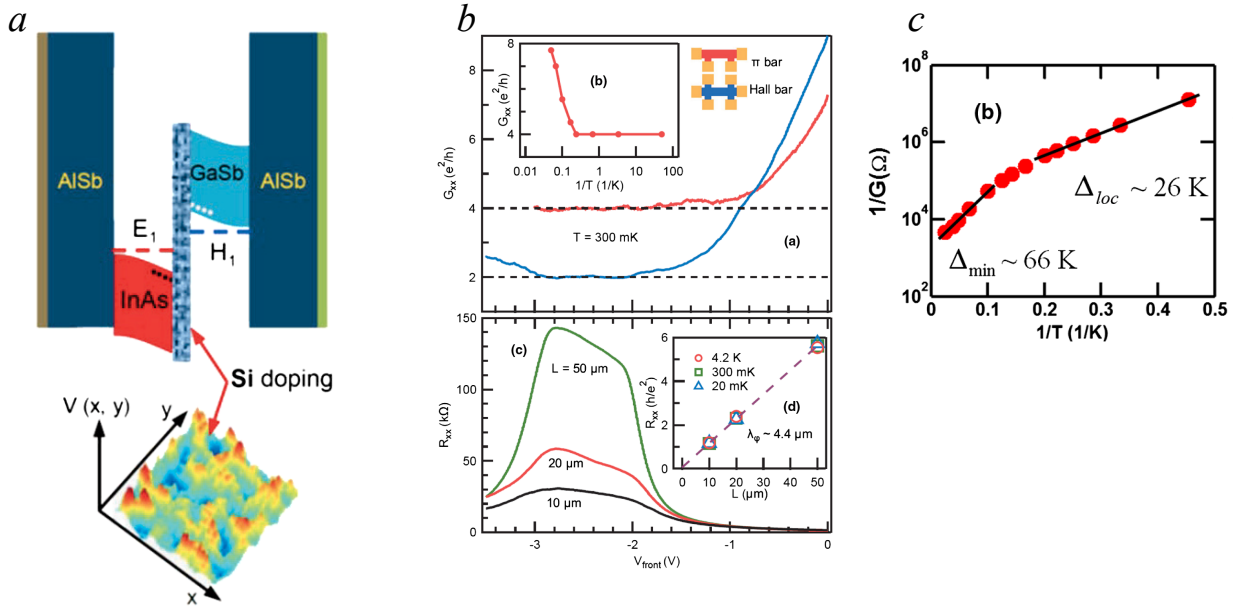


Figure 6.10: [Adapted from Ref. 144] a) Schematic of the InAs/GaSb CQW with Si doping and the potential fluctuations induced by the dopants at the interface. b) Top: Wide conductance plateaus quantized to  $\frac{4e^2}{h}$  for the  $\pi$  bar (red) and  $\frac{2e^2}{h}$  for the Hall bar (blue). Top Inset: Plateau persists to 4 K, but at higher temperatures conductance increases due to delocalized 2D bulk carriers. Bottom: Longitudinal resistance as a function of front gate voltage for longer devices. Bottom Inset: Resistance scales linearly with edge length, indicating a coherence length  $\sim 4.4 \mu\text{m}$ . The coherence length is independent of temperature between 20 mK and 4 K. c) Arrhenius plot for a Corbino geometry, showing bulk conductance vanishes exponentially with decreasing temperature.

Under perpendicular magnetic field, the authors observed an increase in the four-terminal conductance of the Hall bar with increasing field. In this case, the time reversal symmetry is broken as the field pushes the edge states of different chiralities in opposite directions. Consistent with a trend toward chiral transport, the authors also observed increases in Hall

resistance with increasing field. In the two-terminal device, conductance decreased with increasing field, due to the fact that in two-terminal high-field magnetotransport, the signal is dominated by the Hall resistance, which increases with field.<sup>219</sup>

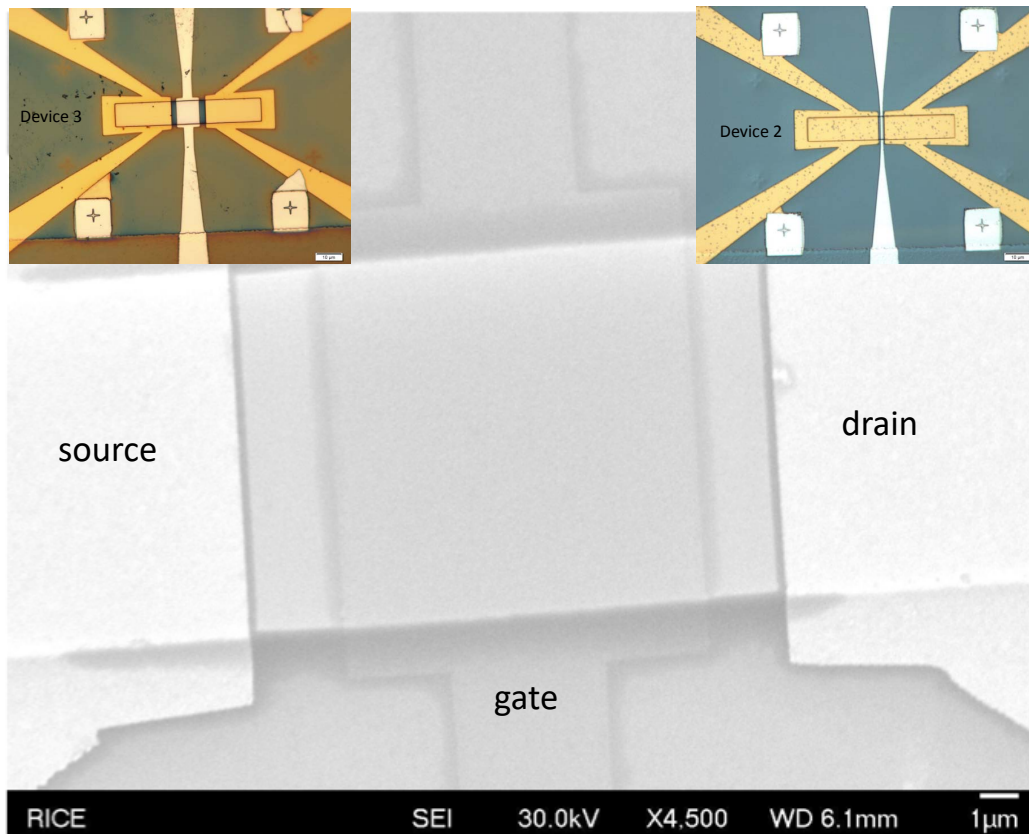
The observations of Knez *et al.*, Du *et al.*, and others with arguments for the presence of the two-dimensional topological insulator state in InAs/GaSb CQWs spurred some controversy over the nature of the edge states. For example, Nichele *et al.* demonstrated the existence of edge states in the trivial regime of InAs/GaSb.<sup>220</sup> They found that resistance scaled linearly with device length, down to values below  $\frac{h}{e^2}$  for short edges, in contrast with the expected length-independent resistance when length is shorter than the coherence length for truly ballistic helical edge modes. Trivial edge modes can form and be populated by carriers due to band bending at the device edges, and in this case, edge conduction would be multi-channel.

Using the methods described in *Section 2.6* and *Section 5.3*, we know it is possible to use shot noise to distinguish between single- and multi-channel transport. Therefore, measurements of the shot noise can differentiate between the trivial edge states due to band bending (multi-channel) and nontrivial, time reversal symmetry protected edge states (single-channel). Furthermore, as discussed in *Section 3.7*, shot noise has been proposed as a useful probe for determining the root cause of backscattering in two-dimensional topological insulators. Aiming to verify the nature of the edge modes, we measured noise in two-terminal InAs/GaSb QWs as a function of temperature, gate voltage, and perpendicular magnetic field.

#### **6.4.1 Sample Structure and Measurement Setup**

*Figure 6.11* is an SEM image and two optical images of example devices, fabricated by Tingxin Li at Peking University. The wafers were grown by molecular beam epitaxy (MBE) at

Teledyne by Gerard Sullivan and included the silicon doping layer discussed in the previous section, with a concentration of about  $10^{11} \text{ cm}^{-2}$ . As seen in *Figure 6.11*, source and drain cover the entire width of the quantum well mesa, so that a two-terminal measurement should include contributions from both edges. A top gate covers the mesa to tune the Fermi level of the CQW.



*Figure 6.11: SEM image of one of the long devices. Insets: Optical images of one of the long devices (left) and one of the short devices (right).*

Samples of two different lengths were measured; for the first, the distance between the source and drain contacts was  $14 \mu\text{m}$ , and the mesa and front gate length was  $10 \mu\text{m}$  (*Figure 6.11 left inset*), and for the second, the contacts were separated by  $2.5 \mu\text{m}$ , and the mesa was approximately  $1 \mu\text{m}$  (*Figure 6.11 right inset*). Based on the previous studies, the short sample

should be within the coherence length of the edge modes, while the longer sample should be subject to backscattering processes due to disorder.

Conductance measurements were performed as described in *Section 4.3*, and noise measurements were executed using the lock-in based broadband rf technique outlined in *Section 4.4*.

## **6.4.2 Results**

### *Temperature Dependence*

In both devices, as temperature is decreased, conductance decreases as expected for the gapping out of bulk states. At low temperatures, both devices also show a zero-bias conductance suppression, with a wider and stronger suppression observed in the shorter device (*Figure 6.12a,b*). In general, the longer device exhibits more Ohmic behavior. At temperatures comparable to the hybridization gap (40-60 K) and above the localization gap of residual bulk conduction, noise in both devices as a function of applied voltage bias is monotonic. In the short device at this higher temperature range, the lock-in detected differential current noise is roughly flat with applied bias voltage up to approximately 30 mV (*Figure 6.12c*) before growing linearly with bias voltage. The overall magnitude of the noise in the longer device at the higher temperatures is lower than that in the short device, consistent with the expectation that there would be more dephasing scattering events for a device longer than the coherence length. As temperature is decreased, an interesting trend occurs for both devices: the nonequilibrium noise power decreases with increasing bias up to some finite bias ( $\sim 30$  mV for the short device,  $\sim 100$  mV for the long device) at which point it then turns upward and begins to increase with increasing bias (*Figure 6.12c,d*). In other words, at low biases, the overall noise power is lower

at finite bias than it is at zero bias. This is interpreted as negative differential current noise power by the lock-in, which returns  $S_I(V) - S_I(V=0)$  (Section 4.4), and this difference grows more negative with increasing bias until the voltage reaches some critical value above which nonequilibrium noise once again increases linearly with bias voltage.

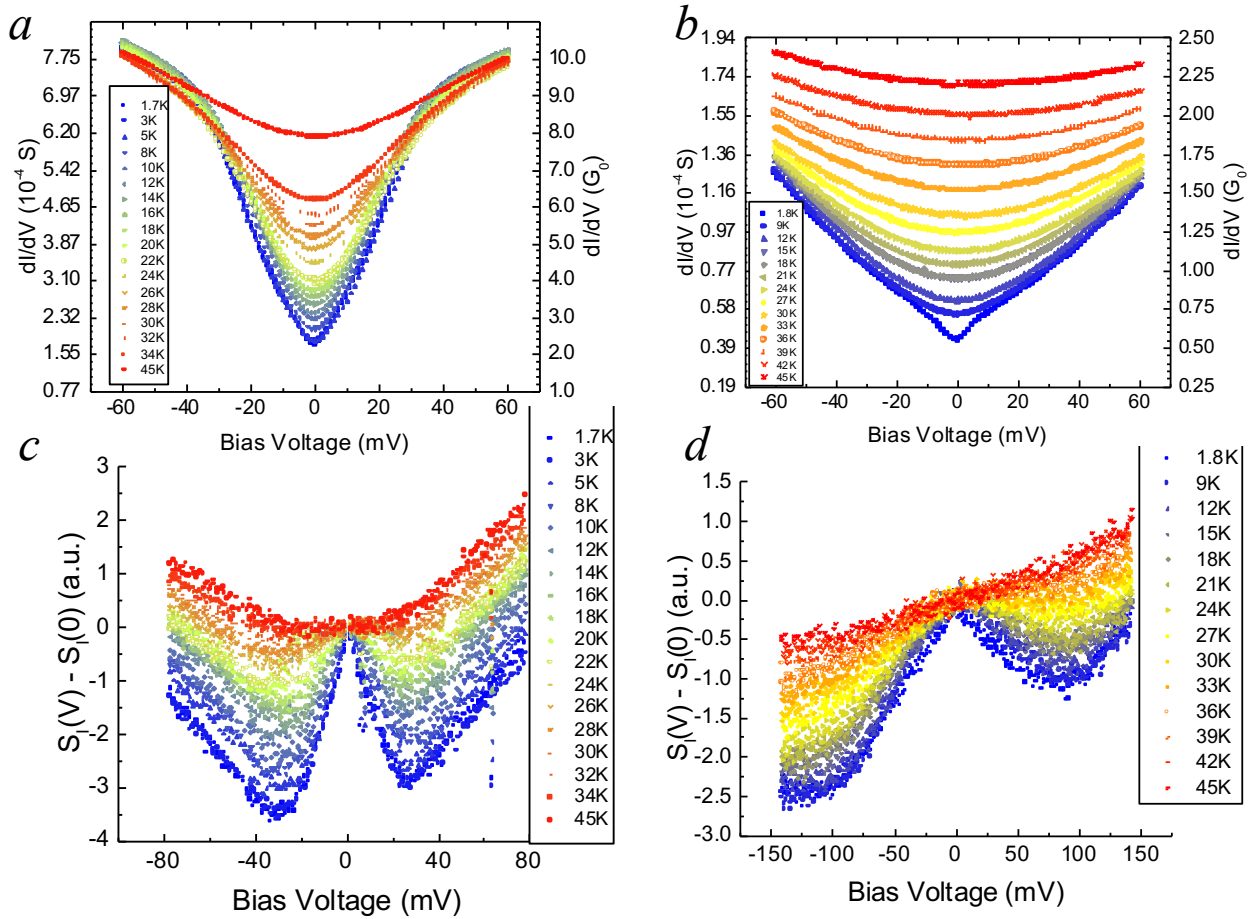


Figure 6.12: *a,b*) Differential conductance as a function of temperature in the short device (*a*) and the long device (*b*) As expected, conductance decreases with decreasing temperature as the 2D bulk states are gapped out. At low temperatures, both devices exhibit a zero-bias conductance suppression. *c,d*) Differential noise power as a function of temperature for the short (*c*) and long (*d*) devices. For both devices, as temperature is decreased, differential noise



is increasingly negative with applied bias until some critical bias voltage, above which the noise begins to increase linearly with bias.

### *Gate Bias Dependence*

At  $T = 1.8$  K, when a positive voltage is applied to the top gate, InAs electron-like states are filled, and consequently, the conductance of both devices slightly increases (*Figure 6.13d*). As a negative gate voltage is applied, the system reaches the topological insulator phase, with the longitudinal resistance plateau manifesting in the range of approximately -0.75 to -1.25 V for both devices. In the noise of the short device, the nonmonotonic trend is weakened as the gate is swept either positively or negatively; by  $V_g = 1$  V, it has disappeared completely, and at  $V_g = -1.2$  V, it is still somewhat present, but is greatly reduced (*Figure 6.13b,c*). In the long device, the nonmonotonic noise shape is not significantly diminished by a positive gate voltage, but it is nearly suppressed by  $V_g = -0.75$  V (*Figure 6.13e*).

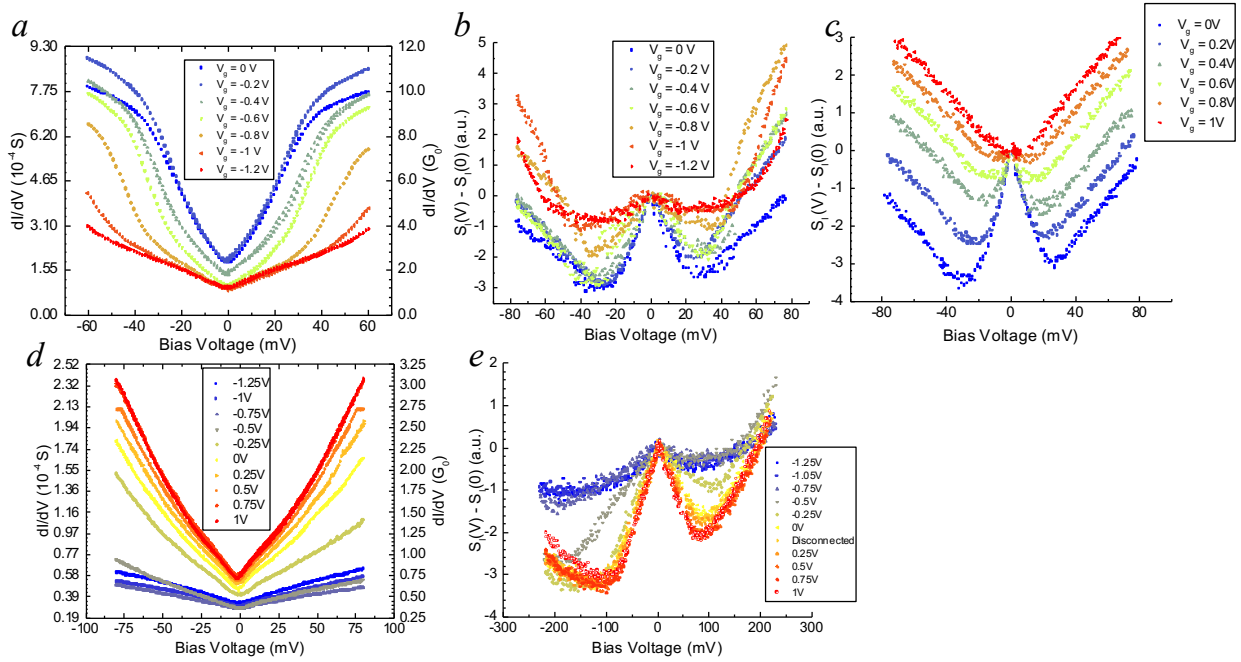


Figure 6.13: *a,d*) Differential conductance as a function of gate voltage for the short (*a*) and long (*d*) devices. Resistance peak for both devices happens around  $V_g = -0.75 - -1.25$  V. *b,c,e*) Differential current noise with varying gate voltage for the short (*b,c*) and long (*e*) devices. Nonmonotonic noise behavior is suppressed as gate voltage is increased positively or negatively.

### Magnetic Field Dependence

With increasing perpendicular magnetic field, conductance in both devices decreased (Figure 6.14*a,b*), consistent with the two-terminal results of Du *et al.*<sup>144</sup> In both devices, the nonmonotonic noise trend is weakened by the perpendicular magnetic field. In the long device, it is completely suppressed by  $B = 4$  T (Figure 6.14*d*), and in the short device, the noise behavior is monotonic by 5 T (Figure 6.14*c*). Additionally, in the high-bias regime, when noise exhibits the expected positive linear dependence with bias, the slope of  $S_I(V)$ , proportional to the Fano factor, also decreases with increased magnetic field.

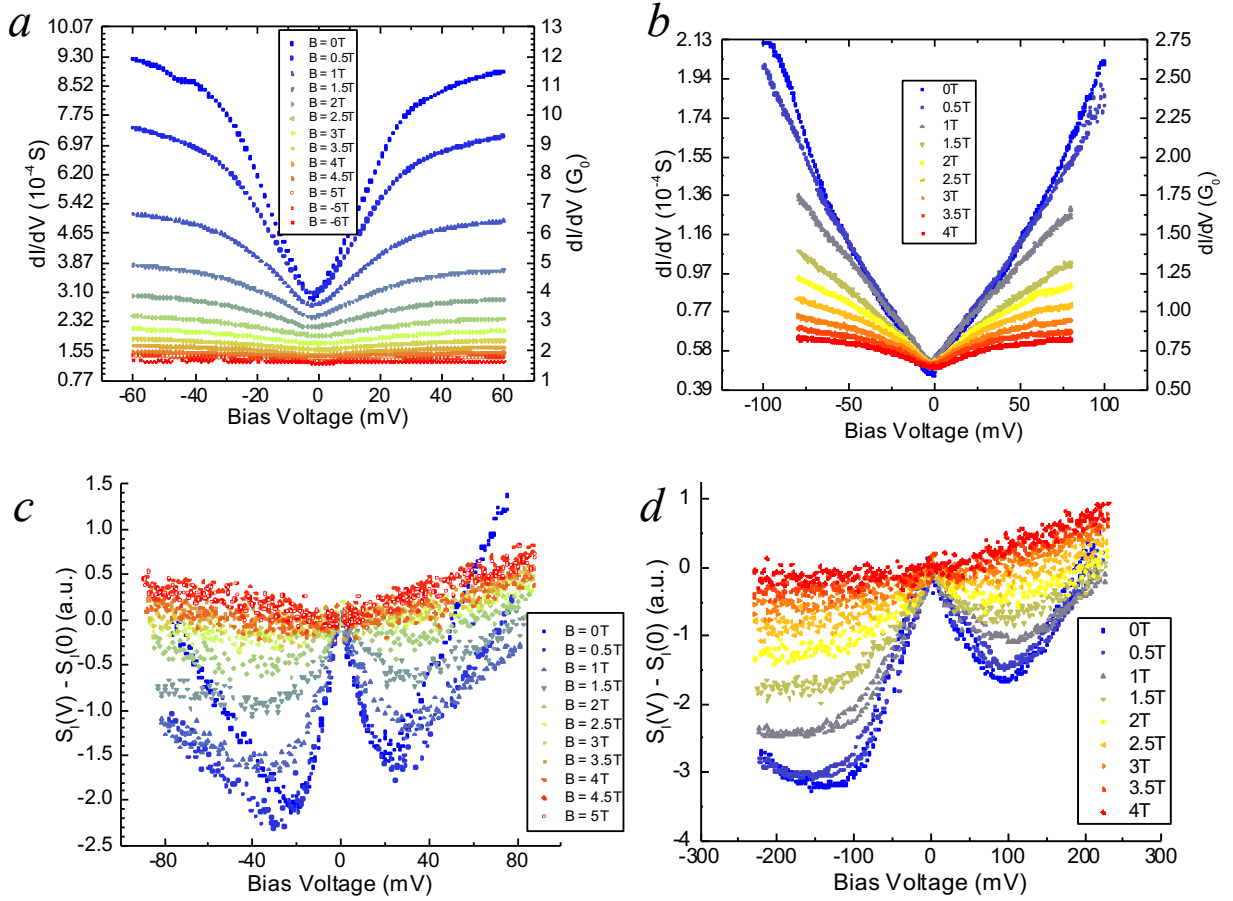


Figure 6.14: *a,b*) Differential conductance with varying perpendicular magnetic field in the short (*a*) and long (*b*) devices. Increasing field decreases conductance in both devices. *c,d*) Differential current noise in the short (*c*) and long (*d*) devices. Negative noise trend decreases with increasing field.

### 6.4.3 Discussion

The decrease in noise power with increasing bias is unique to these doped InAs/GaSb structures. As this behavior was observed in devices both longer than and shorter than the coherence length predicted by Du *et al.* ( $l_\phi \sim 2 \mu\text{m}$ ),<sup>144</sup> it is likely a bulk effect and not due to scattering processes at the edges. The most probable explanation is that the silicon dopant layer

produces generation-recombination (GR) noise in the quantum wells that competes with the shot noise. The electric field that arises between the source and drain when a voltage is applied across the junction leads to ionization of the Si dopants, suppressing this noise, resulting in a lower overall noise power at finite bias than at zero bias, until the bias is large enough to allow shot noise to dominate.

Generation-recombination noise arises due to fluctuations in the number of free carriers associated with random transitions of charge carriers between states in different energy bands, such as transitions between the conduction band and localized states within the energy gap of a semiconductor.<sup>221–224</sup> Figure 6.15 is an example of a current spectral density demonstrating  $1/f$ , GR, and Johnson-Nyquist noise as frequency is increased, based on experiments in  $n$ -type silicon.

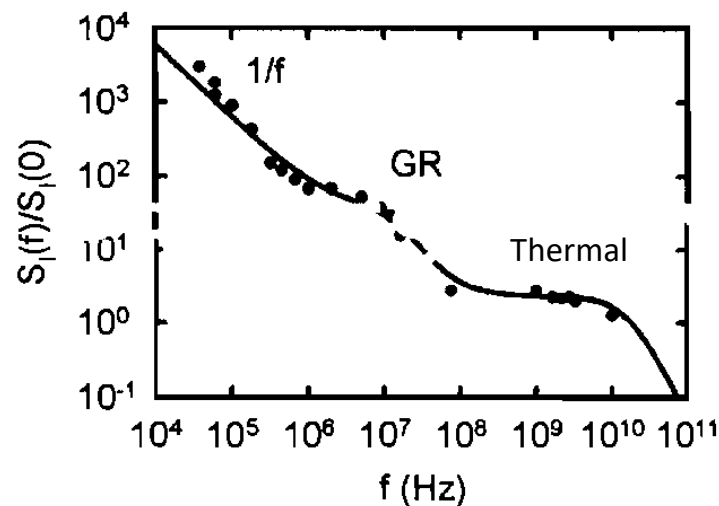


Figure 6.15: [Adapted from Ref. 221] Current spectral density of an example  $n$ -doped Si device, with indications of ranges of  $1/f$ , generation-recombination, and thermal noise.

The standard theory for GR noise is derived for a two-level system, with a probability per unit time for a free carrier to be generated via a transition from the impurity level to the

conduction band,  $g(N)$ , and a probability per unit time that an existing free carrier will undergo a recombination transition from the conduction band to the impurity level,  $r(N)$ .<sup>221</sup> It is assumed the rates  $g$  and  $r$  depend only on the instantaneous number of carriers in the conduction band,  $N(t)$ , and in general,  $g$  tends to decrease with  $N$  and  $r$  tends to increase with  $N$ , with the specifics dependent on the particular system of interest. Under steady state or stationary conditions,

$$g_0 = g(N_0) = r(N_0) = r_0. \quad (6.10)$$

The probability of finding  $N$  electrons in the conduction band at time  $t$  can be approximated with a normal law for  $N(t)$  near the steady state value  $N_0$

$$P(N) = P(N_0)e^{-\frac{(N-N_0)^2}{2\delta N^2}}. \quad (6.11)$$

If the generation and recombination rates are rewritten as  $\frac{1}{\tau_g} = -\frac{dg}{dN_{N=N_0}} = -g'_0$  and  $\frac{1}{\tau_r} =$

$\frac{dr}{dN_{N=N_0}} = r'_0$ , then we have

$$\tau_N = \frac{1}{r'_0 - g'_0}, \quad (6.12)$$

$$\overline{\delta N^2} = \frac{g_0}{r'_0 - g'_0} = g_0 \tau_N, \quad (6.13)$$

and the spectral density of the fluctuations in the number of charge carriers is

$$S_N(\omega) = 4g_0\tau_N^2 \frac{1}{1+(\omega\tau_N)^2}, \quad (6.14)$$

where  $\tau_N$  is the majority charge carrier lifetime. The current noise spectral density due to GR noise can be written as

$$S_I(\omega) = I^2 \frac{S_N(\omega)}{N_0^2} = I^2 \frac{\overline{\delta N^2}}{N_0^2} \frac{4\tau_N}{1+(\omega\tau_N)^2}. \quad (6.15)$$

The typical cutoff frequency range for GR noise is below the GHz region.<sup>221</sup> Therefore, it is reasonable to expect some GR noise in our RF range of approximately 250-600 MHz. Furthermore, for doped, or extrinsic, semiconductors, GR noise is expected to have a strong

temperature dependence, with  $\tau_N$  (and thus  $S_{I,GR}$ ) rapidly decreasing with increasing temperature.<sup>224,225</sup>

Early studies of GR noise in extrinsic semiconductors also demonstrated that GR noise could give way to shot noise, displaying a linear relationship with applied current.<sup>222,226,227</sup> This is possible when the transit time from source to drain,  $\tau_T = \frac{L}{v_d}$ , where  $L$  is the length between source and drain and  $v_d$  is the drift velocity, becomes the shortest time scale compared to the lifetime of the GR processes and the dielectric relaxation time,  $\tau_d = \frac{\epsilon_r \epsilon_0}{en\mu}$ , where  $\epsilon_r$  is the relative static dielectric constant,  $\epsilon_0$  is the permittivity of free space,  $n$  is the free carrier concentration, and  $\mu$  is the carrier mobility. It is conceivable, then, that at low applied biases, GR noise, which is present even at zero bias, decreases with increasing bias until  $v_d$  becomes large enough to allow  $\tau_T$  to be the shortest timescale. As bias increases past this critical point, shot noise becomes detectable with a positive linear bias dependence.

In addition to the temperature dependence agreeing with expectations for GR noise, the gate and magnetic field dependence also seem well-explained by a GR picture. As the gate voltage adjusts the Fermi level of the system, the mobile charge carriers become less vulnerable to trapping. When an applied perpendicular magnetic field is increased, and the two-terminal edge conduction is suppressed due to localization effects, both GR and shot noise also decrease. Furthermore, the larger bias required to suppress the nonmonotonic trend in the longer device is consistent with the scaling of the transition from GR to shot noise with the transit time between source and drain; a shorter  $L$  requires a smaller  $v_d$  to yield a small transit time.

While silicon-doped InAs/GaSb quantum wells are no longer in use for QSH studies, this interesting result demonstrates how other forms of noise can manifest in 2DTIs.

## 6.5 Noise Processes in InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb Corbino Structures

In a newer iteration of InAs/GaSb quantum spin Hall insulators, Du *et al.* explored the enhancement of the bulk gap in strained-layer InAs/GaInSb quantum wells.<sup>190</sup> The strain effect caused by inserting In atoms into the GaSb lattice causes changes in the band structure of the CQW (Figure 6.16),<sup>228</sup> allowing the QWs to be made narrower, which in turn yields stronger overlap between the electron and hole wavefunctions. As a result, insulating hybridization gaps at low temperatures are possible even when the charge density at the CNP,  $n_{cross}$ , exceeds  $3 \times 10^{11} \text{ cm}^{-2}$ .

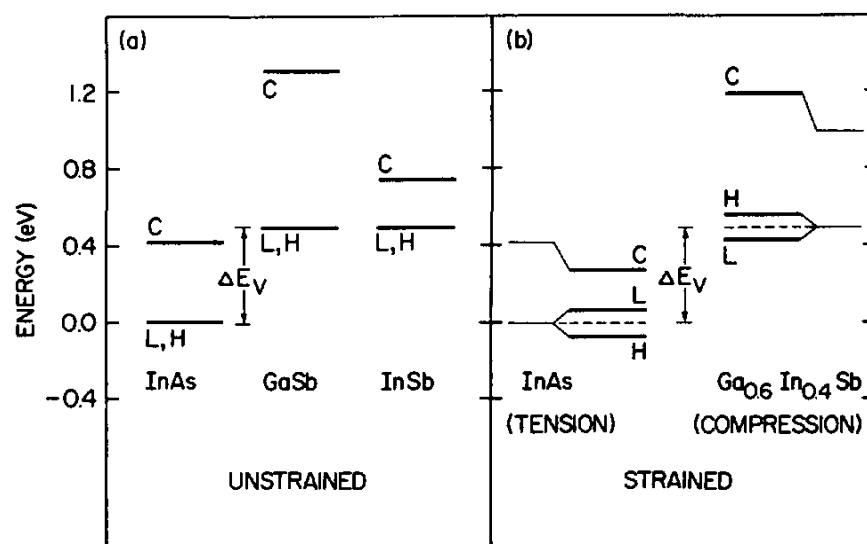


Figure 6.16: [Adapted from Ref. 228] Assumed band positions for unstrained (left) InAs, GaSb, and InSb and in comparison, the effect of strain on InAs and Ga<sub>0.6</sub>In<sub>0.4</sub>Sb (right).

The strained-layer InAs/Ga<sub>1-x</sub>In<sub>x</sub>Sb superlattice (SL) was first proposed by Smith and Maihiot in 1987 for infrared detectors.<sup>228</sup> Alloying GaSb (lattice constant  $\sim 6.1 \text{ \AA}$ ) with InSb (lattice constant  $\sim 6.4 \text{ \AA}$ ) causes strain in the crystal growth plane, resulting in a downward shift of the conduction band (CB) energy of InAs and a split in the energy level of the valence band

(VB) in GaInSb between the heavy hole (HH) and light hole (LH) levels. The HH level rises higher than the original top VB in GaSb. Therefore, the QSHI state can be reached in narrower QWs of strained-layer InAs/GaInSb compared to the unstrained InAs/GaSb case. The hybridization gap should increase in these narrower wells due to the enhanced overlap of the electron and hole wavefunctions. Furthermore, due to the energy splitting of the HH and LH levels in GaInSb, the Fermi surfaces of the electrons and holes should be better matched to each other, which could help reduce residual nonhybridized carriers.

*Figure 6.17a* is the band structure for the InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb QWs used in this study, calculated using the eight-band Kane model. The results indicate a hybridization gap of about 20 meV, a roughly fivefold enhancement over previous unstrained InAs/GaSb QWs with  $\Delta \sim 4$  meV. *Figure 6.17b* is a TEM image of the InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb wafer, indicating the crystalline structure remains coherent across the heterostructure, despite the approximately 1.5% in-plane strain.



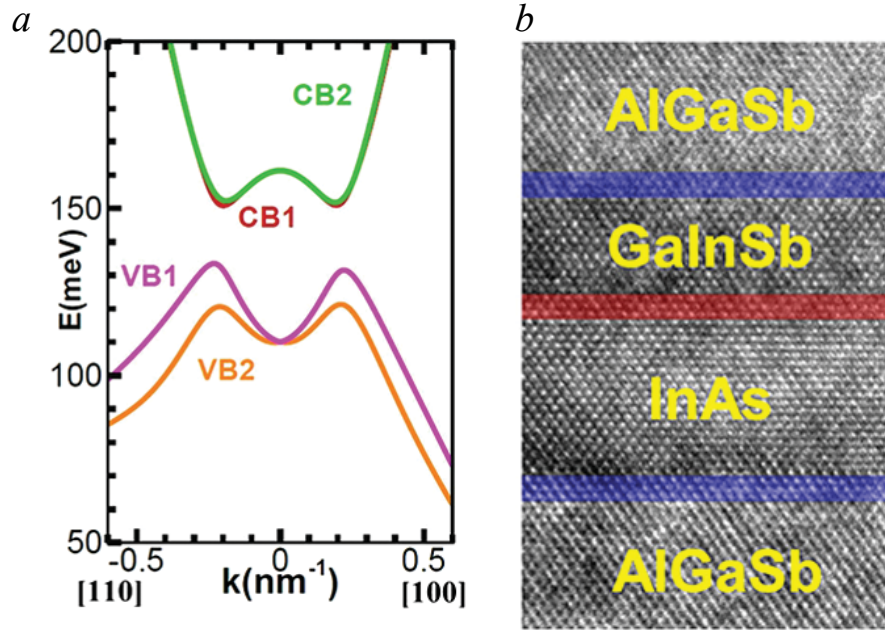


Figure 6.17: [Adapted from Ref. 190] a) Calculated bulk band structure of InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb CQW. The strain effect splits the heavy and light hole bands, and the strain effect creates a hybridization gap of about 20 meV. b) TEM image of the InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb wafer, showing the crystalline structure is intact throughout the heterostructure.

Du *et al.* directly measured the bulk conductance of the strained-layer QWs using Corbino devices, in which any edge conduction is shunted. Figure 6.18 shows the conductance for different concentrations of indium as a function of temperature. Based on the temperature dependence, the hybridization gaps were estimated by fitting the Arrhenius plot to  $e^{-\frac{\Delta}{2k_B T}}$ , with  $\Delta \sim 250$  K for InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb. Furthermore, the coherence lengths of the strained-layer devices, the length scale at which dissipationless edge transport breaks down and the counterpropagating channels equilibrate, were  $\sim 5$ - $11$   $\mu\text{m}$ , a significant improvement on past devices.<sup>121,144</sup>

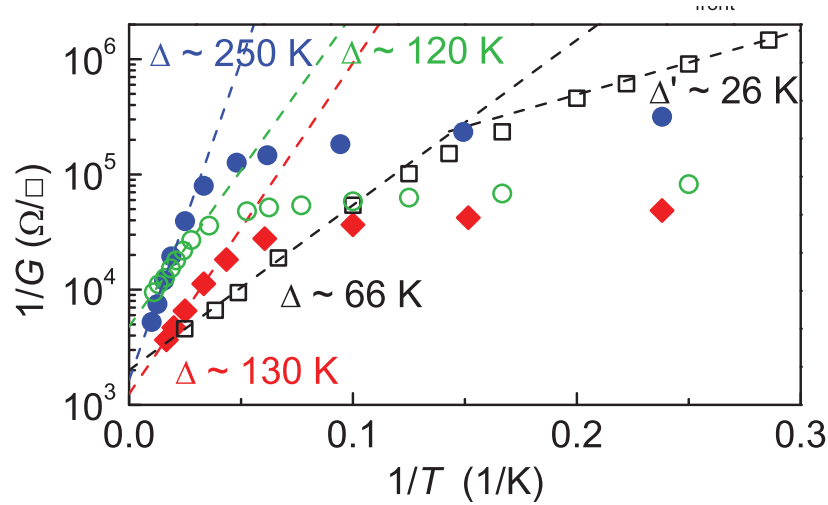


Figure 6.18: [Adapted from Ref. 190] Arrhenius plots for InAs/GaSb (open squares), InAs/Ga<sub>0.8</sub>In<sub>0.2</sub>Sb (open circles), InAs/Ga<sub>0.75</sub>In<sub>0.25</sub>Sb (filled diamonds), and InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb (filled circles). Energy gaps are calculated based on fitting  $G_{xx} \propto e^{-\frac{\Delta}{2k_B T}}$ , shown as the dashed lines in the plot.

As an initial step in characterizing the shot noise in these 2DTIs, we sought to measure the noise in Corbino structures of InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb to identify any noise contributions from the bulk or contacts.

### 6.5.1 Device Structure and Measurement Setup

Figure 6.19a<sup>229</sup> is a colorized SEM image of an example Corbino device, with inner/outer diameters of 800/1200  $\mu\text{m}$ . The wafers used were grown by G. Sullivan by molecular beam epitaxy. The full wafer composition is given in Appendix 1. The structures were fabricated by the photolithography and etching steps outlined in Section 4.1 and Appendix 2. Figure 6.19b is a schematic of the final stage wafer layers, and Figure 6.19c depicts the band structure of the CQWs.

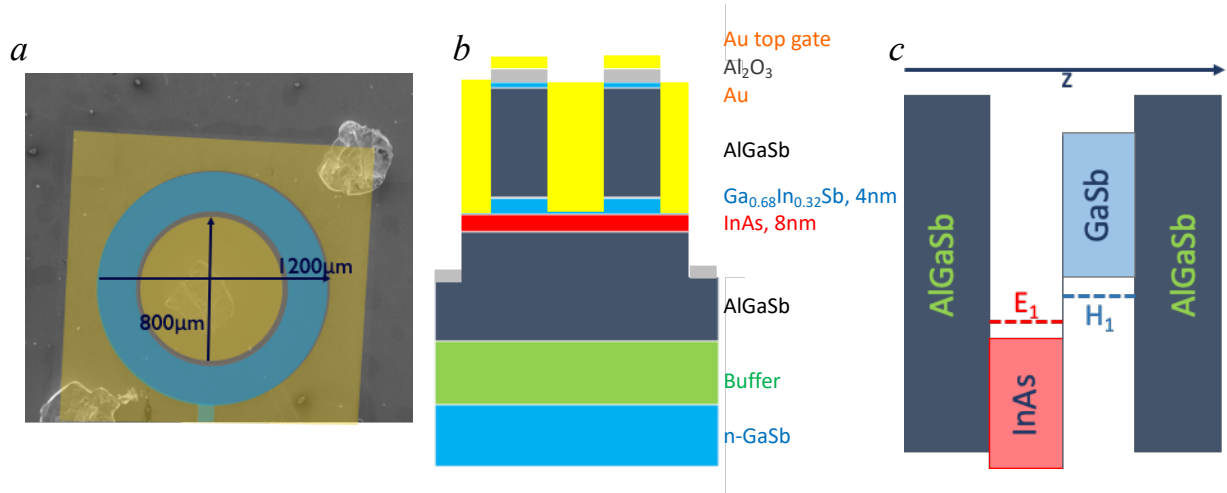


Figure 6.19: [Adapted from Ref. 229] a) Colorized SEM image of an example Corbino device. Yellow denotes the inner and outer electrodes, and blue marks the area of the top gate. b) Diagram of device structure. c) Band structure of the CQW.

From Chapter 2, for a two-terminal device at finite temperature, we expect the total current noise spectral density ( $A^2/Hz$ ) to be given as

$$S_I = F \cdot 2eI \coth\left(\frac{eV}{2k_B T}\right) + (1 - F) \cdot 4k_B T G, \quad (6.16)$$

where  $I$  and  $V$  are the current and voltage bias across the junction,  $G$  is the two-terminal conductance, and  $F$  is the Fano factor, a measure of how the high bias ( $eV \gg 2k_B T$ ) shot noise compares to the ideal Poissonian value,  $S_I = 2eI$ . Equation 6.16 is a phenomenological expression designed to give the correct Johnson-Nyquist noise ( $4k_B T G$ ) at  $I, V = 0$  and a bias-independent Fano factor in the limit  $eV \gg 2k_B T$ . Shot noise and thermal noise are expected to be white over the frequency ranges of interest.

As discussed in Section 2.8, as the length of a diffusive conductor increases beyond the inelastic mean free path for scattering energy out of the charge carriers, electron-phonon interactions cause shot noise to decay with increasing length until it is fully suppressed when the

electrons have thermalized to the phonon temperature.<sup>54,73,230</sup> Due to the relatively long length between contacts, we expect to be in the limit of fully suppressed shot noise from the 2D bulk.

In considering shot noise contributions arising at contact barriers, we define the contacts as a series combination of gold to InAs then InAs to the hybridized InAs/GaInSb interface. From *Section 6.1*, we do not expect any significant barrier between Au and InAs.<sup>191,194</sup> At the InAs to InAs/GaInSb interface, however, there can be some band offset when the Fermi level is in the hybridization gap. This band offset can act as a contact barrier that produces shot noise, analogous to noise-producing Schottky diodes.<sup>231</sup>

Noise measurements were first performed using the broadband RF method described in *Section 4.4*. Then, to clarify the absolute magnitude of the noise, we also employed the low frequency noise spectroscopy method outlined in *Section 4.5*.

### 6.5.2 Results and Analysis

In the initial RF noise measurements, at high temperatures, thermal excitation of carriers across any hybridization gap is sufficient to yield a conductive bulk. At 100K,  $\frac{dI}{dV}$  measurements of the first sample showed a two-terminal resistance of about 200  $\Omega$ , with approximately Ohmic response. As expected, the two-terminal zero-bias resistance increased with decreasing temperature (*Figure 6.20a*). By 5 K, the resistance had increased to 10 k $\Omega$ , which corresponds to a square-resistance,  $R_{sq} = \frac{2\pi}{\ln\left(\frac{r_{out}}{r_{in}}\right)}$ , of roughly 155 k $\Omega$ , where  $r_{in,out}$  correspond to the radii of the of the inner/outer electrodes. At 5 K, when the device is positively top-gated to populate the InAs QW, the two-terminal resistance dropped to only 80  $\Omega$  at  $V_g = 0.9$  V. In the higher temperature range of 30 – 100 K, the lock-in detected change in noise power is essentially zero

and flat with applied bias up to 50  $\mu\text{A}$ . Around 20 K, however, shot noise as a function of current becomes detectable, and increases with decreasing temperature down to 5 K (*Figure 6.20b*). The critical temperature for the onset of shot noise approximately corresponds to the temperature at which conduction in the 2D bulk gaps out.<sup>190</sup> Additionally, when the positive gate voltage is applied to the device, closing the bulk gap, the inferred current noise is again roughly flat with increasing bias (*Figure 6.20b*).

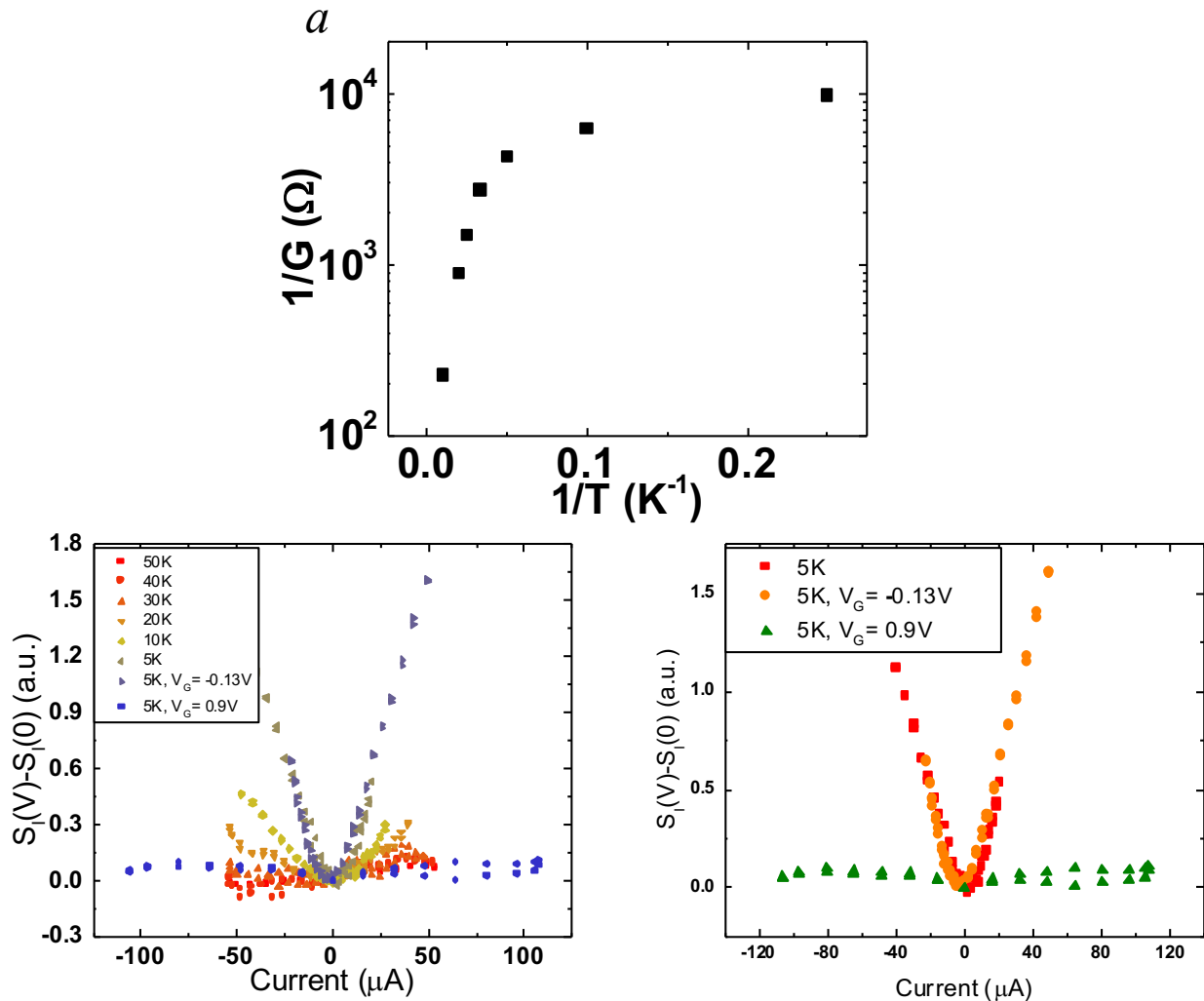


Figure 6.20: [Adapted from Ref. 229] a) Zero-bias two-terminal resistance as a function of temperature. By 5 K, two-terminal resistance is approximately  $10\text{ k}\Omega$ , corresponding to a square resistance of about  $155\text{ k}\Omega$ . b,c) Integrated differential current noise in RF bandwidth ( $\sim 250 - 600\text{ MHz}$ ) as a function of temperature (b) and gate voltage (c). Shot noise is essentially fully suppressed at high temperatures and positive gate voltages, when the bulk is conductive. The shot noise contribution begins to turn on with increasing current around 20 K, as bulk transport gaps out, and continues to grow as temperature is decreased.

Voltage noise was measured in additional devices at low frequencies. At zero bias, the voltage noise was consistent with the Johnson-Nyquist expectations for the resistance obtained from preceding  $\frac{dI}{dV}$  measurements. Again, the noise remained roughly constant with applied current at higher temperatures, but started to increase with increasing current below  $\sim 20$  K (Figure 6.21a), indicative of the onset of some shot noise contribution. The magnitude of the bias-dependent noise, however, was consistently much smaller and had a broader curvature around zero bias than expected if one naively applied Equation 6.16 for a given conductance and temperature, assuming a Fano factor of 1 and that all of the applied voltage contributes to the argument of the coth term (Figure 6.21).

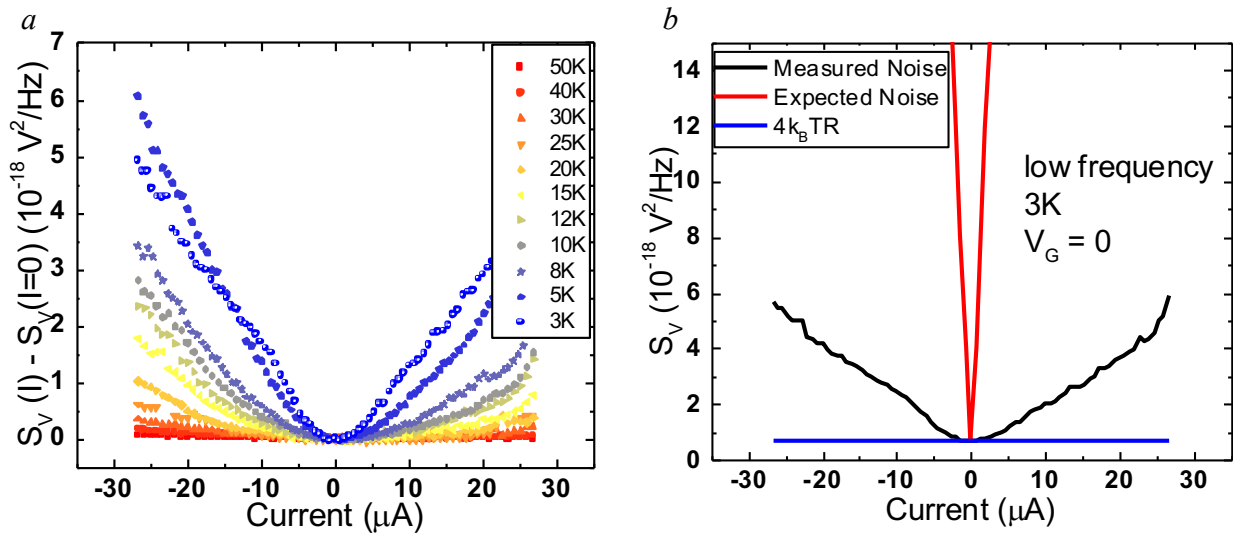


Figure 6.21: [Adapted from Ref. 229] a) Voltage noise minus the zero-bias voltage noise as a function of temperature. The nonequilibrium noise is relatively flat with current above 20 K, but then begins to increase with decreasing temperature. b) Comparison of the measured noise to the expected noise based on Equation 6.16. Assuming all of the applied bias is contributing to the shot noise, the magnitude of the noise at 3 K is much smaller than predicted, and the curvature about zero bias is much broader.

Particularly motivated by the breadth of the curvature of the noise about zero bias, we attempt to model the noise based on the equivalent circuit of *Figure 6.22*. The circuit is essentially three resistors: two contact resistances (taken to be identical for simplicity) and a bulk resistance representing the contribution from the 2D Corbino bulk. We assume the band offset at each InAs-InAs/GaInSb contact is the dominant source of contact resistance that can act as a barrier to produce shot noise. In the simplest case, we assume both the bulk resistance,  $R_b$ , and the contact resistances,  $R_c$ , are Ohmic, with total device resistance  $R = 2R_c + R_b$ . More generally, to account for a non-Ohmic device, the resistances should be differential resistances,  $\frac{dV}{dI}$ , found self-consistently under biasing conditions such that the total voltage across the three resistors in series is the applied dc bias across the device,  $V_{dev} = \frac{RV_{tot}}{R+300\text{ k}\Omega}$ , where  $R$  is the total two-terminal device resistance,  $V_{tot}$  is the total dc voltage applied by the DAQ, and 300 k $\Omega$  accounts for the two 150 k $\Omega$  current-limiting series resistances on either side of the device. In the temperature range covered in this study, the device was still relatively Ohmic, and thus unlikely to be in the limit of Poole-Frenkel hopping transport, a mechanism by which electrons can move through an insulator under an electric field by hopping between localized states due to random thermal fluctuations that allow the electrons to briefly gain sufficient energy to be in the conduction band before relaxing back into another localized state.<sup>232</sup> As temperature approaches zero, and the system reaches a limit in which all carriers are frozen out, it would be necessary to reevaluate the primary charge transport mechanism. In our simplified model, however, the bulk is considered diffusive and large relative to the inelastic electron-phonon scattering length scales, and therefore should only contribute to the Johnson-Nyquist noise, and all nonequilibrium shot noise should arise at the contacts. Factoring in finite temperature contributions, including



thermal noise and that the voltage dropped across each contact is  $IR_c = V_c$ , the total voltage noise in the Corbino should be<sup>33</sup>

$$S_V = 4FeIR_c^2 \coth\left(\frac{eV_c}{2k_B T}\right) + 4k_B T(R_b + 2(1-F)R_c). \quad (6.17)$$

Under this model, at high bias ( $eV_c \gg 2k_B T$ ), the  $S_V(I)$  data can be fit to a line, with slope  $m = 4eFR_c^2$  and intercept  $b = 4k_B T(R_b + 2(1-F)R_c) = 4k_B T(R - 2FR_c)$ . The linear fits can thus be used to derive values for the Fano factor and contact resistance. *Figure 6.22b* is a comparison of Equation 6.17 with  $F$  and  $R_c$  found by the high bias linear fits to the voltage noise measured at 3 K in the low frequency setup. This model seems to explain the voltage scale of the rounding near zero bias, with only a fraction of the applied voltage actually dropping across each shot-noise-producing contact.

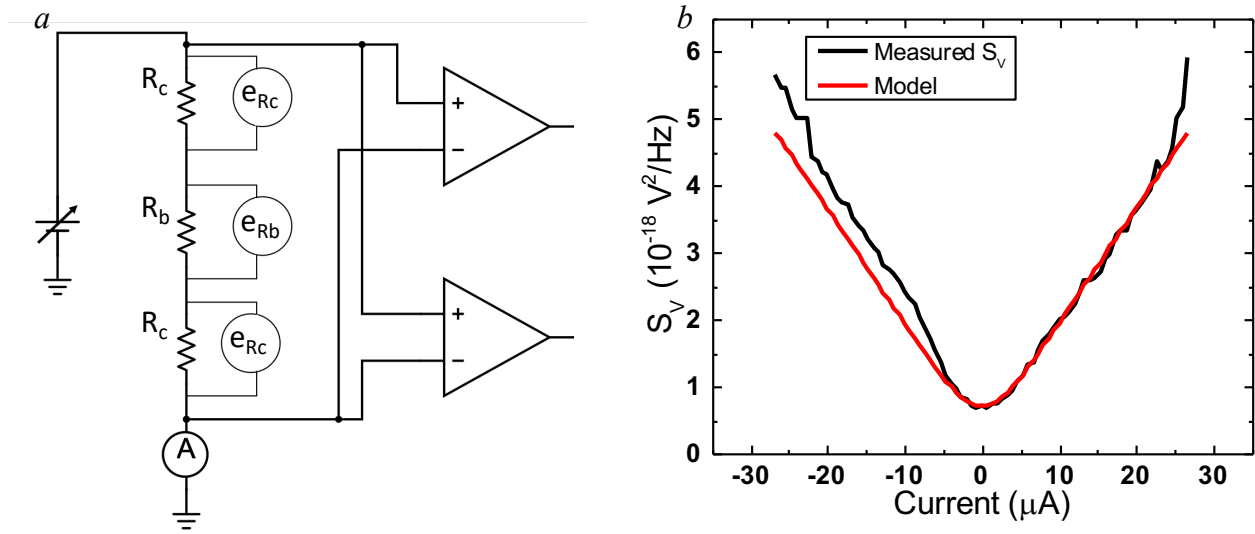


Figure 6.22: [Adapted from Ref. 229] Equivalent circuit (a) and comparison to data (b) of the model described by Equation 6.17, with  $R = 4.34 \text{ k}\Omega$ ,  $F = 5.76 \pm 0.2$ , and  $R_c = 217 \pm 6 \text{ }\Omega$ . Both contact resistances and the resistance of the bulk contribute to the Johnson-Nyquist noise, but any contribution to the shot noise from the bulk is suppressed, leaving only contributions from the two contacts. Voltage division between contacts and bulk accounts for the low bias curvature scale, but implies surprisingly large Fano factors for the shot noise produced by the contacts.

Figure 6.23 plots the Fano factors and contact resistances derived from linear fitting of the high bias data from an example device measured at low frequency. Positive and negative current data were fitted separately due to the asymmetric nature of the  $S_V(I)$  curves (present even when the  $I$ - $V$  response is Ohmic). We attribute this asymmetry to differences in the inner and outer contacts arising during the etching processes of the fabrication. The contact resistances found by the high bias linear fits are generally small, on the order of 10% or less of the total device resistance. While the error bars are obtained from reasonable goodness of fits for each

temperature individually, the degree of scatter among all the data points suggests the systematic uncertainties are comparable to or greater than any trend with temperature. At higher temperatures, the Fermi distribution of carriers in the metal is broadened, and the bulk of the 2D interface (where the GaInSb hole and InAs electron QW states hybridize to form the bulk gap) should have greater thermal activation of carriers between the lower and upper hybridized bands. Additionally, when the Fermi level lies in the gap, disorder creates a spatially varying energy landscape within the 2D bulk, which localizes residual charge carriers as temperature goes to zero.<sup>202</sup> Naively, then, one would expect higher free carrier densities at higher temperatures would yield lower contact resistances as temperature is increased. The contact resistances inferred from our model do not appear to fit this trend, but we have excluded any disorder effects and have not attempted to incorporate the Fermi level pinning or charge transfer at the contacts, which could influence the effective height and width of any barrier at the InAs to InAs/GaInSb interface.

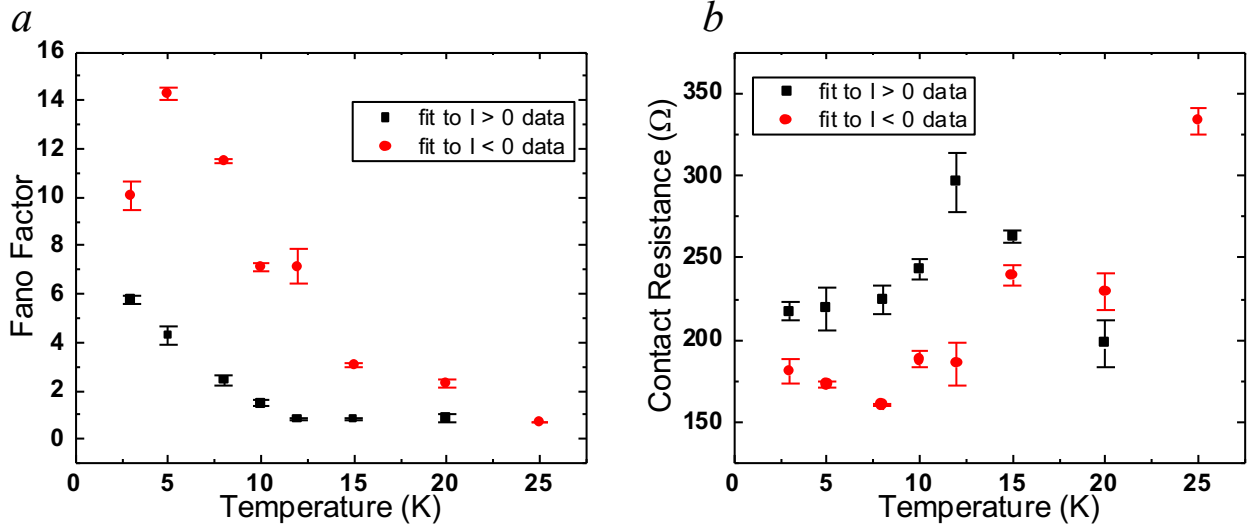


Figure 6.23: [Adapted from Ref. 229] Fano factors (a) and contact resistance (b) derived from linear fits at high bias  $S_V(I)$  data taken at low frequency for the model described by Equation 6.17. Error bars are calculated based on standard deviations of the slope and intercept of the fits. Contact resistance is generally small relative to the total device resistance. Fano factors tend to increase with decreasing temperature.

The inferred Fano factors tend to increase with decreasing temperature, as expected for a Schottky-like contact. In many cases, however, and particularly in the low-temperature limit, the derived values of  $F$  are greater than one, up to over 10 in some cases. This could indicate either some contribution to the shot noise from the bulk that was not captured by the model, or a noise-enhancing process at the contacts. In considering possible contributions from the bulk, we consider that as thermal activation of carriers over the gap in the 2D bulk is suppressed at lower temperatures, disorder can lead to both puddles of charge<sup>149</sup> and residual charge<sup>202</sup> bound to localized states. As discussed in Section 3.7, tunneling through this irregular landscape could contribute to shot noise. The 2D bulk is essentially in the macroscopic limit, however, and if the system were in the limit of many puddles or localized states, the net shot noise should be fully

suppressed, similar to the  $\frac{1}{N}$  reduction of shot noise in  $N$  identical tunnel junctions in series<sup>77</sup> (Section 2.7). This leads us to believe the shot noise arises at a small number of interfaces.

Shot noise-enhancing processes are associated with positive correlations between current pulses. One possibility for a noise-enhancing process at low temperatures in these devices is positive feedback between tunneling electrons and space charge near one of the contacts. This mechanism for enhanced Fano factors was first put forth by Reklaitis and Reggiani,<sup>233</sup> who modelled a single barrier heterostructure of GaAs/Al<sub>0.25</sub>Ga<sub>0.75</sub>As. Enhancement is expected when transport is dominated by tunneling across a barrier, and the space charge region preceding the barrier is sufficiently active to drive the transmission probability.

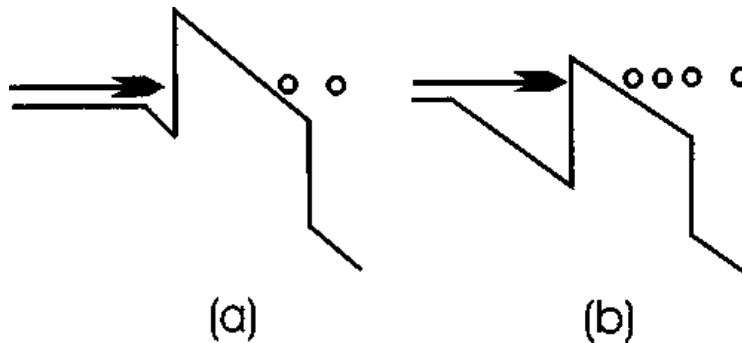


Figure 6.24: [Adapted from Ref. 233] Enhanced noise model of Reklaitis and Reggiani.

Electrons tunneling the barrier (a) decrease the barrier energy as seen by electrons coming from the cathode as a result of the positive feedback between space charge in the layer preceding the barrier and the transmission probability for electrons tunneling through the barrier.

Consequently, more electrons will succeed in tunneling (b).

Figure 6.24 depicts their model, in the limit that the frequency of the current fluctuations is smaller than the inverse of the transit time; an increase in the instantaneous current through the

barrier results in a decrease of the charge density in the potential well, yielding an increase of the potential drop as measured between the left contact and the bottom of the barrier. As a result, the potential barrier seen by electrons from the cathode is reduced, and the electron tunneling probability increases, such that more electrons will pass through the barrier. Conversely, a decrease in the instantaneous current fluctuations, the potential barrier seen by electrons from the cathode is raised, and the electron tunneling probability will decrease, and fewer electrons will tunnel through the barrier. Therefore, the coupling between space charge and tunneling probability amplifies the fluctuations through the Poisson equation, acting as a positive feedback, which should result in an enhancement of the shot noise. Reklaitis and Reggiani found Fano factors up to 7 as a result of the positive feedback between tunneling probability and a build-up of space charge near the barrier. While difficult to precisely model our system for space charge effects, we believe this could be one reasonable explanation for the large Fano factors derived from the data. The tendency for the Fano factor to increase with decreasing temperature, as the bulk becomes more gapped, would be consistent with an increase in space charge near the contacts.

### **6.5.3 Conclusions**

We have performed RF and low frequency noise spectroscopy measurements on InAs/GaInSb Corbino structures to gain a better understanding of the transport properties and shot noise characteristics of the 2D bulk and the contacts. At higher temperatures, voltage noise remains relatively current-independent, staying roughly constant about the Johnson-Nyquist thermal noise level. As temperature is decreased, and the 2D bulk is expected to gap out, however, shot noise becomes detectable. Naively, treating the device as a single noise source,

the magnitude of the shot noise in the high bias, linear limit appears much smaller than  $2eI$ , and the curvature of the noise about zero bias is too broad for the given conductance and temperature of the device. We have found the noise fits well to a model in which the applied bias is dropped over two contacts and the bulk, wherein only the voltage dropped over the contact resistances contributes to the shot noise, while all three contribute to the thermal noise. This model of voltage division between the bulk and contacts reproduces the low bias noise dependence with reasonably small contact resistances, but at the cost of anomalously large Fano factors. The large Fano factors could be the result of some bulk noise contribution not accounted for in the model, or a noise-enhancing process at the contacts. One such process theorized to produce large Fano factors is positive feedback between the tunneling probability and space charge at the interface of the bulk and contacts. These findings show that contacts to bulk 2DTI interfaces can have nontrivial noise response.

## **6.6 Future Directions**

These initial noise studies in InAs/GaSb quantum wells provide an interesting first look at noise processes at work in these two-dimensional topological insulators and are good indicators of promising next steps. For example, it could be informative to perform systematic measurements on  $1/f$  noise in Hall bars of InAs/GaInSb QWs. While there was some  $1/f$  noise at the lowest temperatures of the Corbino structures, we did not observe significantly large  $1/f$  behavior in the spectra. This is in contrast to an ongoing study in which large  $1/f$  signals were measured in samples of monolayer WTe<sub>2</sub> QWs (*Figure 6.25a*). It may be that in the limit of only (ideally single) edge modes,  $1/f$  noise can dominate over shot noise, but in the limit of macroscopic transport through the bulk, samples are less vulnerable to  $1/f$  fluctuators. An

interesting follow-up study could focus on  $1/f$  spectroscopy, particularly in a lower temperature range than that covered above, in the edge transport of InAs/GaInSb bar-geometry devices.

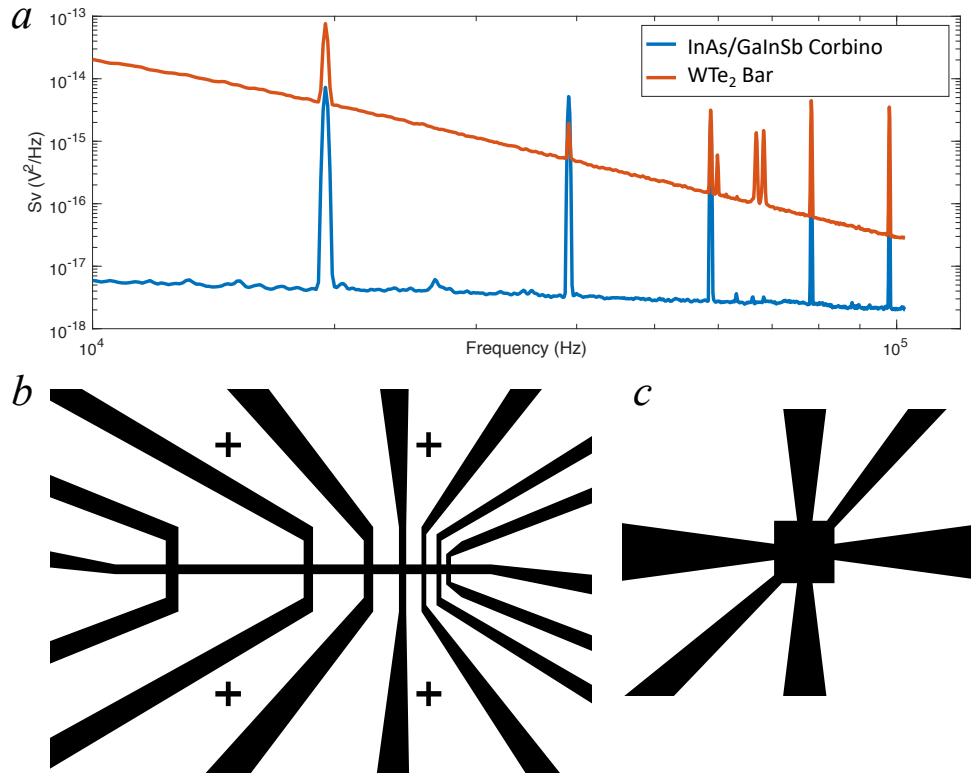


Figure 6.25: a) Comparison of spectra at highest measured current bias for a InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb Corbino device versus a WTe<sub>2</sub> bar structure with edge mode conduction. Despite a much larger current through the Corbino device, the WTe<sub>2</sub> bar exhibits much larger  $1/f$  noise. b) Proposed device configuration for multiple source-drain contact separations along the same edge. This setup could enable direct comparison of noise arising from the edge states when the device length is within the coherence length of the QSHI and when the device length is long, such that the helical edge modes are subject to backscattering processes. c) Another proposed device geometry for comparing noise along the sample edge (using the diagonal contacts) versus through the bulk (using the horizontal or vertical contacts).



Another interesting comparison of the bulk noise versus edge state noise could use a multi-terminal device, similar to *Figure 6.25c*, to measure both  $S_{I,xx}$  and  $S_{I,xy}$ . Any noise arising from the bulk could be directly compared to noise in the edge state while tuning temperature and gate voltage, as the bulk becomes gapped out. Additionally, this may yield more clarity in discerning noise arising from the bulk from that arising at the contact interfaces.

While the fabrication of devices with edge lengths within the coherence length have proven difficult using the mask aligner UV photolithography method, recently acquired instruments may provide a solution for more precise pattern writing. For example, a maskless photolithography system, using high power LED sources, promises submicron resolution and is compatible with many readily available photoresists. The ability to design devices with shorter contact spacing could enable a much wider variety of future studies. One informative example would be to study the shot noise between multiple contacts along the same edge (*Figure 6.25b*), with varying distances between the contacts, ranging from below to beyond the coherence length. This could provide a clear picture of the evolution of noise as the edge modes become vulnerable to backscattering processes. Furthermore, this could be particularly enlightening in understanding the interesting perpendicular magnetic field dependence observed by Du *et al.*,<sup>144,190</sup> who found that the resistance plateaus in InAs/GaInSb QWs under perpendicular magnetic field first increased due to breaking time reversal symmetry, then began to decrease with further increasing of the field. They attribute this behavior to a transition from helical edge states to chiral edge states. Shot noise measurements could be an interesting parameter to observe in this transition, particularly for a long contact separation expected to be vulnerable to scattering in the case of helical transport but not chiral transport. Multiple contacts along a single edge could also allow for more interesting measurement setups including a four-terminal

approach in which two outer contacts are current-biased, while the shot noise is measured between two inner contacts. As fabrication methods improve, the possibilities for noise studies to contribute to the characterization of 2DTIs continue to expand.

# Appendices

## *Appendix 1 1838 Wafer Composition*

Layer # (from bottom)	Layer Name	Thickness (Å)	Material
0	Substrate	-	n-GaSb
1	Smoother	3000	GaSb:Te
2	Nucleation	300	AlSb
3	Relaxer	5000	Al <sub>0.8</sub> Ga <sub>0.2</sub> Sb linearly graded to Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb
4	Lower barrier	500	Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb
5	Channel 1 doping	10	Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb:Te
6	Channel 1 spacer	30	Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb
7	Channel 1	80	InAs
8	Channel 2	40	In <sub>0.32</sub> Ga <sub>0.68</sub> Sb
9	Channel 2 spacer	20	Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb
10	Channel 2 doping	20	Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb:Be
11	Upper barrier	500	Al <sub>0.7</sub> Ga <sub>0.3</sub> Sb
12	Cap 1	30	GaSb
13	Cap 2	30	InAs

## ***Appendix 2 Detailed Procedures for Fabrication of InAs/Ga<sub>0.68</sub>In<sub>0.32</sub>Sb Devices***

### Etchants:

#### *Make ahead of time:*

Citric Acid solution (for use in cocktail etchant and first selective etchant)

- a. 1g:1mL     Monohydrate citric acid:H<sub>2</sub>O
- b. Needs to dissolve at least overnight
- c. Make ~100-200mL at a time
- d. Lasts few weeks

Cocktail etchant

- a. 3:5:55:220 mL     H<sub>3</sub>PO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub>:Citric acid:H<sub>2</sub>O

#### *Make during fabrication:*

Selective etchant #1

- a. 2:1 mL     Citric acid solution:H<sub>2</sub>O<sub>2</sub> (typically 10:5)
- b. Stir, wait 5 min before using

Selective etchant #2

- a. 1:8 mL     NH<sub>4</sub>OH:H<sub>2</sub>O

#### *Use pre-prepared:*

Transetch-N for Al<sub>2</sub>O<sub>3</sub>

- a. Heat in beaker at 70° C, covered in foil

### Fabrication steps:

1. Photolitho the mesa + contacts pattern
  - a. S1813 resist, spin coat at 5000rpm for 1 minute
  - b. Hot plate at 100° C after spin-coating resist
  - c. 135-140 dose
  - d. Photolithography instructions to follow
2. Develop for 30 seconds
  - a. M319 developer
3. Rinse with water, dry with blower, and bake at 105° C for 1 min
4. Let cool for 5 min

5. Soak in cocktail etch for 3-4 min
  - a. Quantum well is 100 nm from the surface
  - b. With back gate, need to go all the way through
  - c. Test the etch rate of new etchants
6. Soak in water for ~ 1 min, then soak in acetone and sonicate to remove resist, IPA, new water, and dry
7. Selective etchant #1 to remove InAs cap
  - a. Mix 10mL citric acid solution with 5mL hydrogen peroxide
  - b. Let sit for ~5 min (can mix up while sample is soaking in water or acetone etc.)
  - c. Soak in selective etch #1 for 45 seconds
8. Soak in water, IPA, then dry
9. Photolitho the contacts
  - a. S1813 resist
  - b. 135 dose
10. Develop for 30-40 seconds
  - a. M319 developer
11. Soak in water, dry, and bake for 1 minute at 105° C
12. Selective etchant #2 to etch down to QWs
  - a. While sample is in photolitho or soaking in water, mix 1:8 of ammonium hydroxide : water
    - i. NH<sub>4</sub>OH smells really bad, only mix under hood
  - b. Soak in SE #2 for 3 minutes
    - i. After ~3 min., the photoresist begins to break down
13. Soak in water, dry
14. Evaporate 10nm Ti / 50-100nm Au
  - a. E-beam evaporator instructions to follow
15. Lift-off
  - a. Soak in acetone in the sonicator for at least 2 minutes
  - b. IPA, water, dry
16. ALD Al<sub>2</sub>O<sub>3</sub>
  - a. 100° C

- b. Choose recipe file (something like Al<sub>2</sub>O<sub>3</sub> 100° C)
  - c. Change inner/outer reactor temps to 100° C
    - i. Takes a bit (at least 20 min) for it to cool down from the default 150° C, so if possible start while doing evaporation or lift-off
  - d. Follow instructions in binder
  - e. Takes about 5 h 40 min to run
    - i. Maximum time able to book on ALDS is 6 hours, so best to do this at night when hopefully no one will need the machine immediately after
  - f. Set heaters back to 150° C when done
17. (Next day) Photolitho windows to contacts
18. Transetch-N for Al<sub>2</sub>O<sub>3</sub>
- a. Pour small amount of etch in beaker, wrap top in foil, and place on hot plate at 70° C
  - b. Soak sample in etch for 4 minutes
19. Soak in water, then acetone to remove resist (sonicate), IPA, fresh water, dry

*Other notes:*

- The process for bar structures is identical, except for the inclusion of an additional photo lithography step at the beginning
  - Before defining any of the device features, pattern and develop a large rectangular surface to remove the rough portions of the resist around the chip edges
- Top side of wafer is teal (at edges can see where layers end)
- Only use the plastic-tipped tweezers
- Wafers are very brittle and will chip easily if using normal tweezers, especially when picking up when taped down
- Have 3 pairs of tweezers on hand: 2 plastic-tipped, one for wet, one for dry, and 1 metal tipped for removing tape from mask aligner, evaporator, spin-coater, etc.
- When soaking sample in etchants, developer, water, acetone, etc. use wet tweezer to move it around a bit
- Be very careful with the water level in the sonicator, if it is too low, vibrations will be too strong and break the wafer

- If it seems like SE #2 didn't reach the QWs, can try annealing in  $H_2N_2$  forming gas at  $250^\circ C$  for 5 min
- In the new cleanroom, use the acid bench under the UV filtered lights for all etching processes except for the Transetch-N, which should be done in the normal light base bench

#### Photolithography Instructions:

1. Spin coat resist onto sample
  - a. Typically S1813
  - b. 5000 rpm for 1 min (standard saved recipe for 1813 in new spin-coater)
  - c. Typically tape off-center (may need to put a larger chip on the vacuum holder in the center)
  - d. Hot plate at  $100^\circ C$  for 1 min
2. Start EVG program
  - a. Login info: user: operator password: retrainme
3. Install 5" holder and plate
4. Follow instructions on computer
  - a. Except: easier to position sample (and tape it down, press tape with glove to make less sticky) when loading mask to get an approximate location
  - b. Also make sure to use other chips on side of sample to prevent too much pressure from breaking sample
5. When finished, remember to log off, take out plate and mask holder, and park the tray

#### E-beam Evaporator Instructions:

1. Grab gold crucible (refill if needed)
2. Press stop. Pressure will reach  $\sim 7E2$
3. Tape sample to plate
4. Open door
5. Use screwdriver to fix plate to ceiling of evaporator
6. Turn to empty gold slot, and load gold crucible (don't forget to put in beads at bottom first)  
Turn back to Ti

7. Make sure substrate shutter is closed
8. Close door and write initials on board
9. Press start
10. Wait to make sure pressure goes down (may need to press on door for a bit)
11. After ~3h, turn on ion gauge to check pressure
12. If  $P \sim 10^{-5-6}$ , it's ok to proceed
13. Flip the two switches on the far left (in new cleanroom, these are no longer in the main room, so skip this step)
14. Turn on main power
15. Turn on power supply controller
16. High voltage: press reset, then on
17. Select film, and check that all the numbers are correct
18. Double click start, then hit manual
19. When rate is steady, open the shutter and zero the thickness
20. Close shutter and bring rate back down to zero, stop
21. Let crucible cool, then repeat for Au
22. Turn everything off and let crucible cool for ~5 min
23. Vent, open door, and remove sample
24. Lift-off in acetone, sonicate



## ***Appendix 3 Procedure for Cooling Down Janis Cryostat***

### *Beginning at Room Temperature*

If the system has been warm for a while (and particularly if it warmed up after being cooled to helium temperatures), it is best to pump the dewar jacket before cooling back down

1. To pump out the jacket, use the red Edwards mechanical pump. Roll it into the Faraday cage by using the wooden ramp.
2. Double check that the valve to the jacket is closed (left side, bottom, with white metal “fences” on either side)
3. Connect the end of tube on the pump to the jacket valve
4. Pump smaller areas first: 1. All knobs closed (make sure pressure is dropping), 2. Open largest, bottom knob, 3. Open knob to tube
  - a. \*Do not open to a high pressure area (e.g. tube area which will be around atmospheric pressure) if the pump has reached max speed. Stop pump, allow motor speed to drop to at most 10k rpm, then open knob slowly and restart pump
5. Let the pressure reach at least  $1\text{E-}4$  mbar before slowly opening the valve to the dewar jacket
6. Pumping out the jacket will take several hours/all day depending on the time since last pumping
  - a. \*Do not pump overnight. If there is a power outage or something, and the jacket reaches atm it will be a big pain to pump down from scratch.
7. When the pressure has reached  $\sim 1\text{E-}5$  mbar (check notes on side for previous recordings), close valve tightly, close all knobs on pump and turn off pump. Remove tube and replace cap on valve.

### *Liquid Nitrogen Fill:*

1. Take off the pipe at the exhaust port
2. Bring the liquid nitrogen dewar up to the door of the Faraday cage
3. Connect the  $\text{LN}_2$  transfer line (rubber tube connected to metal tube) to the liquid port on the dewar. Use the adjustable wrench to ensure a tight fit
4. Open the back filling port and insert metal end of transfer line

5. Open liquid port on the dewar (not too much at first, just so that you can tell liquid is flowing)
6. Hold the rubber tubing near the cryostat so that it doesn't kink until it is frozen
7. After a few minutes, you can open the dewar more.
  - a. It's open too much if the line starts to swing or shake.
8. The fill is done when liquid helium starts to pour out of the exhaust.
9. Close the dewar and warm up the line with the heat gun.
10. When the rubber is no longer rigid, put on gloves, and remove the metal tube from the cryostat. Replace the cap on the fill port.
11. Use the heat gun to warm up the connection at the dewar, and then unscrew the line.
12. Use the heat gun to warm up the exhaust, then wipe down and replace the pipe.

#### *Liquid Helium Fill:*

##### Need:

1. At least 2/3 dewar of liquid helium
2. Helium transfer line (if it has been awhile since last fill, pump the vacuum jacket on the line day before or during first few steps)
3. Helium gas
4. Liquid nitrogen and empty liquid nitrogen dewar
  - a. Preferable to be able to use up a dewar in first step, then re-use it for next part
  - b. At the very least use a dewar that was recently emptied (and is therefore still cold so will be faster to refill)

##### Steps:

1. Top off the liquid nitrogen (see steps for LN<sub>2</sub> fill above) and let sit for at least an hour
2. Open vent on nitrogen dewar, and let pressure drop to zero, keep open
3. Purge the annular space with helium gas
  - a. Close needle valve
  - b. Close the valve to the annular space pump
  - c. Connect the helium gas to the annular helium gas port (gold, in back right)
    - i. Don't forget to let gas run a bit to purge line

- d. Open to the gas until the release valve pops up
  - e. Close gas, and open to pump down
  - f. Repeat 3 times, closing to the pump at the end
4. Remove caps to both filling ports and the pipe extension of the exhaust
5. Replace the pipe extension with a cap
6. Connect copper pipe connector (keep closed at this point) to the helium gas tube and to front filling port
  - a. Fit is not well-sealed, use Teflon tape to prevent leaking (listen for sound, feel for air)
7. Set up the LN<sub>2</sub> filling line as usual, taking extra care to make sure tube reaches the bottom of the reservoir
  - a. If there is hissing around the tube, may also need to Teflon tape this port as well
8. Open liquid valve on the dewar
  - a. Liquid will start flowing quickly after initial high pressure, be quick about getting back to hold the line to prevent kinks after opening the dewar
  - b. If you're too slow, use the heat gun to warm up the line enough to get the tube in a favorable position
9. Open the valve on the copper pipe to allow helium gas to pressurize the reservoir
  - a. Check pressure on He gas regulator, want P~4-5 psi
10. Signs the LN<sub>2</sub> is all out of the reservoir
  - a. Will take ~1.5-2 hours if the dewar is cold (much longer if not)
  - b. If starting with a full cylinder, cylinder pressure will drop from ~1600 to ~1200
  - c. There will be a distinct change in the sound of the nitrogen gas venting from the dewar
  - d. Eventually the tube will go limp (you're very late at this point)
11. Let the helium gas continue to push out any remaining liquid nitrogen while warming the transfer line with the heat gun
12. Remove the LN<sub>2</sub> tube and replace the cap on the back fill port
13. Keep the He gas open until it stops flowing (internal reservoir pressure ~5 psi)
  - a. Meanwhile, close off dewar, put away LN<sub>2</sub> line
14. Close to the helium gas at the copper tube

15. Open the knob closest to the back wall (to the rough pump)
16. When pressure has dropped, VERY gradually open the small, red-topped knob to pump out any remaining nitrogen and the helium gas
17. Pressure will eventually reach a stable pressure (manual says 65 mTorr, but usually we only get to 1-3 Torr)
18. Repeat steps 13-17 at least one more time (This part also takes a couple hours)
19. After pumping, close both knobs, and reopen He gas to backfill the reservoir
20. Meanwhile, in the pump room, connect the right hand pump to the helium recovery line
21. Open both switches to helium recovery
22. Close helium gas and remove copper pipe. Replace cap on front filling port.
23. Connect mesh tube to the exhaust port
24. Set up the ramp and open the ceiling hatch inside the Faraday cage
25. Put the O-ring, inner metal ring, and outer cap to the top of the LHe dewar
26. Bring the dewar up the ramp so that the front wheels are hooked on the doorway of the Faraday cage. Place a chair or step stool at the base of the ramp
27. Grab the gloves, and form a bow and arrow with the transfer line
28. After relieving some pressure from the dewar (using gas port, not top), open the top of the dewar and insert the end of the transfer line without the vacuum port slowly into the dewar while the other end rests across the top of the dewar
  - a. Insert just enough to be able to fit everything through the doorway
29. Connect the helium gas to the vent on the dewar (don't open yet)
  - a. Want a relatively low pressure to start (~0.5-1.5 psi), once there is liquid, can increase pressure
30. Wait until there is white steam flowing from the end of the transfer line, then open back filling port and prepare to insert the line into the reservoir
31. When you can see the characteristic white flame, insert the transfer line as quickly as possible
  - a. Ideally, it should be a straight shot down, but in practice, getting the transfer line down into the reservoir takes some wiggling and searching
32. Open the needle valve, and open the annular space pump to cool the sample space
33. Monitor the temperature and cylinder pressure

34. Once  $T \sim 30\text{K}$ , there should be liquid collecting, and you can increase the pressure and turn on the helium monitor
  - a. Set to continuous monitoring during fill, turn back to on demand after
35. Want at least 20 inches (ideally 23-24")
36. When ready to remove the transfer line, close the helium gas and remove the gas line
37. Using gloves, open the vent on the dewar and pull the line up out of the dewar part of the way (so that it is out of the liquid inside the dewar)
38. Quickly pull the other end out of the fill port (this is when it is important to have the ceiling hatch open)
39. Quickly cap the port
40. Roll dewar back out and pull transfer line out all of the way
41. Put dewar back on the wall and connect to recovery line
  - a. Return dewars and send email to Quinn for refill as soon as they are empty (so they don't have time to warm up) for faster refills
42. To cool down the sample below 4 K, open the needle valve all the way
  - a. Need at least 3-4 Torr to stabilize at low temperatures (below 30 K)

## ***Appendix 4 Conversion from Lock-In to Real Units for RF Measurements***

Let  $X_i$  be the number returned by the lock-in. This number reflects the RMS value, so to get to the amplitude in units of volts  $X_{A,V}$

$$X_{A,V} = \frac{2X_i}{0.901 * 10} * (\text{sensitivity setting of Lock - in})$$

The 2 accounts for going from RMS to full amplitude, the 0.901 comes from taking the Fourier transform of a square wave, and the 10 is the scaling factor of the lock-in.

The next step is to convert to dB based on the weighted average fit of the PE8040 log power detector typical response plot

$$X_{dB} = -41.2587211X_{A,V} + 27.2182545 - 30$$

Then to convert to  $A^2/Hz$

$$X_{SI} = 10^{X_{dB}/10}$$

Now, to arrive at the final value, we need to compare the lock-in value to the value from the DAQ (the DC baseline power)

$$\begin{aligned} DC \text{ Power (V)} - \frac{1}{2}(Lock - in)(V) &\rightarrow dB - GBWP \rightarrow A^2/Hz \\ &- \\ DC \text{ Power (V)} + \frac{1}{2}(Lock - in)(V) &\rightarrow dB - GBWP \rightarrow A^2/Hz \\ &= S_I (A^2/Hz) \end{aligned}$$

where  $GBWP$  is the previously measured gain-bandwidth product (161.639 for Janis system with attocube probe).

Calibration to account for the pick-up efficiency of the measurement system are ongoing and will result in additional considerations in the determination of the accurate value of the noise power.

1. Ohm, G. *Die galvanische Kette, mathematisch bearbeitet*. (T.H. Riemann, 1827).
2. Drude, P. Zur elektronentheorie der metalle. *Ann. Phys.* **306**, 556–613 (1900).
3. Kittel, C. *Introduction to Solid State Physics*. (John Wiley & Sons, 1971).
4. Natelson, D. *Nanostructures and Nanotechnology*. (Cambridge University Press, 2015).
5. Landau, L. D. The Theory of a Fermi Liquid. *Sov. Phys. J. Exp. Theor. Phys.* **3**, 920–925 (1957).
6. Pines, D. & Nozieres, P. *The Theory of Quantum Liquids, Vol. 1*. (Addison Wesley, 1989).
7. Baym, G. & Pethick, C. *Landau Fermi-Liquid Theory: Concepts and Applications*. (Wiley, 1992).
8. Datta, S. *Electronic Transport in Mesoscopic Systems*. (Cambridge University, 1995).
9. Sharvin, Y. & Bogatina, N. Investigation of Focusing of Electron Beams in a Metal by a Longitudinal Magnetic Field. *Sov. Phys. J. Exp. Theor. Phys.* **29**, 419–423 (1969).
10. Landauer, R. Conductance from Transmission: Common Sense Points. *Phys. Scr.* **T42**, 110 (1992).
11. Landauer, R. Electrons as Guided Waves in Laboratory Structures: Strengths and Problems. in *Analogies in Optics and Micro Electronics* (eds. VanHaeringen, W. & Lenstra, D.) 243–257 (1990).
12. Buttiker, M. Four-Terminal Phase-Coherent Conductance. *Phys. Rev. Lett.* **57**, 1761 (1986).
13. Van Wees, B. J. *et al.* Quantized conductance of point contacts in a two-dimensional electron gas. *Phys. Rev. Lett.* **60**, 848–850 (1988).
14. Wharam, D. A. *et al.* One-dimensional transport and the quantisation of the ballistic

- resistance. *J. Phys. C Solid State Phys.* **21**, L209 (1988).
15. Agraït, N., Rodrigo, J. G. & Vieira, S. Conductance steps and quantization in atomic-size contacts. *Phys. Rev. B* **47**, 12345 (1993).
  16. Pascual, J. I., Mendez, J., Gomez-Herrero, J., Baro, A. M. & Garcia, N. Quantum Contact in Gold Nanostructures by Scanning Tunneling Microscopy. *Phys. Rev. Lett.* **71**, 1852–1855 (1993).
  17. Gai, Z., He, Y., Yu, H. & Yang, W. S. Observation of conductance quantization of ballistic metallic point contacts at room temperature. *Phys. Rev. B* **53**, 1042–1045 (1996).
  18. Yanson, A. I., Bollinger, G. R., van den Brom, H. E., Agraït, N. & van Ruitenbeek, J. M. Formation and manipulation of a metallic wire of single gold atoms. *Nature* **395**, 783–785 (1998).
  19. Muller, C. J., van Ruitenbeek, J. M. & de Jongh, L. J. Conductance and Supercurrent Discontinuities in Atomic-Scale Metallic Constrictions of Variable Width. *Phys. Rev. Lett.* **69**, 140–143 (1992).
  20. Krans, J. M., van Ruitenbeek, J. M., Fisunt, V. V., Yansont, I. K. & de Jongh, L. J. The signature of conductance quantization in metallic point contacts. *Nature* **375**, 767–769 (1995).
  21. Muller, C. J., Krans, J. M., Todorov, T. N. & Reed, M. A. Quantization effects in the conductance of metallic contacts at room temperature. *Phys. Rev. B* **53**, 1022–1025 (1996).
  22. Scheer, E., Joyez, P., Esteve, D., Urbina, C. & Devoret, M. H. Conduction Channel Transmissions of Atomic-Size Aluminum Contacts. *Phys. Rev. Lett.* **78**, 3535 (1997).
  23. Scheer, E. *et al.* The signature of chemical valence in the electrical conduction through a



- single-atom contact. *Nature* **394**, 6689 (1998).
24. van den Brom, H. E., Yanson, A. I. & van Ruitenbeek, J. M. Characterization of individual conductance steps in metallic quantum point contacts. *Phys. B* **252**, 69–75 (1998).
  25. Yanson, I. K. *et al.* Atomic-Size Oscillations in Conductance Histograms for Gold Nanowires and the Influence of Work Hardening. *Phys. Rev. Lett.* **95**, 256806 (2005).
  26. Krans, J. M. No Title. (Leidan, 1996).
  27. Rubio, G., Agrait, N. & Vieira, S. Atomic-Sized Metallic Contacts : Mechanical Properties and Electronic Transport. *Phys. Rev. Lett.* **76**, 2302 (1996).
  28. Yanson, A. I. & van Ruitenbeek, J. M. Do Histograms Constitute a Proof for Conductance Quantization? *Phys. Rev. Lett.* **79**, 2157 (1997).
  29. Johnson, J. B. Thermal agitation of electricity in conductors. *Phys. Rev.* **32**, 97–109 (1928).
  30. Nyquist, H. Thermal agitation of electric charge in conductors. *Phys. Rev.* **32**, 110–113 (1928).
  31. Dutta, P. & Horn, P. M. Low-frequency fluctuations in solids: 1/f noise. *Rev. Mod. Phys.* **53**, 497–516 (1981).
  32. Weissman, M. B. 1/f noise and other slow, nonexponential kinetics in condensed matter. *Rev. Mod. Phys.* **60**, 537–571 (1988).
  33. Blanter, Y. M. & Buttiker, M. Shot Noise in Mesoscopic Conductors. *Phys. Rep.* **336**, 1–166 (2000).
  34. Kubo, R. The fluctuation-dissipation theorem. *Reports Prog. Phys.* **29**, 255 (1966).
  35. Callen, H. B. & Welton, T. A. Irreversibility and Generalized Noise. *Phys. Rev. B* **83**,

- (1951).
36. Schiek, B., Rolfes, I. & Siweris, H.-J. *Noise in High-Frequency Circuits and Oscillators*. (John Wiley & Sons, 2006).
  37. Johnson, J. B. The Schottky effect in low frequency circuits. *Phys. Rev.* **26**, 71 (1925).
  38. Johnson, J. B. Thermal Agitation of Electricity in Conductors. *Nature* **119**, 50 (1927).
  39. Wu, Z. *et al.* Scaling of 1/f noise in tunable break junctions. *Phys. Rev. B* **78**, 235421 (2008).
  40. Voss, R. F. & Clarke, J. '1/f noise' in music and speech. *Nature* **258**, 317–318 (1975).
  41. Holweg, P. A. M., Caro, J., Verbruggen, A. H. & Radelaar, S. Ballistic electron transport and two-level resistance fluctuations in noble-metal nanobridges. *Phys. Rev. B* **45**, 9311 (1992).
  42. Ralls, K. S., Ralph, D. C. & Buhrman, R. A. Individual-defect electromigration in metal nanobridges. *Phys. Rev. B* **40**, 11561 (1989).
  43. van Harlingen, D. J. *et al.* Decoherence in Josephson-junction qubits due to critical-current fluctuations. *Phys. Rev. B* **70**, 064517 (2004).
  44. Vandamme, L. 1/f noise in CMOS transistors. in *Proceedings of the 10th International Conference on Noise in Physical Systems* 21–25 (1989).
  45. Ishigami, M. *et al.* Hooge 's constant for carbon nanotube field effect transistors. *Appl. Phys. Lett.* **88**, 203116 (2006).
  46. Schottky, W. Uber spontane Stromschwankungen in verschiedenen Electrizaritsleitern. *Ann. Phys.* **362**, 541 (1918).
  47. von Bortkiewicz, L. *Das Gesetz der kleinen Zahlen*. (B.B. Teubner, 1898).
  48. Student. On the Error of Counting with a Haemocytometer. *Biometrika* **5**, 351 (1907).

49. Rice, F. A frequency-domain derivation of shot-noise. *Am. J. Phys.* **84**, 44 (2016).
50. Fano, U. Ionization Yield of Radiations II. The Fluctuations of the Number of Ions. *Phys. Rev.* **72**, 26 (1947).
51. van den Brom, H. E. & van Ruitenbeek, J. M. Quantum suppression of shot noise in atom-size metallic contacts. *Phys. Rev. Lett.* **82**, 1526 (1999).
52. Djukic, D. & van Ruitenbeek, J. M. Shot Noise Measurements on a Single Molecule. *ACS Nano Lett.* **6**, 789–793 (2006).
53. Vardimon, R., Klionsky, M. & Tal, O. Experimental determination of conduction channels in atomic-scale conductors based on shot noise measurements. *Phys. Rev. B* **88**, 161404 (2013).
54. Steinbach, A. H., Martinis, J. M. & Devoret, M. H. Observation of Hot-Electron Shot Noise in a Metallic Resistor. *Phys. Rev. Lett.* **76**, 3806 (1996).
55. Kumar, M., Avriller, R., Yeyati, A. L. & van Ruitenbeek, J. M. Detection of vibration-mode scattering in electronic shot noise. *Phys. Rev. Lett.* **108**, 146602 (2012).
56. Kumar, A., Saminadayar, L., Glatli, D. C., Jin, Y. & Etienne, B. Experimental Test of the Quantum Shot Noise Reduction Theory. *Phys. Rev. Lett.* **76**, 2778 (1996).
57. Reznikov, M., Heiblum, M., Shtrikman, H. & Mahalu, D. Temporal Correlation of Electrons-Suppression of Shot Noise in a Ballistic Quantum Point Contact. *Phys. Rev. Lett.* **75**, 3340–3343 (1995).
58. van den Brom, H. E. & van Ruitenbeek, J. M. Shot noise suppression in metallic quantum point contacts. in *Statistical and Dynamical Aspects of Mesoscopic Systems* (eds. Reguera, D., Platero, G., Bonilla, L. L. & Rubi, J. M.) 114–122 (Springer, 2000).
59. Chen, R. *et al.* Shot noise variation within ensembles of gold atomic break junctions at

- room temperature. *J. Phys. Condens. Matter* **26**, 474204 (2014).
60. Tal, O., Krieger, M., Leerink, B. & van Ruitenbeek, J. M. Electron-Vibration Interaction in Single-Molecule Junctions: From Contact to Tunneling Regimes. *Phys. Rev. Lett.* **100**, 196804 (2008).
  61. Kiguchi, M. *et al.* Highly Conductive Molecular Junctions Based on Direct Binding of Benzene. *Phys. Rev. Lett.* **101**, 046801 (2008).
  62. Kumar, M. *et al.* Shot noise and magnetism of Pt atomic chains : Accumulation of points at the boundary. *Phys. Rev. B* **88**, 245431 (2013).
  63. Vardimon, R., Klionsky, M. & Tal, O. Indication of Complete Spin Filtering in Atomic-Scale Nickel Oxide. *ACS Nano Lett.* **15**, 3894 (2015).
  64. Zhou, P., Hardy, W. J., Watanabe, K., Taniguchi, T. & Natelson, D. Shot noise detection in hBN-based tunnel junctions Shot noise detection in hBN-based tunnel junctions. *Appl. Phys. Lett.* **110**, 133106 (2017).
  65. Spietz, L., Lehnert, K. W., Siddiqi, I. & Schoelkopf, R. J. Primary Electronic Thermometry Using the Shot Noise of a Tunnel Junction. *Science (80-. )*. **300**, 1929 (2003).
  66. Li, Y. P., Zaslavsky, A., Tsui, D. C., Santos, M. & Shayegan, M. Noise characteristics of double-barrier resonant-tunneling structures below 10 kHz. *Phys. Rev. B* **41**, 8388 (1990).
  67. Liu, H. C. *et al.* Shot-noise suppression in resonant tunneling. *Phys. Rev. B* **51**, 5116 (1995).
  68. Safonov, S. S. *et al.* Enhanced Shot Noise in Resonant Tunneling via Interacting Localized States. *Phys. Rev. Lett.* **91**, 136801 (2003).
  69. Birk, H., de Jong, M. J. . & Schonberger, C. Shot-Noise Suppression in the Single-

- Electron Tunneling Regime. *Phys. Rev. Lett.* **75**, 1610 (1995).
70. Chen, L. Y. & Ting, C. S. Theoretical investigation of noise characteristics of double-barrier resonant-tunneling systems. *Phys. Rev. B* **43**, 4534 (1991).
  71. Davies, J. H., Hyldgaard, P., Hershfield, S. & Wilkins, J. W. Classical theory of shot noise in resonant tunneling. *Phys. Rev. B* **46**, 9620 (1992).
  72. de Jong, M. J. M. & Beenakker, C. W. J. Semiclassical theory of shot noise in mesoscopic conductors. *Physica A* **230**, 219–248 (1996).
  73. Beenakker, C. W. J. & Büttiker, M. Suppression of shot noise in metallic diffusive conductors. *Phys. Rev. B* **46**, 1889 (1992).
  74. de Jong, M. J. . & Beenakker, C. W. J. Semiclassical theory of shot noise suppression. *Phys. Rev. B* **51**, 16867 (1995).
  75. Beenakker, C. W. J. & Melsen, J. A. Conductance fluctuations, weak localization, and shot noise for a ballistic constriction in a disordered wire. *Phys. Rev. B* **50**, 2450 (1994).
  76. Nazarov, Y. V. Limits of Universality in Disordered Conductors. *Phys. Rev. Lett.* **73**, 134 (1994).
  77. Landauer, R. Mesoscopic noise: Common sense view. *Phys. B* **227**, 156–160 (1996).
  78. Kozub, V. I. & Rudin, A. M. Shot noise in mesoscopic diffusive conductors in the limit of strong electron-electron scattering. *Phys. Rev. B* **52**, 7853 (1995).
  79. Kanskar, M. & Wybourne, M. N. Crossover between Dissipative and Nondissipative Electron Transport in Metal Wires. *Phys. Rev. Lett.* **73**, 2123 (1994).
  80. Prober, D. E. Comment on ‘Crossover Between Dissipative and Nondissipative Electron Transport in Metal Wires’. *Phys. Rev. Lett.* **75**, 3964 (1995).
  81. Nagaev, K. E. Influence of electron-electron scattering on shot noise in diffusive contacts.

- Phys. Rev. B* **52**, 4740 (1995).
82. Oberholzer, S., Sukhorukov, E. V., Strunk, C. & Schonenberger, C. Shot noise of series quantum point contacts intercalating chaotic cavities. *Phys. Rev. B* **66**, 233304 (2002).
  83. Landau, L. D. On the theory of phase transitions. *J. Exp. Theor. Phys.* **7**, 19–32 (1937).
  84. von Klitzing, K., Dorda, G. & Pepper, M. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance. *Phys. Rev. Lett.* **45**, 494 (1980).
  85. Thouless, D. J., Kohmoto, M., Nightingale, M. P. & den Nijs, M. Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Phys. Rev. Lett.* **49**, 405–408 (1982).
  86. Qi, X. L. & Zhang, S. C. Topological insulators and superconductors. *Rev. Mod. Phys.* **83**, 1057 (2011).
  87. Hall, E. H. On a New Action of the Magnet on Electric Currents. *Am. J. Math.* **2**, 287–292 (1879).
  88. Cage, M. E., Dziuba, R. F. & Field, B. F. A Test of the Quantum Hall Effect as a Resistance Standard. *IEEE Trans. Instrum. Meas.* **34**, 301 (1985).
  89. von Klitzing, K. Developments in the quantum Hall effect. *Philos. Trans. R. Soc. A* **363**, 2203–2219 (2005).
  90. Davies, J. H. *The Physics of Low-dimensional Semiconductors: An Introduction*. (Cambridge University Press, 1998).
  91. Tong, D. Lectures on the Quantum Hall Effect. (2016).
  92. Halperin, B. I. Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential. *Phys. Rev. B* **25**,

- 2185 (1982).
93. Laughlin, R. B. Quantized Hall conductivity in two dimensions. *Phys. Rev. B* **23**, 5632 (1981).
  94. Yoshioka, D. *The Quantum Hall Effect*. (Springer, 1998).
  95. Laughlin, R. B. Fractional Quantization. (1998).
  96. Prange, R. E. Quantized Hall resistance and the measurement of the fine-structure constant. *Phys. Rev. B* **23**, 4802 (1981).
  97. Kubo, R. Statistical-Mechanical Theory of Irreversible Processes I. *J. Phys. Soc. Japan* **12**, 570 (1957).
  98. Thouless, D. J. *Topological Quantum Numbers in Nonrelativistic Physics*. (World Scientific Publishing Company, 1998).
  99. Avron, J. E. & Seiler, R. Quantization of the Hall Conductance for General, Multiparticle Schrodinger Hamiltonians. *Phys. Rev. Lett.* **54**, 259 (1985).
  100. Avron, J. E., Osadchy, D. & Seiler, R. A Topological Look at the Quantum Hall Effect. *Phys. Today* **56**, 38 (2003).
  101. Niu, Q. & Thouless, D. J. Quantum Hall effect with Realistic boundary conditions. *Phys. Rev. B* **35**, 2188 (1987).
  102. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. London A* **392**, 45–57 (1984).
  103. Chern, S. Characteristic Classes of Hermitian Manifolds. *Ann. Math. Second Ser.* **47**, 85–121 (1946).
  104. Hasan, M. Z. & Kane, C. L. Colloquium : Topological insulators. *Rev. Mod. Phys.* **82**, 3045–3067 (2010).

105. Dresselhaus, M. S., Dresselhaus, G. & Jorio, A. *Group Theory: Application to the Physics of Condensed Matter*. (Springer, 2008).
106. Haldane, F. D. M. Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the ‘Parity Anomaly’. *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
107. DiVincenzo, D. P. & Mele, E. J. Self-consistent effective-mass theory for intralayer screening in graphite intercalation compounds. *Phys. Rev. B* **29**, 1685 (1984).
108. Semenoff, G. W. Condensed-Matter Simulation of a Three-Dimensional Anomaly. *Phys. Rev. Lett.* **55**, 2449 (1984).
109. Hsieh, D. *et al.* A tunable topological insulator in the spin helical Dirac transport regime. *Nature* **460**, 1101–1105 (2009).
110. Wright, A. R. Realising Haldane’s vision for a Chern insulator in buckled lattices. *Nat. Sci. Reports* **3**, 2736 (2013).
111. Kim, H. & Kee, H. Realizing Haldane model in Fe-based honeycomb ferromagnetic insulators. *Quantum Mater.* **20**, (2017).
112. Jotzu, G. *et al.* Experimental realization of the topological Haldane model with ultracold fermions. *Nature* **5115**, 237 (2014).
113. Kane, C. L. & Mele, E. J. Z<sub>2</sub> Topological Order and the Quantum Spin Hall Effect. *Phys. Rev. Lett.* **95**, 146802 (2005).
114. Kane, C. L. & Mele, E. J. Quantum Spin Hall Effect in Graphene. *Phys. Rev. Lett.* **95**, 226801 (2005).
115. Novoselov, K. S. *et al.* Electric Field Effect in Atomically Thin Carbon Films. *Science* (80-. ). **306**, 666–670 (2004).
116. Geim, A. K. & Novoselov, K. S. The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007).



117. Murakami, S., Nagaosa, N. & Zhang, S.-C. Spin-Hall Insulator. *Phys. Rev. Lett.* **93**, 6–9 (2004).
118. Murakami, S., Nagaosa, N. & Zhang, S.-C. Dissipationless Quantum Spin. *Science (80-. )*. **301**, 1348 (2003).
119. Sinova, J. *et al.* Universal Intrinsic Spin Hall Effect. *Phys. Rev. Lett.* **92**, 126603 (2004).
120. Bernevig, B. A., Hughes, T. L. & Zhang, S.-C. Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells. *Science (80-. )*. **314**, 1757–1761 (2006).
121. König, M. *et al.* Quantum Spin Hall Insulator State in HgTe Quantum Wells. *Science (80-. )*. **318**, 766 (2007).
122. Qi, X.-L. & Zhang, S.-C. Spin-Charge Separation in the Quantum Spin Hall State. *Phys. Rev. Lett.* **101**, 086802 (2008).
123. Fu, L. & Kane, C. L. Time reversal polarization and a Z<sub>2</sub> adiabatic spin pump. *Phys. Rev. B* **74**, 195312 (2006).
124. Fukui, T. & Hatsugai, Y. Quantum Spin Hall Effect in Three Dimensional Materials: Lattice Computation of Z<sub>2</sub> Topological Invariants and Its Application to Bi and Sb. *J. Phys. Soc. Japan* **76**, 053702 (2007).
125. Fukui, T., Fujiwara, T. & Hatsugai, Y. Topological Meaning of Z<sub>2</sub> Numbers in Time Reversal Invariant Systems. *J. Phys. Soc. Japan* **77**, 123705 (2008).
126. Moore, J. E. & Balents, L. Topological invariants of time-reversal-invariant band structures. *Phys. Rev. B* **75**, 121306 (2007).
127. Qi, X.-L., Hughes, T. L. & Zhang, S.-C. Topological field theory of time-reversal invariant insulators. *Phys. Rev. B* **78**, 195424 (2008).
128. Roy, R. Z<sub>2</sub> classification of quantum spin Hall systems : An approach using time-reversal

- invariance. *Phys. Rev. B* **79**, 195321 (2009).
129. Wang, Z., Qi, X.-L. & Zhang, S.-C. Equivalent topological invariants of topological insulators. *New J. Phys.* **12**, 065007 (2010).
  130. Sheng, D. N., Weng, Z. Y., Sheng, L. & Haldane, F. D. M. Quantum Spin-Hall Effect and Topologically Invariant Chern Numbers. *Phys. Rev. Lett.* **97**, 036808 (2006).
  131. Wu, C., Bernevig, B. A. & Zhang, S.-C. Helical Liquid and the Edge of Quantum Spin Hall Systems. *Phys. Rev. Lett.* **96**, 106401 (2006).
  132. Qi, X.-L. & Zhang, S.-C. The quantum spin Hall effect and topological insulators. *Phys. Today* **63**, 33 (2010).
  133. Xu, C. & Moore, J. E. Stability of the quantum spin Hall effect: Effects of interactions, disorder, and  $Z_2$  topology. *Phys. Rev. B* **73**, 045322 (2006).
  134. Yao, Y., Ye, F., Qi, X.-L., Zhang, S.-C. & Fang, Z. Spin-orbit gap of graphene: First-principles calculations. *Phys. Rev. B* **75**, 041401 (2007).
  135. Fu, L. & Kane, C. L. Topological insulators with inversion symmetry. *Phys. Rev. B* **76**, 045302 (2007).
  136. König, M. *et al.* The Quantum Spin Hall Effect: Theory and Experiment. *J. Phys. Soc. Japan* **77**, 031007 (2008).
  137. Daumer, V. *et al.* Quasiballistic transport in HgTe quantum-well nanostructures. *Appl. Phys. Lett.* **83**, 1376 (2003).
  138. Roth, A. *et al.* Nonlocal Transport in the Quantum Spin Hall State. *Science (80-. )*. **325**, 294 (2009).
  139. Nowack, K. C. *et al.* Imaging currents in HgTe quantum wells in the quantum spin Hall regime. *Nat. Mater.* **12**, 787–791 (2013).

140. Liu, C., Hughes, T. L., Qi, X.-L., Wang, K. & Zhang, S.-C. Quantum Spin Hall Effect in Inverted Type-II Semiconductors. *Phys. Rev. Lett.* **100**, 236601 (2008).
141. Knez, I., Du, R.-R. & Sullivan, G. Evidence for Helical Edge Modes in Inverted InAs/GaSb Quantum Wells. *Phys. Rev. Lett.* **107**, 136603 (2011).
142. Li, X.-B. *et al.* Experimental Observation of Topological Edge States at the Surface Step Edge of the Topological Insulator ZrTe<sub>5</sub>. *Phys. Rev. Lett.* **116**, 176803 (2016).
143. Fei, Z. *et al.* Edge conduction in monolayer WTe<sub>2</sub>. *Nat. Phys.* **13**, 677–683 (2017).
144. Du, L., Knez, I., Sullivan, G. & Du, R.-R. Robust helical edge transport in gated InAs/GaSb bilayers. *Phys. Rev. Lett.* **114**, 096802 (2015).
145. Maciejko, J. *et al.* Kondo Effect in the Helical Edge Liquid of the Quantum Spin Hall State. *Phys. Rev. Lett.* **102**, 256803 (2009).
146. Kondo, J. Resistance Minimum in Dilute Magnetic Alloys. *Prog. Theor. Phys.* **32**, 37–49 (1964).
147. König, M. Spin-related transport phenomena in HgTe-based quantum well structures. (University of Würzburg, 2007).
148. Tkachov, G. & Hankiewicz, E. M. Ballistic quantum spin hall state and enhanced edge backscattering in strong magnetic fields. *Phys. Rev. Lett.* **104**, 1–4 (2010).
149. Väyrynen, J. I., Goldstein, M. & Glazman, L. I. Helical edge resistance introduced by charge puddles. *Phys. Rev. Lett.* **110**, 216402 (2013).
150. Väyrynen, J. I., Goldstein, M., Gefen, Y. & Glazman, L. I. Resistance of helical edges formed in a semiconductor heterostructure. *Phys. Rev. B* **90**, 115309 (2014).
151. Schmidt, T. L. Current correlations in quantum spin hall insulators. *Phys. Rev. Lett.* **107**, 096602 (2011).

152. Dolcini, F. Noise and current correlations in tunnel junctions of quantum spin Hall edge states. *Physical* **92**, 155421 (2015).
153. Rizzo, B., Arrachea, L. & Moskalets, M. Transport phenomena in helical edge state interferometers : A Green's function approach. *Phys. Rev. B* **88**, 155433 (2013).
154. Edge, J. M., Li, J., Delplace, P. & Buttiker, M. Z<sub>2</sub> Peak of Noise Correlations in a Quantum Spin Hall Insulator. *Phys. Rev. Lett.* **110**, 246601 (2013).
155. Maestro, A. Del. Backscattering between helical edge states via dynamic nuclear polarization. *Phys. Rev. B* **87**, 165440 (2013).
156. Aseev, P. P. & Nagaev, K. E. Shot noise in the edge states of two-dimensional topological insulators. *Phys. Rev. B* **94**, 045425 (2016).
157. Väyrynen, J. I. & Glazman, L. I. Current Noise from a Magnetic Moment in a Helical Edge. *Phys. Rev. Lett.* **118**, 106802 (2017).
158. Kurilovich, P. D., Kurilovich, V. D., Burmistrov, I. S., Gefen, Y. & Goldstein, M. Unrestricted electron bunching at the helical edge. *Phys. Rev. Lett.* (2019).
159. Nagaev, K. E., Remizov, S. V. & Shapiro, D. S. Noise in the Helical Edge Channel Anisotropically Coupled to a Local Spin. *J. Exp. Theor. Phys. Lett.* **108**, 664–669 (2018).
160. König, M. *et al.* Spatially resolved study of backscattering in the quantum spin Hall state. *Phys. Rev. X* **3**, 021003 (2013).
161. Grabecki, G. *et al.* Nonlocal resistance and its fluctuations in microstructures of band-inverted HgTe/(Hg,Cd)Te quantum wells. *Phys. Rev. B* **88**, 165309 (2013).
162. Tikhonov, E. S. *et al.* Shot noise of the edge transport in the inverted band HgTe quantum wells. *J. Exp. Theor. Phys. Lett.* **101**, 708–713 (2015).
163. Cascales, J. P. *et al.* Band structure of topological insulators from noise measurements in

- tunnel junctions. *Appl. Phys. Lett.* **107**, 252402 (2015).
164. Kimme, L., Rosenow, B. & Brataas, A. Backscattering in helical edge states from a magnetic impurity and Rashba disorder. *Phys. Rev. B* **93**, 081301 (2016).
  165. Kurilovich, P. D., Kurilovich, V. D., Burmistrov, I. S. & Goldstein, M. Helical edge transport in the presence of a magnetic impurity. *J. Exp. Theor. Phys. Lett.* **106**, 593 (2017).
  166. Wu, F. *et al.* Setup for shot noise measurements in carbon nanotubes. in *24th International Conference on Low Temperature Physics. AIP Conference Proceedings* 1482–1483 (2006).
  167. Reznikov, M., Heiblum, M., Shtrikman, H. & Mahalu, D. Temporal Correlation of Electrons: Suppression of Shot Noise in a Ballistic Quantum Point Contact. *Phys. Rev. Lett.* **75**, 3340 (1995).
  168. Stevens, L. A., Zolotavin, P., Chen, R. & Natelson, D. Current noise enhancement: Channel mixing and possible nonequilibrium phonon backaction in atomic-scale Au junctions. *J. Phys. Condens. Matter* **28**, 495303 (2016).
  169. Wheeler, P. J., Russom, J. N., Evans, K., King, N. S. & Natelson, D. Shot Noise Suppression at Room Temperature in Atomic-Scale Au Junctions. *ACS Nano Lett.* **10**, 1287–1292 (2010).
  170. Hashisaka, M., Ota, T., Yamagishi, M., Fujisawa, T. & Muraki, K. Cross-correlation measurement of quantum shot noise using homemade transimpedance amplifiers Cross-correlation measurement of quantum shot noise using homemade transimpedance amplifiers. *Rev. Sci. Instrum.* **85**, 054704 (2014).
  171. Zhou, P. *et al.* Electron pairing in the pseudogap state revealed by shot noise in copper-

- oxide junctions. *Nature* Submitted (2019).
172. Chen, R., Wheeler, P. J., Di Ventra, M. & Natelson, D. Enhanced noise at high bias in atomic-scale Au break junctions. *Sci. Rep.* **4**, 4221 (2014).
  173. Henny, M., Oberholzer, S., Strunk, C. & Schonberger, C. 1/3-shot-noise suppression in diffusive nanowires. *Phys. Rev. B* **59**, 2871 (1999).
  174. Novotný, T., Haupt, F. & Belzig, W. Nonequilibrium phonon backaction on the current noise in atomic-sized junctions. *Phys. Rev. B - Condens. Matter Mater. Phys.* **84**, (2011).
  175. Novotný, T. & Belzig, W. Large-voltage behavior of charge transport characteristics in nanosystems with weak electron – vibration coupling. *Belstein J. Nanotechnol.* **6**, 1853–1859 (2015).
  176. D’Agosta, R., Sai, N. & Di Ventra, M. Local Electron Heating in Nanoscale Conductors. *ACS Nano Lett.* **6**, 2935–2938 (2006).
  177. Roy, D., Vignale, G. & Di Ventra, M. Viscous corrections to the resistance of nanojunctions : A dispersion relation approach. *Phys. Rev. B* **83**, 075428 (2011).
  178. Chen, R., Wheeler, P. J. & Natelson, D. Excess noise in STM-style break junctions at room temperature. *Phys. Rev. B* **85**, 235455 (2012).
  179. Armstrong, J. N., Schaub, R. M., Hua, S. Z. & Chopra, H. D. Channel saturation and conductance quantization in single-atom gold constrictions. *Phys. Rev. B* **82**, 195416 (2010).
  180. Bürki, J., Stafford, C. A., Zotos, X. & Baeriswyl, D. Cohesion and conductance of disordered metallic point contacts. *Phys. Rev. B* **50**, 5000 (1999).
  181. Bürki, J. & Stafford, C. A. Comment on ‘Quantum Suppression of Shot Noise in Atom-Size Metallic Contacts’. *Phys. Rev. Lett.* **83**, 3342 (1999).

182. Mitra, A., Aleiner, I. & Millis, A. J. Phonon effects in molecular transistors: Quantal and classical treatment. *Phys. Rev. B* **69**, 245302 (2004).
183. Chen, Y.-C. & Di Ventra, M. Effect of Electron-Phonon Scattering on Shot Noise in Nanoscale Junctions. *Phys. Rev. Lett.* **95**, 166802 (2005).
184. Schmidt, T. L. & Komnik, A. Charge transfer statistics of a molecular quantum dot with a vibrational degree of freedom. *Phys. Rev. B* **80**, 041307 (2009).
185. Avriller, R. & Yeyati, A. L. Electron-phonon interaction and full counting statistics in molecular junctions. *Phys. Rev. B* **80**, 041309 (2009).
186. Haupt, F., Novotný, T. & Belzig, W. Phonon-Assisted Current Noise in Molecular Junctions. *Phys. Rev. Lett.* **103**, 136601 (2009).
187. Urban, D. F., Avriller, R. & Yeyati, A. L. Nonlinear effects of phonon fluctuations on transport through nanoscale junctions. *Phys. Rev. B* **82**, 121414 (2010).
188. Wheeler, P. J., Chen, R. & Natelson, D. Noise in electromigrated nanojunctions. *Phys. Rev. B* **87**, 155411 (2013).
189. Naveh, Y. & Laikhtman, B. Band-structure tailoring by electric field in a weakly coupled electron-hole system. *Appl. Phys. Lett.* **66**, 1980 (1995).
190. Du, L. *et al.* Tuning Edge States in Strained-Layer InAs/GaInSb Quantum Spin Hall Insulators. *Phys. Rev. Lett.* **119**, 056803 (2017).
191. Kroemer, H. The 6.1 A family (InAs, GaSb, AlSb) and its heterostructures : a selective review. *Phys. E* **20**, 196–203 (2004).
192. Sakaki, H. *et al.* In<sub>1-x</sub>GaxAs-GaSb<sub>1-y</sub>As<sub>y</sub> heterojunctions by molecular beam epitaxy. *Appl. Phys. Lett.* **31**, 211 (1977).
193. Nakagawa, A., Kroemer, H. & English, J. H. Electrical properties and band offsets of

- InAs/AlSb n-N isotype heterojunctions grown on GaAs. *Appl. Phys. Lett.* **54**, 1893 (1989).
194. Bhargava, S., Blank, H.-R., Narayanamurti, V. & Kroemer, H. Fermi-level pinning position at the Au – InAs interface determined using ballistic electron emission microscopy. *Appl. Phys. Lett.* **70**, 759 (1997).
  195. Altarelli, M. Electronic structure and semiconductor-semimetal transition in InAs-GaSb superlattices. *Phys. Rev. B* **28**, 842 (1983).
  196. Yang, M. J., Yang, C. H., Bennett, B. R. & Shanabrook, B. V. Evidence of a Hybridization Gap in “Semimetallic” InAs/GaSb Systems. *Phys. Rev. Lett.* **78**, 4613 (1997).
  197. Lakrimi, M. *et al.* Minigaps and Novel Giant Negative Magnetoresistance in InAs/GaSb Semimetallic Superlattices. *Phys. Rev. Lett.* **79**, 3034 (1997).
  198. Naveh, Y. & Laikhtman, B. Tunneling and nonresonant negative differential resistance in narrow-well interband-tunneling devices. *Phys. Rev. B* **49**, 16829 (1994).
  199. Cooper, L. J. *et al.* Resistance resonance induced by electron-hole hybridization in a strongly coupled InAs/GaSb /AlSb heterostructure. *Phys. Rev. B* **57**, 11915 (1998).
  200. Kono, J. *et al.* Far-infrared magneto-optical study of two-dimensional electrons and holes in InAs/Al<sub>x</sub>Ga<sub>1-x</sub>Sb quantum wells. *Phys. Rev. B* **55**, 1617 (1997).
  201. Yang, M. J., Yang, C. H. & Bennett, B. R. Magnetocapacitance and far-infrared photoconductivity in GaSb / InAs composite quantum wells. *Phys. Rev. B* **60**, R13958 (1999).
  202. Naveh, Y. & Laikhtman, B. Magnetotransport of coupled electron-holes. *Europhys. Lett.* **55**, 545 (2001).
  203. Quinn, J. J. & Quinn, J. J. Semimetal-semiconductor transition in InAs-GaSb



- heterostructures. *Surf. Sci.* **361/362**, 930–932 (1996).
204. Laikhtman, B., de Leon, S. & Shvartsman, L. D. Energy Spectrum of InAs/GaSb Heterostructures. *Solid State Commun.* **104**, 257–262 (1997).
205. de Leon, S., Shvartsman, L. D. & Laikhtman, B. Band structure of coupled InAs/GaSb quantum wells. *Phys. Rev. B* **60**, 1861 (1999).
206. Aizin, G. R., Laikhtman, B. & Gumbs, G. Plasmons in coupled electron-hole double quantum wells. *Phys. Rev. B* **64**, 125317 (2001).
207. Döhler, G. H. Electron-hole subbands at the GaSb/InAs interface. *Surf. Sci.* **98**, 108–116 (1980).
208. Knez, I., Du, R. R. & Sullivan, G. Finite conductivity in mesoscopic Hall bars of inverted InAs / GaSb quantum wells. *Phys. Rev. B* **81**, 201301 (R) (2010).
209. Kane, E. O. Band structure of indium antimonide. *J. Phys. Chem. Solids* **1**, 249–261 (1957).
210. Suzuki, K., Harada, Y., Onomitsu, K. & Muraki, K. Gate-controlled semimetal-topological insulator transition in an InAs/GaSb heterostructure. *Phys. Rev. B - Condens. Matter Mater. Phys.* **91**, 245309 (2015).
211. Mueller, S. *et al.* Nonlocal transport via edge states in InAs / GaSb coupled quantum wells. *Phys. Rev. B* **92**, 081303 (2015).
212. Couëdo, F., Irie, H., Suzuki, K., Onomitsu, K. & Muraki, K. Single-edge transport in an InAs/GaSb quantum spin Hall insulator. *Phys. Rev. B* **94**, 035301 (2016).
213. Caldeira, A. O. & Leggett, A. J. Influence of Dissipation on Quantum Tunneling in Macroscopic Systems. *Phys. Rev. Lett.* **46**, 4–7 (1981).
214. Zhou, B., Lu, H.-Z., Chu, R.-L., Shen, S.-Q. & Niu, Q. Finite Size Effects on Helical Edge

- States in a Quantum Spin-Hall System. *Phys. Rev. Lett.* **101**, 246807 (2008).
215. Väyrynen, J. I. & Ojanen, T. Electrical Manipulation and Measurement of Spin Properties of Quantum Spin Hall Edge States. *Phys. Rev. Lett.* **106**, 076803 (2011).
216. Qu, F. *et al.* Electric and Magnetic Tuning Between the Trivial and Topological Phases in InAs / GaSb Double Quantum Wells. *Phys. Rev. Lett.* **115**, 036803 (2015).
217. Abrahams, E., Anderson, P. W., Licciardello, D. C. & Ramakrishnan, T. V. Scaling Theory of Localization: Absence of Quantum Diffusion in Two Dimensions. *Phys. Rev. Lett.* **42**, 673 (1979).
218. König, M. *et al.* The quantum spin Hall effect: Theory and experiment. *J. Phys. Soc. Japan* **77**, 031007 (2008).
219. Fang, F. F. & Stiles, P. J. Quantized magnetoresistance in two-dimensional electron systems. *Phys. Rev. B* **27**, 6487 (1983).
220. Nichele, F. *et al.* Edge transport in the trivial phase of InAs / GaSb. *New J. Phys.* **18**, 083005 (2016).
221. Mitin, V., Reggiani, L. & Varani, L. Generation-Recombination Noise in Semiconductors. in *Noise and Fluctuations Control in Electronic Devices* (ed. Balandin, A.) 2–18 (American Scientific Publishers, 2002).
222. Suris, R. A. & Fuks, B. I. Noise in finite compensated semiconductors. I. Generation-recombination noise. *Sov. Phys. Semicond.* **14**, 641 (1980).
223. Bosman, G. & Zijlstra, R. J. J. Generation-Recombination Noise in p-Type Silicon. *Solid State Electron.* **25**, 273 (1982).
224. van Vleit, K. M. Noise in Semiconductors and Photoconductors. *Proc. Inst. Radio Eng.* **101**, 1004 (1958).

225. Christensen, C. J. & Pearson, G. L. Spontaneous Resistance Fluctuations in Carbon Microphones and Other Granular Resistances. *Bell Syst. Tech. J.* **15**, 197–223 (1936).
226. Hill, J. E. & van Vleit, K. M. AMBIPOLAR TRANSPORT OF CARRIER DENSITY FLUCTUATIONS IN GERMANIUM. *Physica* **24**, 709–720 (1958).
227. van Vleit, K. M. & Fassett, J. Fluctuation Phenomena in Solids. in *Fluctuation Phenomena in Solids* (ed. Burgess, R.) 267 (Academic, 1965).
228. Smith, D. L. & Mailhot, C. Proposal for strained type II superlattice infrared detectors. *J. Appl. Phys.* **62**, 2545 (1987).
229. Stevens, L. A., Li, T., Du, R.-R. & Natelson, D. Noise processes in InAs/Ga(In)Sb Corbino structures. *Appl. Phys. Lett.* **115**, 052107 (2019).
230. Shimizu, A. & Ueda, M. Effects of Dephasing and Dissipation on Quantum Noise in Conductors. *Phys. Rev. Lett.* **69**, 1403–1406 (1992).
231. Cowley, A. M. & Zettler, R. A. Shot Noise in Silicon Schottky Barrier Diodes. *IEEE Trans. Electron Devices* **15**, 761–769 (1968).
232. Frenkel, Y. On Pre-Breakdown Phenomena in Insulators and Electronic Semi-Conductors. *Phys. Rev.* **54**, 647 (1938).
233. Reklaitis, A. & Reggiani, L. Enhanced shot noise from tunneling and space-charge positive feedback. *Phys. Rev. B* **62**, 16773 (2000).
234. Spanton, E. M. *et al.* Images of edge current in InAs/GaSb quantum wells. *Phys. Rev. Lett.* **113**, 1–5 (2014).