Disorder Effects in Charge Transport and Spin Response of Topological Insulators

Lukas Zhonghua Zhao
Graduate Center, City University of New York

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/gc_etds
Part of the Physics Commons

Recommended Citation
https://academicworks.cuny.edu/gc_etds/1200
Disorder Effects in Charge Transport and Spin Response of Topological Insulators

by

Lukas Zhonghua Zhao

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of requirements for the degree of Doctor of Philosophy, The City University of New York.

2015
This manuscript has been read and accepted for the Graduate Faculty in Physics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date
Professor Lia Krusin-Elbaum
Chair of Examining Committee

Date
Professor Igor L. Kuskovsky
Executive Officer

Supervisory Committee:
Professor James C. Hone
Professor Vadim Oganesyan
Doctor Jonathan Sun
Professor Maria C. Tamargo

THE CITY UNIVERSITY OF NEW YORK
Abstract

Disorder Effects in Charge Transport and Spin Response in Topological Insulators

by

Lukas Zhonghua Zhao

Thesis Advisor: Professor Lia Krusin-Elbaum

Topological insulators are a class of solids in which the non-trivial inverted bulk band structure gives rise to metallic surface states that are robust against impurity backscattering. First principle calculations predicted Bi$_2$Te$_3$, Sb$_2$Te$_3$ and Bi$_2$Se$_3$ to be three-dimensional (3D) topological insulators with a single Dirac cone on the surface. The topological surface states were subsequently observed by angle-resolved photoemission (ARPES) and scanning tunneling microscopy (STM). The investigations of charge transport through topological surfaces of 3D topological insulators, however, have faced a major challenge due to large charge carrier densities in the bulk donated by randomly distributed defects such as vacancies and antisites. This bulk disorder intermixes surface and bulk conduction channels, thereby complicating access to the low-energy (Dirac point) charge transport or magnetic response and resulting in the relatively low measured carrier mobilities. Moreover, charge inhomogeneity arising from bulk disorder can result in pronounced
nanoscale spatial fluctuations of energy on the surface, leading to the formation of surface ‘puddles’ of different carrier types. Great efforts have been made to combat the undesirable effects of disorder in 3D topological insulators and to reduce bulk carriers through chemical doping, nanostructure fabrication, and electric gating. In this work we have developed a new way to reduce bulk carrier densities using high-energy electron irradiation, thereby allowing us access to the topological surface quantum channels. We also found that disorder in 3D topological insulators can be beneficial. It can play an important part in enabling detection of unusual magnetic response from Dirac fermions and in uncovering new excitations, namely surface superconductivity in Dirac ‘puddles’.

In Chapter 3 we show how by using differential magnetometry we could probe spin rotation in the 3D topological material family (Bi$_2$Se$_3$, Bi$_2$Te$_3$ and Sb$_2$Te$_3$), and describe our detection of paramagnetic singularity in the magnetic susceptibility at low magnetic fields that persists up to room temperature, and which we have demonstrated to arise from the surfaces of the samples. The singularity is universal to the entire family, largely independent of the bulk carrier density, and consistent with the existence of electronic states near the spin-degenerate Dirac point of the 2D helical metal. The exceptional thermal stability of the signal points to an intrinsic surface cooling process, probably of thermoelectric organ, and establishes a sustainable platform for the singular field-tunable Dirac spin response.

In Chapter 4 we describe our discovery of surface superconductivity in a hole-conducting topological insulator Sb$_2$Te$_3$ with transition to zero resistance induced through a minor tuning of growth chemistry that depletes bulk conduction channels. The depletion shifts Fermi energy towards the Dirac point as witnessed by over two orders of magnitude
ABSTRACT

reduced bulk hole density and by the largest carrier mobility (∼ 25,000 cm²V⁻¹s⁻¹) found in any topological material. Direct evidence from transport, the unprecedentedly large diamagnetic screening, and the presence of up to ∼ 25 meV gaps in differential conductance detected by scanning tunneling spectroscopy (STM) reveal the superconducting condensate to emerge first in surface puddles at unexpectedly high temperature, near 50 K. Percolative Josephson paths mediated by diffusing quasiparticles establish global phase coherence around 9 K. Rich structure of this state lends itself to manipulation and tuning via growth conditions and the topological material’s parameters such as Fermi velocity and mean free path.

In Chapter 5 we describe a new approach we have developed to reaching stable charge neutrality in 3D topological materials. The technique uses swift (∼ 2.5 MeV energy) electron beams to compensate charged bulk defects and bring the Fermi level back into the bulk gap. By controlling the beam fluence we could tune bulk conductivity from p- (hole-like) to n-type (electron-like), crossing the Dirac point and back, while preserving the robust topological signatures of surface channels. We establish that at charge neutrality conductance has a two-dimensional (2D) character with a minimum value on the order of ten conductance quanta \( G_0 = e^2/h \). From quantum interference contribution to 2D conductance we demonstrate in two systems, Bi₂Te₃ and Bi₂Se₃, that at charge neutrality only two quantum channels corresponding to two topological surfaces are present. The charge neutrality point achieved using electron irradiation with long penetration range shows a route to intrinsic quantum transport of the topological states unconstrained by the bulk size.
In loving memory of my grandparents Guangjun Zhao and Yuanqing Quan.
Acknowledgments

I want to thank my mentor Professor Lia Krusin-Elbaum for all her teaching and guidance. Her knowledge, expertise, and perseverance made it possible for me to finish my Ph.D. studies and this thesis. Lia has become a role model scientist for me, and a dear friend. Her caring and love have given me enormous support and joy in the course of my graduate studies.

I would like to thank Professor Vadim Oganesyan for his consistent theoretical support and advice.

Haiming Deng has been a great partner in the lab. His talent and hard work have been an important part in the work we have done. Doctor Milan Begliarbekov has given me good advice and practical assistance. I thank Zhiyi Chen, Inna Korzhovska and Jeff Secor for being great lab partners and all their help. I am grateful to Doctor Alexey Bykov for his technical support.

I would also like to thank my thesis supervisory committee members Professor Lia Krusin-Elbaum, Professor Vadim Oganesyan, Doctor Jonathan Sun, Professor Maria Tamargo and Professor James Hone for their advice and thank Doctor Daniel Moy for his help and guidance through the years of my graduate work.

Last but not the least, I want to thank my parents Baoli Wang and Hongquan Zhao, and my brother Boying Zhao for their love and support. I thank Justin Cruz for his love and for always standing by me.
Contents

Abstract ...................................................................................................................................................... iv
Acknowledgments .......................................................................................................................................... viii
Table of Figures ........................................................................................................................................ xii
Chapter 1 Introduction to topological insulators ..................................................................................... 1
  1.1 Classification of matter ....................................................................................................................... 1
  1.2 Topology ......................................................................................................................................... 2
  1.3 Relating topology to band theory .................................................................................................... 3
  1.4 Three-dimensional topological insulators and the topological surface states ..................................... 6
  1.5 Disorders in three-dimensional topological insulators ...................................................................... 8
  1.6 Weak localization and weak anti-localization .................................................................................... 10
  1.7 Superconductivity in topological insulators ..................................................................................... 13
Chapter 2 Experimental methodologies ................................................................................................ 14
  2.1 Bridgman-Stockbarger growth of single crystalline topological insulators ...................................... 14
  2.2 Fabricating samples of topological insulator crystals for transport measurement ................................ 17
  2.3 Electrical transport measurement instrumentation setups ............................................................... 19
  2.4 AC susceptibility measurement system ............................................................................................ 20
Chapter 3 Singular robust Dirac spin response ....................................................................................... 22
  3.1 Introduction ..................................................................................................................................... 22
  3.2 Spin response from Dirac fermions .................................................................................................. 24
CONTENTS

3.3 Landé $g$-factor from Shubnikov de Haas oscillations .............................................. 37
3.4 The low frequency limit of the AC magnetic susceptibility ......................................... 37
3.5 Phenomenological theory .................................................................................................. 40
3.6 Efficient cooling of Dirac fermions ................................................................................. 44

Chapter 4 Surface superconductivity of Dirac puddles .......................................................... 49

4.1 Introduction ....................................................................................................................... 49
4.2 Experimental results ......................................................................................................... 50
4.3 Discussion ........................................................................................................................... 59
4.4 Structural and elemental analysis ..................................................................................... 62
4.5 Te pressure during crystal growth and Te doping ............................................................. 65
4.6 Superconducting gap mapping using scanning tunneling spectroscopy ......................... 67
4.7 Determination of carrier density from Hall resistivity and dHvA oscillations .............. 69
4.8 Lifshitz-Kosevich analysis and determination of spin-orbit splitting ............................ 70
4.9 Estimating interpuddle separation .................................................................................... 73
4.10 Frequency and temperature dependence in the inductive linear response ................... 75
4.11 Resistive signature at 50 K consistent with superconducting puddles ...................... 78

Chapter 5 Using electron irradiation to achieve surface transport in topological insulators ........................................................................................................................................ 79

5.1 Introduction ....................................................................................................................... 79
5.2 Achieving surface quantum transport using high energy electron irradiation ............. 80
5.3 Penetration range of electron beams in Bi$_2$Te$_3$ and Bi$_2$Se$_3$ ........................................ 91
5.4 Transmission Electron Microscopy of Bi$_2$Te$_3$ exposed to 2.5 MeV e-beams ............ 93
5.5 Tuning through charge neutrality point (CNP) in Bi$_2$Se$_3$:Ca .................................... 94
5.6 Material’s parameters from longitudinal and Hall resistivities and from SdH...... 96
5.7 Isochronal annealing experiments and energy barriers for defect migration........ 96
5.8 Variable range hopping bulk charge transport at low carrier density................. 98
5.9 Stable charge neutrality point in Bi2Te3 achieved through annealing............... 99

Chapter 6 Open questions .................................................................................. 101

List of publications – Lukas Zhonghua Zhao ...................................................... 103

Bibliography ........................................................................................................ 105
Table of Figures

Figure 1-1 The surfaces of a sphere and a bowl are topologically identical (g = 0). The surface of a doughnut and a coffee mug are also topologically identical (g = 1). The former two and the latter two are distinguished topologically by their genus g.

Figure 1-2 Berry phase in a two band theory is equal to half the solid angle swept out by the unit vector of \( \mathbf{d} \).

Figure 1-3 a, Crystal structure of Bi\(_2\)Te\(_3\). A quintuple layer (a stack of Te1-Bi-Te2-Bi-Te1) is indicated by the red box. b, Angle-resolved photoemission spectroscopy (ARPES) maps the energy states in momentum space. Spin-resolved ARPES reveals that the spins (red) of the surface states lie in the surface plane and are perpendicular to the momentum, as shown in Xia et al., Nature Physics 5, 398 (2009) and Hsieh et al., Science 323, 919 (2009). c. ARPES plot of energy versus wavenumber in Bi\(_2\)Te\(_3\) shows the linearly dispersing surface-state band (SSB) above the bulk valence band (BVB). The dashed white line indicates the Fermi level. The blue lines meet at the tip of the Dirac cone, as published in Science 325, 178 (2009).

Figure 1-4 Shou-Cheng Zhang and Zhong Fu’s group showed calculation of energy and momentum dependence of the local density of states (LDOS) for Sb\(_2\)Te\(_3\), Bi\(_2\)Te\(_3\) and Bi\(_2\)Se\(_3\), and predicted these systems are three-dimensional topological insulators, published in Nature Physics, 5, 438 (2009).
Figure 1-5 a, Electron and hole ‘puddles’ formation on the surface of topological insulator with bulk disorders, observed via scanning tunneling microscopy (STM). b, High energy resolution local density of states (LDOS) taken along a line cut that crosses several charge puddles (white line) on topological insulator with disorders in the bulk, demonstrating the absence of resonances between the puddles and the dopants on the surface. c, Line cut along the yellow line in a showing electron and hole ‘puddles’ on the surface. These three figures are from Beidenkopf et al., Nature Physics 7, 939 (2011).

Figure 1-6 Weak localization is due to self-intersecting paths of electron scattering in disordered metals or semi-metals. When spin-orbit coupling is present, the clock-wise and counter-clock-wise loop interfere destructively, leading to weak anti-localization.

Figure 2-1 Sketches of (a) horizontal Bridgman growth and (b) vertical Bridgman-Stockbarger growth of single crystalline topological insulators.

Figure 2-2 (a) an eight-pin chip carrier carrying a van der Pauw sample. (b) An exfoliated 3D topological insulator flake with lithographical van der Pauw contacts.

Figure 3-1 a, The energy-momentum relation of the surface states in a 3D topological insulator. b, Schematic for the measurement of magnetic susceptibility.

Figure 3-2 Magnetic susceptibility of Bi$_2$Se$_3$ measured by applying a small AC excitation field $H_{ac}$ shows that the spin response is cusp-like and large near zero applied DC magnetic field.

Figure 3-3 a, Diamagnetic susceptibility of the sample holder used in the experiments is two orders of magnitude smaller than the typical signal from the sample signal. Susceptibility of precursor materials used in the crystal growth, Te and Sb, is depicted in panels b, d and c, respectively. d, Susceptibility of Te after annealing at 450 ºC in the
growth furnace is also featureless. The well-known large diamagnetism of Sb is enhanced at low temperatures (similar to Bi) while the diamagnetism of Te is only weakly temperature dependent. Both are in good agreement with the literature values. The signal from the paramagnetic Pd calibration sample is shown in e. Susceptibility of the layered 2H– NbSe$_2$ shows a well-known behavior in the normal and superconducting states$^{57}$ and no paramagnetic cusp, see f. These essentially featureless back-ground/calibration checks are in contrast with the cusplike paramagnetic field dependence at low fields consistently observed in the crystals of Sb$_2$Te$_3$, Bi$_2$Te$_3$, and Bi$_2$Se$_3$.

Figure 3-4 Universality of singular spin response near zero magnetic field. a-c, The zero-field susceptibility cusp is found in all three topological insulators. d-f, The susceptibility surface in the H-T phase space for fields above H~0.5T is shown for Sb$_2$Te$_3$ (d), Bi$_2$Te$_3$ (e) and Bi$_2$Se$_3$ (f). The most pronounced temperature dependence is found in Sb$_2$Te$_3$ (d) which has the smallest bulk bandgap of ~100 meV. g-i, Corresponding schematic band structures (insets) indicate noticeable differences in the location of the Dirac point relative to the bulk valence and conduction bands. Measurements of Hall resistivity (blue lines in g-i) show that the Te-based TIs, Sb$_2$Te$_3$ (g) and Bi$_2$Te$_3$ (h), are intrinsically p-type, whereas the Se-base TI, Bi$_2$Se$_3$ (i), is n-type.

Figure 3-5 Susceptibility cusp for a Bi$_2$Se$_3$ crystal at several temperatures measured (a) an hour after the crystal growth and (b) two weeks later after the crystal was stored in flowing nitrogen. While the temperature robustness is intact, the overall cusp height has been reduced with time likely due to surface reconstruction and the formation of two-dimensional electron gas (2DEG) associated with the band bending of the bulk states at the surfaces$^{54,55}$. 


Figure 3-6 a, Shubnikov-de Haas (SdH) oscillations for a Bi2Te3 crystal with carrier density $n \sim 10^{17}$/cc; b, The oscillations are fitted to Lifshitz-Kosevich formula using Monte Carlo technique. Below 8 Tesla field, the fit gives cyclotron mass $m_c = 0.0767m_e$, where $m_e$ is bare electron mass. The obtained value of the g-factor is $g \sim 2m_e/m_c \approx 30$, of the order of g-factors $\sim 60$ reported in other experiments.

Figure 3-7 Signature of the surface origin of the cusp. a, Susceptibility cusp for two Bi2Te3 crystals with carrier densities differing by two orders of magnitude. The slope of the cusp is independent of the bulk carrier density n. Here the diamagnetic background was subtracted and the height of the cusp was normalized to (B=0), which for the $n \sim 10^{19}$ cm$^{-3}$ crystal was $3 \times 10^{-5}$ emu/cc, and for the $n \sim 10^{17}$ cm$^{-3}$ crystal was $3.5 \times 10^{-5}$ emu/cc. b, Left: The susceptibility cusp before and after cutting the crystal thickness by factors of 0.63, 0.29 and 0.15 seems to be independent of the thickness t. The diamagnetic background scales with thickness. Right: The data for all thicknesses shown on the left shifted to match the diamagnetic background.

Figure 3-8 a, The simple Dirac model of equation (3-1) produces a very good match to the data, as illustrated for the case of Sb2Te3. Here $\chi = x\chi_A/L_z$ and $\chi_A$ is the 2D susceptibility of the Dirac state, $L_z \approx 10^{-3}$ m, thickness of our samples, and $x < 1$ the effective areal fraction occupied by the ungapped Dirac state ($x$ is used as a fitting parameter). Other parameter values used to generate this plot are $\mu = k_BT= 0$, $g = 60$, $v_F = 2 \times 10^3$ m s$^{-1}$, which are known from our own studies (Supplementary Information) and those of others. Both $x$ and $\Lambda$ (effective radius of k-space contributing to singular response) were adjusted to match the data, producing $x \approx 0.002$ and $\Lambda = 5 \times 10^8$ m$^{-1}$. The cusp is preserved even when hexagonal warping (inset in a) is taken into account—it is merely subsumed into $\Lambda$. 
b. Rare regions of chemical potential $\mu \approx 0$ (grey) can exist between electron (blue) and hole (yellow) droplets owing in part to the electrostatic potential established by the charged defects in the bulk. Such fluctuations of the local surface charge are probably ‘healing’ in the course of the ageing process as the mean chemical potential steadily floats away from the Dirac point towards bulk conduction or valence bands, as has been documented in ARPES studies. This is qualitatively consistent with the observed decrease in the amplitude of the paramagnetic anomaly over time.

Figure 3-9 a, The in-phase component of the susceptibility containing the singular cusp is independent of frequency (shown here for Sb$_2$Te$_3$). However, the diamagnetic susceptibility is slightly dependent on frequency. b, The nonlinearity of the surface–bulk connection is witnessed by the observed second harmonic of $\chi$. It is consistent with the existence of ‘rectifying’ paths in the putative thermoelectric cooling elements required for the cooling of a small fraction of the sample’s surface and thus suppressing thermalization of Dirac surfaces with the bulk, as explained in the text. The effective cooling of the surface is naturally achieved by the electron and hole puddles in the subsurface region, forming a Peltier element (inset) owing its cooling efficiency partly to nanoconstriction and partly to frequency-dependent transport coefficients.

Figure 3-10 DC magnetization M measured in the Superconducting Quantum Interference Device (SQUID) magnetometer shows clear nonlinearity near zero field. Numerical derivative $dM/dH$ of M gives a ‘cusp’, as shown in the inset. Consistent with our model in Section II the dc cusp singularity is rounded and diminished with decreasing temperature. We note that taking numerical derivatives of magnetization in the vicinity of zero DC magnetic field expectedly produces a spurious numerical noise. Hence, not surprisingly,
data taken at a finite frequency, where the derivative is taking ‘in situ’ using a small AC oscillation gives a much more accurate record of magnetic susceptibility $\chi = \frac{dM}{dH}$ near $H \sim 0$. The derivative noise is partly controlled by the lock-in amplifier in the ac detection circuit and it decreases at higher excitation frequencies.

Figure 3-11 (a) Frequency dependence of the background and (b) independence of the singular contributions.

Figure 3-12 (a) The out-of-phase susceptibility component of $\chi(\omega)$ is routinely recorded simultaneously with the in-phase component (here in Sb2Te3). It is purely dissipative and regular in the vicinity of $H = 0$; it does not display the cuspy behavior. (b) Assuming the standard eddy current mechanism is responsible for dissipation, the out-of-phase component of $\chi(\omega)$ is proportional to conductivity, which is confirmed in our measurements: the observed value is consistent, up to geometric factors and closely follows the temperature dependence in-plane resistivity $\rho_{xx}$, according to the standard formula for power $P = \pi^2(\hbar_0^2d^2f^2/2\rho_{xx}(T))$ dissipated during the AC excitation cycles. Here $f = \omega/2\pi$ and $d$ is the sample thickness.

Figure 3-13 Temperature dependence predicted in Equation (3-6) is explored here by plotting traces at $T = 0$ (black), $T = 1\text{K}$ (red) and $T = 10\text{K}$ (green and red, see below). We compare the first plot to choose experimental parameters, i.e. $g = 60$, $\Lambda = 5 \cdot 108^{-1}$, $v_F = 2000 \text{ m/s}$, $x = 0.18 \cdot 1^{-3}$, also $L_z \approx 1 \text{ mm}$ is the measured thickness of these samples. These parameters are sample dependent—they can vary with sample growth and preparation, as well as with the level of surface reconstruction and/or ‘aging’. The value of $\Lambda$, the characteristic size of the momentum space of the surface states, is a few percent of the corresponding bulk quantity, its microscopic meaning remains to be established.
Here, we pick it to be some fraction of the bulk Brillouin zone. Lastly, the Fermi velocity we use is consistent with that obtained from a low energy probe\textsuperscript{60}. If these parameters are rescaled upon raising temperature (from 1K to 10K), $\Lambda \rightarrow \Lambda$, $v_F \rightarrow 10v_F$, $x \rightarrow x/10$, $g \rightarrow 10g$, $\chi$ is invariant (red trace). Otherwise, in the green trace, we explore whether approximate invariance may be maintained under less fine-tuned rescaling, e.g. with $\Lambda \rightarrow \Lambda/2$, $v_F \rightarrow 2v_F$, $x \rightarrow 1.5x$, $g \rightarrow 3g$. This discussion is not intended as analysis of the experiments, but rather to explore potential thermal effects theoretically................. 41

Figure 3-14 The basic idea is that a local fluctuation in doping, e.g. of the kind that may create a $\mu \sim 0$ patch at the surface (see above), will also create adjacent p and n doped regions. With the ac excitation field oriented along the c-axis we expect induced currents primarily flowing in the ab plane. For the experimental parameters, such as material’s conductivity and probe’s frequency range, the dissipative out-of-phase (eddy current) component dominates the in-phase component by about 3 orders of magnitude. These finite frequency currents are expected to exhibit inhomogeneities induced by variations in electronic structure. In particular, it is natural to find current loops traversing through the bulk, the surface and the said p-n regions. Under favorable conditions such current loops will act as mesoscopic Peltier coolers. These favorable conditions include relatively low local resistance (and smooth disorder) along the current path that helps focus the current flow through thermoelectrically asymmetric region and also direct contact between p and n regions that result in formation of a depletion layer and rectification path (diode shunt in the figure).................................................................................................................. 45

Figure 3-15 The nonlinearity of the surface-bulk connection is witnessed by the observed 2nd harmonic $\chi_2$. (b) As expected, $\chi_2$ is quadratic in frequency. It is consistent with the
existence of ‘rectifying’ paths in the putative thermoelectric cooling elements required for the cooling of small fraction of sample’s surface and thus suppressing thermalization of Dirac surfaces with the bulk. The generation of 2nd harmonic (c) is associated with (d) the rectified signal. 47

Figure 4-1 a, Resistivity of an exfoliated 100 nm thin Sb$_2$Te$_3$ crystal synthesized under ~ 1.4 MPa Te vapor pressure shows onset of transition to zero resistance at $T_{CR}$ $\approx$ 8.6 K. b, Diamagnetic susceptibility (left) at 1.9 K measured in a 0.2 T field shows huge diamagnetism (Meissner effect) only in the narrow Te vapor pressure range. It is anticorrelated with the measured hole density (right) which in the superconducting state is decreased by a factor of nearly 300. Outside this range samples are nonsuperconducting. c, Longitudinal sheet resistance $R_{xx}$ (red) showing a discernible downstep on decreasing temperature. This downstep is localized within ~ 5 K in temperature, as indicated by the derivative d$R_{xx}$/dT (blue). d, Illustration of superconductive puddles (blue, size 2R) of Dirac bands, where pairing occurs at high temperature ($T_{CD}$), connected to the 2DEG metallic matrix (grey) which establishes a percolative path (dashed red line) at $T_{CR}$. The electronic Dirac dispersion and spin-orbit split dispersion of 2DEG are also sketched, see text. d,e,f, STM and STS scans were performed at 4.8 K in areas ~ 25 nm each, spaced ~ 200 nm apart. e, A typical topograph of a scanned area shows well-ordered hexagonal lattice, with differential conductance dI/dV (from the average of 500 scans) shown in f. Depending on the scan area, the gaps $2\Delta$ vary from zero to $\&$ 20 meV. 51

Figure 4-2 a, Resistive transition temperature downshifts with increasing magnetic field applied (top) along the c-axis of the crystal, $H \parallel c$, and (bottom) parallel to the ab-plane, $H \parallel ab$. The onset of superconductivity is indicated by the arrows. b, H - T phase diagram
of the superconducting state has relatively small zero-temperature anisotropy of ~ 1.3, the same as that of the square root of g-factor. The critical field values at $T \to 0$ agree with the paramagnetic (Zeeman) depairing field $H_p$. Inset: de Haas van Alfen quantum oscillations (dHvA) shown for $H \parallel c$ at 1.9 K persist below $H_p$ down to the temperature dependent onset field $H^*$. Figure 4-3 a, dHvA oscillations show beats arising from two very close oscillation periods. The beats are a signature of spin-splitting by a strong spin-orbit interaction in 2DEG surface regions. b, dHvA oscillations vs. inverse transverse component of magnetic field $H_\perp = H \cos \theta$ shown for three values of magnetic tilt angle $\theta$. The observed constancy of beat nodes scaling with $H_\perp$ is a characteristic signature of 2D. c, Hole carrier density $n$ vs. Te pressure in the superconducting Sb$_2$Te$_3$ (in the narrow $\Delta P$ vicinity of ~ 1.4 MPa of Te pressure) and non-superconducting Sb$_2$Te$_3$ (outside $\Delta P$ ) states. In the superconducting region $n$ is reduced, bringing the chemical potential from the deep inside the bulk valence band to just below the Dirac point, as illustrated in sketches of the band structures in all three regions. The conductivity of the superconducting Sb$_2$Te$_3$ remains hole-like (p-type), with Fermi level crossing both, the Dirac bands and a 2DEG sliver of the valence band (upper middle sketch). Figure 4-4 a, Differential magnetic susceptibility vs. temperature for several values of magnetic field $\mu_0 \parallel c$-axis. Meissner-like signal is ‘flat’ below $T_{CR}$ and monotonically vanishes at much higher (> 100 K) temperatures. $\chi$ was measured by applying a small ac excitation field $h_{ac} = 10^{-5}$T at $f = 10$kHz. b, Full temperature and field dependence of $\chi$. c, $\chi$ vs. magnetic field $\mu_0 \parallel c$-axis for a series of temperatures. The pronounced dHvA oscillations are apparent at 1.9 K. d, strongly depends on frequency. e, Full frequency and
temperature dependence of \( \chi \). Inset in g: Analysis of the frequency dependence of \( \chi \) shows it to be quadratic in \( \omega = 2\pi f \). The prefactor \( b(T) \) in the \( \omega^2 \) term dominates the total variation of \( \chi \) at finite frequencies. This variation is consistent with kinetic inductance of the patchy distributed 2DEG network discussed in the text. g, Main panel: The zero frequency \( \chi_0(T) \) obtained from the fits to \( \chi(T) = \chi_0(T) + b(T)\omega^2 \) shows a sharp onset at \( T_{CD} \). h, dc magnetization \( M(T) \) measured at 0.5 Oe using SQUID shows a sharp diamagnetic onset at the same \( T_{CD} \approx 50 \text{ K} > T_{CR} \). The data taken under zero-field-cooled (ZFC) and field-cooled (FC) conditions are essentially identical, consistent with negligible vortex pinning in 2D. We note that \( M(T) \propto T_{CD} - T \) as well the onset of \( \chi_0 \) variation near \( T_{CD} \) shown in (g) are both approximately linear in temperature, which within Ginzburg-Landau phenomenology (valid at high temperatures) would be related to magnetic penetration depth \( \lambda, M(T) \propto 1/\lambda^2(T) \).

Figure 4-5 (a) X-ray diffraction spectrum for Sb\(_2\)Te\(_3\) grown in a sealed quartz tube under Te vapor pressure \( P = 1.42 \text{ MPa} \) collected in Panalytical diffractometer using \( \text{Cu K\(\alpha \) (}\lambda = 1.5405\text{\AA}\text{)} \) line from Philips high intensity ceramic sealed tube (3 kW) X-ray source with a Soller slit (0.04 rad) incident and diffracted beam optics. Rietveld refinement lines shown in red are in full correspondence with the measured spectra. The obtained lattice parameters \( a = b = 4.26 \text{\AA} \) and \( c = 30.46 \text{\AA} \) remain unchanged in the 0.5 - 2 MPa pressure range, indicating that in this pressure range there is no structural change. (b) The robustness of the structure is also apparent in the identical micro-Raman spectra for superconducting (red) a non-superconducting (black) Sb\(_2\)Te\(_3\). The E phonon modes are doubly degenerate modes in the ab-plane and the A modes are nondegenerate vibrations with atomic motion along the c-axis. The spectra were taken in ambient conditions in a backscattering geometry with
linearly polarized excitation in the ab plan and normalized to the out of plane vibration at 70 cm\(^{-1}\).

Figure 4-6 Scanning tunneling spectroscopy obtained on cleaved surfaces of superconducting Sb\(_2\)Te\(_3\). (a) Typical STM topography (sample 1) showing atomically flat terraces of individual Sb\(_2\)Te\(_3\) layers. Topographs of different areas (sample 2) taken approximately 200 nm away from each other display the triangular lattice of tellurium atoms on the surface, and sub-surface defects and dopants. While the topography in (c) and (d) is indistinguishable, the area in (c) displays the superconducting gap while the area in (d) does not. Differential conductance dI/dV (V) was obtained in 25 \(\times\) 25 nm\(^2\) areas of sample 2, at the locations where topographs were obtained. Spectra were obtained in 500 locations in each area. Individual spectra are shown in grey in each panel, and the average spectrum of each area is shown in red. (e) Spectra showing well-articulated coherence peaks and the gap \(2\Delta \approx 10\) meV corresponding to \(T_C \sim 30\) K estimated from BCS gap equation. (f) Spectra showing a more complex behavior with two gaps: a smaller one \(2\Delta \approx 8\) meV and a larger one \(2\Delta \approx 20\) meV corresponding to \(T_C \sim 70\) K. (g) Spectra showing gapless metallic behavior.

Figure 4-7 (a) Main panel: Hall resistivity of non-superconducting Sb\(_2\)Te\(_3\) shows it to be p-type. Carrier density determined from Hall is \(n = 3.9 \times 10^{20}/\text{cc}\). However, as determined from dHvA \(n = 2.56 \times 10^{19}/\text{cc}\), is over an order of magnitude lower. Similar differences are found in other TIs when chemical potential is located deeply inside either valence or conduction bands. This is reconciled when taking into account the incoherent addition of contributions to Hall conductivity from hexagonal “pockets” (factor of 6), while the Fermi cross-sections in dHvA are sampled coherently. The additional factor of 2 comes from
spin-splitting in the bulk bands, fully accounting for the differences in n. (b) Main panel: Hall resistivity of superconducting Sb$_2$Te$_3$ shows it also to be p-type. Here, however, carrier densities determined from Hall and dHvA (top outset) are identical $n = 1.4 \times 10^{18}$/cc, and over an order of magnitude lower than in the non-superconducting samples. This is consistent with the location of the Fermi level just below the Dirac point, and nearly on top of the valence band (Figure 4-3c).

Figure 4-8 Comparison of de Haas van Alfen (dHvA) oscillation amplitude damping for V$_3$Si and Sb$_2$Te$_3$. Plot shows field dependence of $D = \ln[\alpha \sinh(X)B^{1/2}T^{-1}]$ (known as Dingle plot) that shows the change in the dHvA oscillation amplitude upon crossing the superconducting limiting field. Here $\alpha(T, B)$ is the amplitude of quantum oscillations, $X = 2\pi^2k_B T/\hbar \omega_c$ and $\omega_c = eB/m_c$ is the cyclotron frequency. The estimated error of D is 10%.

The data shown for V$_3$Si are from Ref $^{109}$ where the field scale was normalized to upper critical field $B_{c2}$. In conventional superconductor V$_3$Si there is an additional attenuation of the oscillation amplitude at the transition into the superconducting state observed in many extreme type II superconductors$^{110}$. Here we shifted the Dingle scale (in the same units) to overlay the data for both systems in their normal states (red line). The observed modulation of the Dingle factor D in Sb$_2$Te$_3$ is a result of beats in dHvA.

Figure 4-9 (a) Determination of spin-splitting due to s spin-orbit coupling from (b) the beats in de Haas van Alfen quantum oscillations in superconducting Sb$_2$Te$_3$. (c) Determination of cyclotron mass from the temperature dependence of dHvA oscillations in (b). (d) The dHvA beats in two different Sb$_2$Te$_3$. While some oscillation amplitude variations are present, the beat pattern is robust.
Figure 4-10 (a) The in-phase component of $\chi$ is quadratic in frequency $\chi(T) = \chi_0(T) + b(T)\omega^2$. (b) The dissipative (out-of-phase) component of $\chi(\omega)$ is strictly frequency linear, as expected. (c) The standard eddy current mechanism is responsible for dissipation dominated by the bulk, the out-of-phase component $\chi''$ is proportional to conductivity: the observed value is consistent, up to geometric factors and closely follows the temperature dependence in-plane resistivity $\rho_{xx}$ of the bulk, according to the standard formula for power $P = \pi^2(h_{ac}^2d^2f^2/2\rho_{xx}(T))$ dissipated during the AC excitation cycles. Here $f = \omega/2\pi$ and $d$ is the sample thickness. ................................................................. 75

Figure 4-11 (a) A sharp diamagnetic transition at $\sim 50K$ in the zero frequency response $\chi_0(T)$ obtained from the fits to $\chi(T) = \chi_0(T) + b(T)\omega^2$ of another Sb$_2$Te$_3$ crystal. (b) The prefactor $b(T)$ in the $\omega^2$ term is varying smoothly, dominating the total variation of $\chi$ at finite frequencies. This variation is consistent with kinetic inductance of the patchy distributed 2DEG network. .............................................................................................. 76

Figure 4-12 Longitudinal sheet resistance $R_{xx}$ (red) on decreasing temperature in (a) 0 T field, (b) 5 T, and (c) 14 T fields. At zero field this downstep is localized within $\sim 5$ K in temperature, as indicated by the derivative $dR_{xx}/dT$ (blue). Consistent with superconductivity, this downstep is smaller and broader at 5 T and is not detected at 14 T. ....................................................................................................................... 77

Figure 5-1 a, Energetic electron beams can penetrate solids to a depth of many tens of microns. Electron irradiation affects the bulk but not the robust topological surfaces. b, Impinging electrons induce formation of Frenkel vacancy-interstitial pairs (inset), which act to compensate the intrinsic bulk defects. Main panel: Calculated crosssections $\sigma$ for Frenkel pair production in Bi, Te and Se sublattices as a function of electron energy E,
assuming displacement energy ~ 25 eV. Energy thresholds, \( E_{th} \), in \( \sigma \) are set by atomic weight; choosing \( E < E_{th}^{Bi} \) or \( E > E_{th}^{Bi} \) allows to tune Fermi level in both p- and n-type TIs.

c, Transmission electron microscopy image of Bi\(_2\)Te\(_3\) with electron dose \( \phi = 1 \) C/cm\(^2\); the atomic displacements of ~ 1 per 5,000 are not seen.

Figure 5-2 Resistivity of p-type Bi\(_2\)Te\(_3\) irradiated with 2.5 MeV electrons vs. dose \( \phi \) (red squares) measured in situ at 20 K shows about three orders of magnitude increase at the charge neutrality point (CNP) where the conduction is converted from p- to n-type, moving the Fermi level \( E_F \) across the Dirac point (see cartoon). Cycling to room temperature reverses the process, which can be recovered by further irradiation (blue circles) and stabilized.

Figure 5-3 a, Bi\(_2\)Te\(_3\) crystals are irradiated to different terminal doses and measured at 4.2 K. b, Hall resistance. c,d,e, The SdH oscillations in longitudinal resistance and Hall resistance. f,g,h, low-field \( R_{xy} \) and the Fermi level for the samples in c, d, and e. i, The Landau level index plot vs. filed minima in the SdH oscillations. j, A cartoon of a TI with conducting bulk before irradiation (top) and after, with ideally only topological surface conducting (bottom).

Figure 5-4 a, Evolution of longitudinal resistivity \( \rho_{xx} \) measured at 4.2 K after cycling to room temperature (RT) for a crystal irradiated with dose \( \phi = 90 \) mC/cm\(^2\). Each RT dwell time is coded with a different color. Resistivity is seen to cross charge neutrality point (CNP) in reverse from n-type back to p-type. It is consistent with slow migration (hundreds of hours at RT) of vacancies (in accord with ~ 0.8 eV migration barriers, see text) and shows that CNP can be reached by designing a suitable thermal protocol. Insets show \( \rho_{xx}(T) \) for n-type region (upper left) and p-type region (lower right). b, Change in
magnetoconductance (MC) at different RT dwell times; here MC evolves from a quadratic field dependence of a typical bulk metal at short RT dwell times, through a complex region dominated by the charge-inhomogeneous bulk, to a weak antilocalization (WAL) region showing the characteristic low-field cusp near CNP. Here the data were normalized to the value at zero field. A fit to 2D localization theory is shown as red line.  

Figure 5-5 a, Sheet resistance $R_{xx}$ vs. temperature of Bi$_2$Te$_3$ crystal at the charge neutrality point (red line) exhibits a plateau at low temperatures - a thumbprint of 2D surface conduction, see right panels in (c) and (d). Magnetic field breaks time reversal symmetry and gaps Dirac bands, resulting in localizing behavior shown as dash. Inset: Optical image of the crystal showing van der Pauw contact configuration used.  

b, Annealing protocol with time steps $\Delta t = 30$ min implemented to tune Bi2Te3 crystal with dose $1 \text{ C/cm}^2$ back to stable CNP. Inset: Magnetoresistance at 1.9 K after each annealing step, with colors matched to indicate different annealing temperatures.  

c, Left: WAL low-field quantum interference correction to the linear-in-field magnetotransport (in the right panel) at CNP in a Bi$_2$Te$_3$ crystal at 1.9 K with its characteristic low-field cusp. The 2D character of WAL is evident in its scaling with transverse field $H_\perp= H \cos \theta$, where $\theta$ is the tilt angle of the field measured from sample's c-axis. A t to 2D localization (HLN) theory (solid line) confirms that the contribution is only from two surfaces and yields a dephasing $B_\Phi= \frac{\hbar}{(4e \ell_0^2)} \sim 0.01 \text{ T}$. Right: Linear magnetoresistance at CNP shows 2D scaling with $H_\perp$.  

d, Left: WAL contribution at CNP in a Bi$_2$Se$_3$:Ca(0.09%) crystal at 1.9 K also scales with $H_\perp$. At high fields outside the cusp, the scaling is seen to fail for $\theta \geq 60^\circ$. A fit to HLN theory (solid line) again confirms the contribution only from two surfaces and yields a smaller
dephasing field $B_\phi \sim 0.004$ T (and almost twice as long dephasing length $l_\phi \sim 220$ nm).

Right: Linear magnetoresistance at CNP also scales with $H_\perp$. ................................. 88

Figure 5-6 Penetration depth of electrons in (a) Bi$_2$Te$_3$ and (b) Bi$_2$Se$_3$ calculated using NIST ESTAR simulator (http://physics.nist.gov/PhysRefData/Star/Text/ESTAR.html). The penetration depth at 2.5 MeV (red dots) used in our experiments is $> 2,000$ µm and the resulting depth profile of vacancies is uniform over hundreds of microns................. 92

Figure 5-7 TEM images and diffraction spots of (a) pristine and (b) irradiated Bi$_2$Te$_3$ show the same hexagonal lattice in the ab-plane. (c) Layered van der Walls structure along the c-axis (normal to the cleavage plane) after irradiation. Inset: Rhombohedral layered structure of Bi$_2$Te$_3$ constructed using lattice parameters ($a = 4.38$ Å and $c = 30.45$ Å) from the X-ray diffraction (XRD). The van der Walls structure has three quintuple layers per unit cell.......................................................... 93

Figure 5-8 (a) Longitudinal resistivity $\rho_{xx}$ of the initially p-type Bi$_2$Se$_3$ doped with 0.09% Ca measured in situ in the irradiation chamber maintained at 20 K as a function of electron irradiation dose $\phi$. Initial (1st) irradiation is shown as red squares. The conversion of conductance from p- to n-type takes place at $\phi_{\text{max}} \approx 18$ mC/cm$^2$. Warming up to room temperature partially reverses the compensation process, which can be fully recovered with repeated irradiation (blue circles). The accumulated dose on the 2nd irradiation at $\rho_{xx}^{\text{max}}$ is $\phi_{\text{max}} \approx 24$ mC/cm$^2$. Inset: The change of sign of the slope $\partial R_{xy} / \partial H$ of Hall resistance $R_{xy}$ from below $\phi_{\text{max}}$ to above. (b) Conductivity type can be changed by annealing (see main text) after irradiating to a very low dose $\phi_{\text{max}} \approx 12.7$ mC/cm$^2$ as shown here in another Bi$_2$Se$_3$:Ca crystal by the changing slope of Hall resistivity................................. 95
Figure 5-9 (a) Isochronal annealing of Bi$_2$Te$_3$ crystal irradiated to the dose of 76 mC/cm$^2$ in 30min intervals followed by the resistivity measurement at 4.2 K. (b) Determination of the energy barrier to defect migration from the isochronal annealing step around 100 ºC. ................................................................. 97

Figure 5-10 (a) Sheet resistance $R_{xx}$ vs. temperature of Bi$_2$Te$_3$ crystal at the charge neutrality point. Inset: A fit to a simple activation law near room temperature. (b) A fit to Efros-Shklovskii variable range hopping (VRH) law$^{89}$ in the temperature region marked in light blue. ......................................................................................................................... 98

Figure 5-11 Sheet resistance $R_{xx}$ vs. temperature of another Bi$_2$Te$_3$ crystal at the charge neutrality point reached through the annealing schedule shown in Figure 5-4. Inset: Magnetoresistance at 1.9 K in a pristine Bi$_2$Te$_3$ crystal (blue) and of Bi$_2$Te$_3$ crystal irradiated with electron dose of 1 C/cm$^2$ after annealing step at 120ºC, which returned the crystal to CNP (red), see also Figure 5-4. At CNP a sharp weak antilocalization (WAL) cusp appears, which is not seen in either pristine crystals or after irradiation.................. 100
Chapter 1

Introduction to topological insulators

1.1 Classification of matter

In condensed matter physics, an important way to categorize phases of matter is by recognizing the symmetries they spontaneously break. For instance, magnets break rotational symmetry spontaneously, and superconductors gauge symmetry. All the states of matter in condensed matter systems could be classified by the principle of broken symmetry, until the discovery of the quantum Hall (QH) effect\(^1\,2\). In the quantum Hall States, there is no spontaneous symmetry breaking, however, some properties of quantum Hall states are fundamentally distinct, such as the quantized Hall conductivity and the number of conducting edge modes, and these properties remain intact as the materials parameters go through smooth changes. This phenomenon can be explained only by the topological structure of these quantum states.

The recent discovery of topological insulators rekindled the interest in the topological orders\(^3\). A topological insulator is ideally insulating in the bulk with an energy gap separating the lowest empty electronic conduction band and the highest electron fully occupied valence band. The surface states of a three-dimensional, or the edge states of two-
dimensional topological insulator are conducting with the gap less electronic states. These surface or edge states are topologically protected by time reversal symmetry against back-scattering and distinctive from any previously known electronic systems\textsuperscript{4}.

1.2 Topology

Topology studies the geometric properties and spatial relations unaffected by the continuous change of shape or size of figures. A simple example is that a sphere can be smoothly deformed into a cylinder or a bowl, but cannot be smoothly deformed into a doughnut or a coffee mug, while the doughnut and the coffee mug can smoothly deformed into each other. The former two and the latter two shapes are distinguished by and integer topological invariant called the genus, \( g \), which essentially equals to the number of handles (holes) in the object (Figure 1-1). Surfaces associated with different genuses cannot be smoothly deformed into each other because the integer \( g \) cannot continuously change its value. Surfaces with the same genus can be smoothly deformed into each other, and are said to be topologically equivalent. The geometry of a surface is connected to its topology by the Gauss-Bonnet theorem. For the case of a closed surface in three-dimensional space (Gaussian surface), the integral of the Gaussian curvature \( K \) over such a closed surface \( S \) is equal to \( 2\pi \) times the Euler characteristic of that surface\textsuperscript{5},

\[
\chi = \frac{1}{2\pi} \int_S K dA \quad \text{(1-1)}
\]

One example is a sphere with radius \( R \), where the Gaussian curvature \( K = 1/R^2 \), and \( \chi = 2 \). In general, Euler characteristic and the genus are related by \( \chi = 2 - 2g \), and are both
quantized. In the studies of topological insulators, a topological invariant\textsuperscript{5}, similar to Euler characteristic and genus, is introduced.

![Image](image.png)

*Figure 1-1* The surfaces of a sphere and a bowl are topologically identical ($g = 0$). The surface of a doughnut and a coffee mug are also topologically identical ($g = 1$). The former two and the latter two are distinguished topologically by their genus $g$.

### 1.3 Relating topology to band theory

An energy gap exists in the electronic band structure\textsuperscript{6} of an insulator, separating the fully occupied ground state from all empty excited states. By slowly changing the Hamiltonian, the gapped band structures of insulators can be slowly deformed into each other based on the principle of adiabatic continuity—a physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough and if there is a gap between
the eigenvalue and the rest of the Hamiltonian’s spectrum. Therefore, there is a notion of topological equivalence. Two insulators are topologically equivalent if there exists an adiabatic path connecting the energy structures of the two, along which the energy gap remains finite (non-zero). It follows that the process of an insulator changing to another state topologically inequivalent to its original state is a phase transition, in which the energy gap vanishes.

Crystalline solids where independent electron approximation well applies can be described by the band theory of solids. Also to take into account translation symmetry, single particle states can be labeled by their crystal momentum $k$. Bloch’s theorem tells us these states can be written as $|\psi(k)\rangle = e^{ik \cdot r} |u(k)\rangle$, where $|u(k)\rangle$ is the eigenstate of the Bloch Hamiltonian $H(k) = e^{ik \cdot r} H e^{-ik \cdot r}$. The Bloch Hamiltonian $H(k)$ defines the band structure. Lattice translation symmetry of the crystalline solids dictates $H(k + G) = H(k)$, where $G$ is the reciprocal lattice. It follows that $k$ is defined in the periodic Brillouin zone, which has the topology of a torus $T^d$ in $d$ dimensions. Therefore, an insulator’s band structure is a mapping from Brillouin zone torus to the space of Bloch Hamiltonians with an energy gap. Thus one can classify distinct electronic phases by classifying topologically distinct Hamiltonians $H(k)$.

From the intrinsic phase ambiguity of a quantum mechanical wavefunction arises the Berry phase. Because the Bloch states are invariant under the transformation $|u(k)\rangle \rightarrow e^{i\phi(k)} |u(k)\rangle$, which is similar to the electromagnetic gauge transformation, Berry connection can be defined as $A = -i \langle u(k) | \nabla_k | u(k) \rangle$, which is similar to the electromagnetic vector potential, $A \rightarrow A + \nabla_k \phi(k)$. Berry phase is then defined as $\gamma_c = \oint_c A \cdot dk = \int_S \mathcal{F} d^2k$, where $\mathcal{F} = \nabla \times A$ is the Berry curvature.
In the case of a two level system, the Hamiltonian can be written in terms of Pauli matrices $\sigma$ as

$$H(k) = d(k) \cdot \sigma = \begin{pmatrix} d_z & d_x - id_y \\ d_x + id_y & -d_z \end{pmatrix}$$

(1-2)

The eigenvalue of the Hamiltonian is $\pm |d|$, and the unit vector $\hat{d}$ can be seen as a point on a sphere $S$.

![Diagram of a sphere with a vector $\hat{d}(k)$ and a loop $C$](image)

*Figure 1-2* Berry phase in a two band theory is equal to half the solid angle swept out by the unit vector of $d$.

Berry showed in his work\(^8\) that for a loop $C$ the Berry phase associated with the ground state is

$$\gamma_C = \frac{1}{2} \Omega,$$

(1-3)

where $\Omega$ is the solid angle swept out by the unit vector $\hat{d}$ (Figure 1-2). In the special case where $C$ represents to a $2\pi$ rotation ($\Omega = 2\pi$), the berry phase is $\gamma_C = \pi$. The Berry
curvature $\mathcal{F}$ integrated over a closed two-dimensional space (e.g. a 2D Brillouin Zone $T^2$) is equal to $2\pi$ multiplied by the number of times $\mathbf{d}(\mathbf{k})$ sweeps around the sphere as a function of $\mathbf{k}$. This number is called Chern number, $n = \frac{1}{2\pi} \int_{S} \mathcal{F} d^2 \mathbf{k}$, and it is topologically invariant.

### 1.4 Three-dimensional topological insulators and the topological surface states

In our work, we mainly focus on studying the second generation three-dimensional topological insulators, Sb$_2$Te$_3$, Bi$_2$Te$_3$, and Bi$_2$Se$_3$. This A$_2$B$_3$ materials family shares the same rhombohedral crystal structure with space group $D_{3d}^5 (R\bar{3}m)$ and five atoms per unit cell. For example, the crystal structure of Bi$_2$Te$_3$, shown in Figure 1-3a, consists of a layered structure where individual layers form a triangular lattice, and five-atom layers stacked along the $c$-axis, known as quintuple layers. Each quintuple layer consists of five atoms per unit cell with two equivalent Te atoms (Te1), two equivalent Bi atoms, and a third Te atom (Te2). Atoms within a quintuple layer are coupled by a strong chemical bond, whereas weak van der Waals forces exist between these quintuple layers allowing easy mechanical exfoliation of these materials.

The A$_2$B$_3$ family of materials were first predicted to be topological insulators by Shou-Cheng Zhang and Zhong Fu’s group (Figure 1-4), who using first-principle calculations predicted that these materials should have a single Dirac cone on the surface. The spin of the surface state lies in the surface plane and is always perpendicular to the momentum. These predictions were confirmed by angle-resolved photoemission
Figure 1-3 a. Crystal structure of Bi$_2$Te$_3$. A quintuple layer (a stack of Te1-Bi-Te2-Bi-Te1) is indicated by the red box. b, Angle-resolved photoemission spectroscopy (ARPES) maps the energy states in momentum space. Spin-resolved ARPES reveals that the spins (red) of the surface states lie in the surface plane and are perpendicular to the momentum, as shown in Xia et al., Nature Physics 5, 398 (2009) and Hsieh et al., Science 323, 919 (2009). c. ARPES plot of energy versus wavenumber in Bi$_2$Te$_3$ shows the linearily dispersing surface-state band (SSB) above the bulk valence band (BVB). The dashed white line indicates the Fermi level. The blue lines meet at the tip of the Dirac cone, as published in Science 325, 178 (2009).

spectroscopy (ARPES$^{12,10,11}$ (Figure 1-3c), and scanning tunneling microscopy (STM)$^{13,14,15}$, where the topological surface states with a single Dirac cone were observed.

Furthermore, spin-resolved ARPES measurements detected the left-handed helical spin texture of the massless Dirac fermions$^{10,11}$ (Figure 1-3b) and the absence of backscattering from non-magnetic impurities has been confirmed through the interference pattern in momentum space by STM$^{16,17}$. 
CHAPTER 1 INTRODUCTION TO TOPOLOGICAL INSULATORS

Figure 1-4 Shou-Cheng Zhang and Zhong Fu’s group showed calculation of energy and momentum dependence of the local density of states (LDOS) for Sb$_2$Te$_3$, Bi$_2$Te$_3$ and Bi$_2$Se$_3$, and predicted these systems are three-dimensional topological insulators, published in Nature Physics, 5, 438 (2009).

1.5 Disorders in three-dimensional topological insulators

Unlike the surface-sensitive measurements ARPES and STM, it is much more difficult to investigate transport properties of the topological surface states in three-dimensional topological insulators Bi$_2$Te$_2$, Sb$_2$Te$_3$ and Bi$_2$Se$_3$, due to the large residual bulk carriers, which tend to dominate in the transport measurement$^{18}$. For example, Bi$_2$Se$_3$ is usually intrinsically $n$-type due to Se vacancies. Bi-Te anti-site defects which are promoted by the weakly polarized bond between Bi and Te with similar electronegativity can result in rather large bulk conductivity in Bi$_2$Te$_3$, which can be either $n$- or $p$-type. Sb$_2$Te$_3$ is usually $p$-type due to the Sb-Te anti-sites. These crystal defects introduce large bulk carrier densities, and lower the carrier mobilities. Great efforts have been made to reduce the bulk carriers through chemical doping$^{19,20,21,22,23}$, nanostructure fabrication$^{24,25,26}$, and electric gating$^{27,28}$. Also, bulk resistivity was found to be much higher in ternary compounds, such as Bi-Sb-
Figure 1-5  

Figure 1-5  

a, Electron and hole ‘puddles’ formation on the surface of topological insulator with bulk disorders, observed via scanning tunneling microscopy (STM). b, High energy resolution local density of states (LDOS) taken along a line cut that crosses several charge puddles (white line) on topological insulator with disorders in the bulk, demonstrating the absence of resonances between the puddles and the dopants on the surface. c, Line cut along the yellow line in a showing electron and hole ‘puddles’ on the surface. These three figures are from Beidenkopf et al., Nature Physics 7, 939 (2011).

Te compound\textsuperscript{29}. In Chapter 5, we will describe a new technique we have developed to gain access to the topological surface states in transport. We will show that by using high-energy electron irradiation we can compensate intrinsic charge defects in the bulk via a process
that creates Frenkel (vacancy-interstitial) pairs, thereby reducing bulk conductivity by orders of magnitude to charge neutrality point.

Charge inhomogeneity arising from bulk disorder can result in pronounced nanoscale spatial fluctuations of energy on the surface, promoting the formation of surface ‘puddles’ of different carrier types. Such ‘puddles’ were experimentally observed in graphene\(^3\) and were recently detected in several topological insulators through direct imaging spatial variation of Dirac point via scanning tunneling microscopy (STM)\(^3\). Energy resolved differential conductance (dI/dV) was mapped on the surface of single crystalline topological insulators Bi\(_2\)Te\(_3\) and Bi\(_2\)Se\(_3\) doped with either Ca or Mn (Figure 1-5a), where n- and p-type ‘puddles’ on the surface were observed (Figure 1-5c). The energy fluctuations on the topological surfaces were attributed to the underlining bulk disorder, identified primarily as poorly screened charged defects. It should be noted that no direct correlation of the surface Dirac puddles resolved in STM and the locations of charged surface defects was established (Figure 1-5b). We will show in Chapter 3 that surface ‘puddles’ in the A\(_2\)B\(_3\) topological family do enable us to observe a robust singular magnetic response from the Dirac spins.

### 1.6 Weak localization and weak anti-localization

In disordered system at low temperatures, a quantum phenomenon known as weak localization manifests itself as a *positive* correction to the resistivity of a metal or semiconductor\(^3\). In a disordered electronic system the electrons move in a diffusive motion rather than ballistic. The interference between different scattering paths increases the probability of electron ‘wandering around in a circle’ than it would otherwise, due to the
constructive interference of the clock-wise and counter-clock-wise loops, thus leading to an increase in the net resistivity. Since it is much more likely to find a self-crossing trajectory in low dimensions, the weak localization effect happens more easily in a lower-dimensional system, such as thin films and wires\textsuperscript{33}.

In a system with spin-orbit coupling (such as A$_2$B$_3$ family of three-dimensional topological insulators), the spin of the charge carrier is coupled to its momentum. The spin rotates as it goes around a self-intersection path, and the direction of this rotation is opposite for two directions about the loop, thus the two paths interfere destructively with each other, leading to the lowering of the net resistivity\textsuperscript{34} (Figure 1-6). In such cases the phenomenon is called weak anti-localization.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure1-6}
\caption{Weak localization is due to self-intersecting paths of electron scattering in disordered metals or semi-metals. When spin-orbit coupling is present, the clock-wise and counter-clock-wise loop interfere destructively, leading to weak anti-localization.}
\end{figure}

Hikami, Larkin and Nagaoka showed in their work\textsuperscript{34} that in two-dimensional systems, the change in the conductivity in an external magnetic field can be described by the following Hikami-Larkin-Nagaoka (HLN) equation:
\[ \sigma(B) - \sigma(0) = \frac{e^2}{2\pi^2\hbar} \left[ \ln \left( \frac{B\phi}{B} \right) - \psi \left( \frac{1}{2} + \frac{B\phi}{B} \right) \right] \]

\[ + \frac{e^2}{\pi^2\hbar} \left[ \ln \left( \frac{B_{SO} + B_e}{B} \right) - \psi \left( \frac{1}{2} + \frac{B_{SO} + B_e}{B} \right) \right] \]

\[ - \frac{3e^2}{2\pi^2\hbar} \left[ \ln \left( \frac{(4/3)B_{SO} + B\phi}{B} \right) \right] \]

\[ - \psi \left( \frac{1}{2} + \frac{(4/3)B_{SO} + B\phi}{B} \right) \] (1-4)

where \( \psi \) is the digamma function, \( B\phi \) is the phase coherence characteristic field, which is roughly the magnetic field required to destroy phase coherence, \( B_{SO} \) is the spin-orbit characteristic field which can be considered a measure of the strength of the spin-orbit interaction and \( B_e \) is the elastic characteristic fields.

In the limit of strong spin-orbit coupling \( B_{SO} \gg B\phi \), the HLN equation reduces to:

\[ \sigma(B) - \sigma(0) = \alpha \frac{e^2}{2\pi^2\hbar} \left[ \ln \left( \frac{B\phi}{B} \right) - \psi \left( \frac{1}{2} + \frac{B\phi}{B} \right) \right] \] (1-5)

The characteristic phase coherence field \( B\phi = \hbar/4e l_{\phi}^2 \), where \( l_{\phi} \) is the dephasing length, which can be understood as the distance traveled by electrons before they lose phase coherence. In this equation, \( \alpha = -1 \) for weak localization and \( \alpha = 1/2 \) for weak anti-localization.34

HLN equation provides a means for us to estimate the number of quantum channels corresponding to the surface states in the three-dimensional topological insulators with insulating bulk, and we shall see in Chapter 5 that this is indeed the case.
1.7 Superconductivity in topological insulators

The discovery of topological insulators stimulated the search for even more exotic particles and excitations, such as the topological superconductor\textsuperscript{35,36}. Unconventional superconductivity involving electron bands with nontrivial topological character has been predicted\textsuperscript{35–39} to arise on the surfaces of three dimensional topological insulators. The superconducting states in the topological insulators are expected to host on the surface unusual Majorana fermions with special non-Abelian statistics\textsuperscript{40}. Majorana fermions are their own antiparticles\textsuperscript{35}, and are currently of significant interest in condensed matter physics because of their novelty and their potential of leading to a new paradigm change in approaches to quantum computing, quantum device engineering and materials science.

The superconducting phases in topological insulators have recently been found in Cu-doped Bi\textsubscript{2}Se\textsubscript{3}\textsuperscript{40–42}, and in Bi\textsubscript{2}Te\textsubscript{3}\textsuperscript{43} and Sb\textsubscript{2}Te\textsubscript{3}\textsuperscript{44} under very high (GPa range) pressures. In these studies, the materials were found to have increased bulk carrier densities and thus their superconductivity appear to be of bulk origin.

In our study, we found Sb\textsubscript{2}Te\textsubscript{3} synthesized under a narrow range of modest Te pressure exhibited superconductivity. We will present in Chapter 4 strong evidence that this striking superconductivity originates in surface puddles and thus can be a potential candidate for hosting Majorana modes.
Chapter 2

Experimental methodologies

2.1 Bridgman-Stockbarger growth of single crystalline topological insulators

We use Bridgman-Stockbarger technique to synthesize single crystalline topological insulators Bi$_2$Te$_3$, Sb$_2$Te$_3$, Bi$_2$Se$_3$ and ternary topological insulators. The starting materials are sealed inside a quartz tube under vacuum ($< 10^{-3}$ Torr), and then are heated above their melting temperatures (to above ~1000 °C) in a tube furnace. One side of the quartz tube is slowly cooled (~ 1 – 10 °C/h) so that a seed crystal is grown at that end. As the quartz tube is cooled from one end to the other, the precursors in the quartz tube form a single crystal along the length of the tube that matches the crystallographic orientation of the seed crystal.

Typically, a quartz tube of inner diameter 8 mm and outer diameter 12 mm was melted in the middle with a torch (using propane and oxygen), separating the tube into two segments, each with one end sealed in the shape of a sharp cone that was suitable for a seed crystal to grow. The starting materials, centimeter-sized chunks of Sb, Te, and/or Bi, Se
(purchased from Alfa-Aesar, purity 99.9995% to 99.9999%) used in stoichiometric ratios, were placed inside one of the tube segments with one of its ends sealed. The open side of the quartz tube was attached to a vacuum pump to evacuate air from the tube’s interior. When the pressure inside the tube was sufficiently low ($< 10^{-3}$ Torr), the tube was torch-sealed under vacuum at a place along the length of the tube segment pre-determined by the tube volume chosen for the given crystal growth run. The torch-sealing is a slow process, with quartz walls shrinking and eventually melting and closing at a torching point—this makes a completely sealed chamber under vacuum with all starting materials inside. The sealed quartz tube was subsequently pulled further and separated at the molten section from the rest of the quartz.

The final quartz tube segment with starting materials sealed inside at both ends under vacuum was then placed in either a horizontal or vertical furnace. In the horizontal furnace (Figure 2-1a), a temperature gradient is naturally established from the center (the highest temperature) to the edge (the lowest temperature) of the furnace. The furnace was heated above the melting temperatures of the materials. The temperature was maintained for one day to a week depending on the type of crystals grown. The starting materials were in liquid or gas states, and thoroughly mixed during this period of time. The furnace was then slowly cooled to room temperature for a period ranging from a week to a month. During the cooling, a seed crystal was first formed at the colder end of the tube. On further cooldown the starting materials crystalized, matching the lattice of the seed crystal. The growth progressed as the solid-liquid interface controlled by the freezing temperature front that was moving along the quartz tube. Sometimes a 2\textsuperscript{nd} regrowth was performed to
eliminate defects associated with lattice mismatch to the seed, thereby obtaining crystals of better quality.

Figure 2-1 Sketches of (a) horizontal Bridgman growth and (b) vertical Bridgman-Stockbarger growth of single crystalline topological insulators.

In the case of vertical growth, a baffle was used to separate two zones in our homemade vertical tube furnace (Figure 1-2b). The upper zone had a temperature above the melting temperature of the crystal grown, and the lower zone below. The sealed quartz tube with the starting materials was hooked with a metal string and hang vertically inside the tube furnace so it could move freely. The string was routed by a pulley and connected to a motor, which controlled the position of the quartz tube inside the furnace. The quartz tube was kept in the hot zone for a period of time needed for the starting materials to mix in liquid and/or gas forms. The quartz tube was then lowered slowly (~ 0.5 – 2 cm/day) with the control of the motor during a time period from one week to a month, depending
on the crystal synthesized. The seed crystal was formed when the lower end of the quartz tube crossed the baffle entering the cold zone. As the quartz tube moved further into the cold zone, more of liquid crystalized at the baffle, matching the crystal lattice of the seed crystal until the whole quartz tube moved into the cold zone. Again, sometimes a second growth was required as well to eliminate any remnant crystal lattice mismatch. In general, our vertical growth produced better quality of crystals than the horizontal growth.

2.2 Fabricating samples of topological insulator crystals for transport measurement

Two types of samples of single crystals grown from Bridgman-Stockbarger were used in our experiments. One type was a chunky crystal of rectangular prism shape cut with a tungsten wire saw. The other type was an exfoliated thin flake. The exfoliation was done using Scotch tape, following the technique developed previously for graphene\(^4\)\(^5\).

Before cutting the crystal with a wire saw, the \(c\)-axis and the \(a\)-\(b\) plane of the topological insulator crystal were identified, and the crystal was fixed on a mount with melted wax, so that the desired orientation cut could be easily made. We cut the crystal into about 1mm \(\times\) 1mm \(\times\) 15µm sized samples, with the \(c\)-axis pointing along the 15µm thickness. Electrical contacts to the sample were then made by spark-welding as follows. A 200 µF capacitor was charged with a DC voltage source, usually of 4 V to 10 V depending on the material, and the voltage source was subsequently removed. One electrode of the capacitor was connected to a pair of tweezers that clamped the crystal, and the other electrode was connected to another pair of tweezers that could pick up a short 12
µm diameter gold wire and bring this wire in contact with the sample. On contacting the crystal, the electric charge stored in the capacitor would flow through the now closed circuit and melt the wire at the contact point due to the large contact resistance. The contacts made this way to large crystals were very robust. In these samples the spark-welded contacts were made at the four corners of the crystal sample, designed for van der Pauw transport measurements. The contacted samples were fixed on a chip carrier (Figure 2-2) with GE varnish and the gold wires were attached to the electrodes of the chip carrier with silver paint or silver epoxy.

![Figure 2-2](image)

Figure 2-2 (a) an eight-pin chip carrier carrying a van der Pauw sample. (b) An exfoliated 3D topological insulator flake with lithographical van der Pauw contacts.

The topological insulators Bi₂Te₃, Bi₂Se₃, Sb₂Te₃ are easily exfoliated due to their van der Waals layered structure. Pressing a Scotch tape on a crystal easily transfers thin crystal flakes to the tape which is then pressed onto a clean SiO₂/Si wafer. Only a moderate pressure needs to be applied from the back of the Scotch tape for about one minute before
the tape is peeled off. The flakes successfully transferred to the wafer are typically 20 nm to 100 nm thick and are 20 µm to 50 µm wide. The residue of the glue from the Scotch tape is removed by acetone and alcohol. In our samples, contacts in Hall bar or van der Pauw geometry are defined by either photolithography or e-beam lithography. Once the resist is developed, the sample is etched in 0.1 mol/L HCl solution for several minutes, rinsed with distilled water, dried with nitrogen gas, and a layer of 25 nm to 50 nm gold is then sputtered; this we found to ensure good Ohmic contacts. A layer of 3 nm titanium and 50 nm to 100 nm gold is then evaporated onto the sample before lifting off the resist. The gold electrodes are wired to the sample holder’s pins using silver epoxy, silver paint, or wire bonding.

2.3 Electrical transport measurement instrumentation setups

Transport measurements were performed primarily in Quantum Design Physical Property Measurement System (PPMS) with a 14 Tesla superconducting magnet and a temperature range from 1.9 K to 400 K. PPMS’s Model 6000 provides source-measurement units for simple four-terminal resistivity measurements. However, Model 6000 has only three sets of source-measurement channels, and it is not capable of measuring van der Pauw samples to obtain both longitudinal resistance $R_{xx}$ and Hall resistance $R_{xy}$ in one sample, where four sets of source-measurement channels are required. Moreover, when measuring samples with demanding conditions, such as very high resistance samples, the source-measurement units in Model 6000 are not adequate or capable to perform.
To overcome these problems we designed an external measurement setup. Keithley 6221 was used as both DC and AC current source. HP 34401A and Keithley 2002 were used for the voltage measurements when sample resistance was far less than 10 MΩ. HP 34401A has two options for its internal impedance, 100 MΩ and 10 GΩ. It is capable to measure samples of resistance larger than 10 MΩ but far less than 10 GΩ when its internal impedance is set to 10 GΩ. To measure higher resistances we used Keithley 6517B electrometer, which is capable of measuring resistances up to $10^{16}$ Ω. It is noteworthy that the ultra-high internal impedance of the Keithley electrometer may cause electric charge accumulation and can damage fragile samples during the measurement. Unless the ultra-high internal impedance is necessary and the sample is robust, HP 34001A was preferred as a safer practice for both low and high resistance samples. For low resistance samples, the voltage signal can be very small and the signal-to-noise ratio is not satisfactory in DC measurements, in which case AC transport measurements were preferred. Keithley 6221 provided AC current ranging from 2 pA to 100 mA at frequencies up to 100 kHz. We used Signal Recovery 5210 lock-in amplifier with analog demodulators; it has an excellent signal-to-noise ratio, and is widely referenced as the benchmark lock-in amplifier.

### 2.4 AC susceptibility measurement system

AC magnetic susceptibilities of the topological insulators were measured in Quantum Design PPMS with AC measurement system (ACMS) option. The ACMS probe has a sample space allowing 5 mm wide sample to be measured. The sample space is rigged with an AC drive (excitation) coil that provides uniform sine AC excitation field of amplitude up to 15 Oe and frequency up to 10 kHz. Two coils of identical geometry are connected in
series but with opposite winding direction, so the emf induced by the uniform AC excitation field in these two coils cancel each other out. The sample is then placed in one of the two coils, so that the emf induced by the sample’s magnetic response can be picked up only by this coil (called the pickup coil, the other coil is called the compensation coil, see Figure 3 1b). The magnetic response of the sample is then sent to a lock-in with the excitation voltage that drives the excitation coil as the reference, so that the sample’s AC susceptibility $\chi_{AC} = dM/dH$ is obtained. The lock-in outputs the phase difference between the excitation and the measured magnetic response, the real component of the susceptibility $\chi''$ and the imaginary component of the susceptibility $\chi'$. Higher order harmonics (up to the 10th harmonics) were also routinely recorded.

The AC excitation field can induce large Eddy current in metallic samples, and this Eddy heating can raise the temperature of the sample. Therefore when performing AC magnetization measurements at low temperature near 1.9 K we used excitation field of 1 Oe amplitude at 10 kHz.
Chapter 3

Singular robust Dirac spin response

3.1 Introduction

The interplay between condensed-matter physics and materials science is often driven by established technological materials that turn out to be remarkably good model systems for fundamentally new physical phenomena which in turn can lead to disruptive technological advances. Topological insulators are one recent example—prized thermoelectrics\textsuperscript{46} since the 1950s, they also host topologically protected spin-helical surface states, as predicted by theory\textsuperscript{47} (Figure 3-1a) and subsequently confirmed in a series of angularly resolved photoemission spectroscopy (ARPES) experiments\textsuperscript{11,48,49}. Much of the activity since has been inspired by prospects of harvesting exotic properties of these helical states for the electrical manipulation of magnetic memory\textsuperscript{50} and error-free topological quantum computing\textsuperscript{36}. 
At present, considerable effort is being aimed at improving the synthesis and characterization of these compounds with the goal of realizing materials with strongly suppressed bulk conduction channels—the latter tend to obscure surface physics, a problem particularly severe in charge transport. Indeed, complex intermixing (hybridization) of the bulk and surface states is clearly observed by a variety of surface probes; for example, recent time-resolved ARPES experiments reveal strong phonon-assisted coupling between the surface and bulk electronic states at high lattice temperature and a unique cooling of Dirac fermions by acoustic phonons. ‘Ageing’ effects arising from complex surface reconstruction processes are also observed—they tend to promote the formation of 2D electron gas states of bulk origin in close proximity to the topological Dirac surfaces. Thus, existing materials continue to present a number of challenges to a complete understanding of the physics of topological Dirac metals, especially at low frequencies and on mesoscales. Magnetic susceptibility measurements reported in this work witness the singular magnetic

*Figure 3-1* a, The energy-momentum relation of the surface states in a 3D topological insulator. b, Schematic for the measurement of magnetic susceptibility.
response of topological surface states, but also hint at an intriguing cooling process involving these surface states and bulk carriers, thereby paving the way for the systematic exploration of low-energy electrodynamics of these transformative materials.

3.2 Spin response from Dirac fermions

The experiments were performed using a weak low-frequency AC excitation field (see Figure 3-1b and Methods) to probe the linear response, focusing on its in-phase component, which is the equilibrium susceptibility $\chi(B) = \partial M(B) / \partial H$ in the limit of zero frequency and in a range of DC fields $B = \mu_0 H$, including the vicinity of $B = 0$. Figure 3-2 shows the susceptibility of the canonical second generation topological insulator Bi$_2$Se$_3$ measured in DC fields $H \parallel c$ axis (normal to the (00\bar{1}) cleavage surface) of a platelike shaped crystal. Above ~0.5 T the response is diamagnetic, consistent with a decades old magnetic susceptibility measurement. At lower fields, however, we detect a large cusp-like paramagnetic susceptibility that sharply rises above the diamagnetic ‘floor’ in a narrow DC field range of ~0.2 T and approaches $\chi(H \to 0)$ in a straight line (Figure 3-2). This singularity arises from the sample’s surface, is robust across all topological samples measured, and is most naturally ascribed to the opening of a Zeeman gap at the Dirac point of the helical metal. Before we turn to substantiating these claims we note one particularly spectacular aspect of our data—its thermal stability. Indeed, the singular field dependence of the susceptibility shows no discernible signs of rounding up to the highest (room) temperature measured. This persistence of singular response to elevated temperature is remarkable and surprising when confronted with a rough conservative
estimate of the expected thermal smearing, for example, obtained from the ratio of thermal energy at 300 K (≈27 meV) to the rather small bulk gap of these materials ~100-300 meV.

![Graph of magnetic susceptibility vs. magnetic field]

*Figure 3-2* Magnetic susceptibility of Bi$_2$Se$_3$ measured by applying a small AC excitation field $H_{ac}$ shows that the spin response is cusp-like and large near zero applied DC magnetic field.

The presence of the cusp in near-zero-field susceptibility is universal—it is observed in all three topological insulators: Sb$_2$Te$_3$, Bi$_2$Te$_3$, and Bi$_2$Se$_3$ (Figure 3-4 a-c). It is absent in all our calibration and background materials (Figure 3-3), which were carefully screened for any spurious signals. At higher fields, $H \gtrsim 0.5$T, the temperature-dependent diamagnetism dominates (Figure 3-4 d-f); it seems to correlate with the details of the bulk band structure, but less clearly with the particulars of donor (n-type) or acceptor (p-type) intrinsic defects (Figure 3-4 g-i) present in the bulk.
Figure 3.3. Diamagnetic susceptibility of the sample holder used in the experiments is two orders of magnitude smaller than the typical signal from the sample signal. Susceptibility of precursor materials used in the crystal growth, Te and Sb, is depicted in panels b, d, and c, respectively. d. Susceptibility of Te after annealing at 450 °C in the growth furnace is also featureless. The well-known large diamagnetism of Sb is enhanced at low temperatures (similar to Bi) while the diamagnetism of Te is only weakly temperature dependent. Both are in good agreement with the literature values. The signal from the paramagnetic Pd calibration sample is shown in e. Susceptibility of the layered 2H– NbSe₂ shows a well-known behavior in the normal and superconducting states⁵⁷ and no paramagnetic cusp, see f. These essentially featureless background/calibration checks are in contrast with the cusplike paramagnetic field dependence at low fields consistently observed in the crystals of Sb₂Te₃, Bi₂Te₃, and Bi₂Se₃.
Figure 3-4 Universality of singular spin response near zero magnetic field. a-c, The zero-field susceptibility cusp is found in all three topological insulators. d-f, The susceptibility surface in the H-T phase space for fields above \( H \approx 0.5T \) is shown for Sb\(_2\)Te\(_3\) (d), Bi\(_2\)Te\(_3\) (e) and Bi\(_2\)Se\(_3\) (f). The most pronounced temperature dependence is found in Sb\(_2\)Te\(_3\) (d) which has the smallest bulk bandgap of \( \sim 100 \) meV. g-i, Corresponding schematic band structures (insets) indicate noticeable differences in the location of the Dirac point relative to the bulk valence and conduction bands. Measurements of Hall resistivity (blue lines in g-i) show that the Te-based TIs, Sb\(_2\)Te\(_3\) (g) and Bi\(_2\)Te\(_3\) (h), are intrinsically p-type, whereas the Se-base TI, Bi\(_2\)Se\(_3\) (i), is n-type.

The height of the cusp is evidently slightly sensitive to the density of defects quenched in during the crystal growth, and there is an ageing effect\(^5\) that can reduce the height over time by an appreciable factor (up to five; an example is shown in Figure 3-5). The absolute magnitude of the cusp in different crystals varies slightly with the intrinsic
bulk carrier density, which in any particular crystal is determined from measurements of Hall conductivity (Figure 3-4) or Shubnikov-de Haas (SdH) quantum oscillations (Figure 3-6). However, the ‘cuspiness’, as quantified by the $B = \mu_0 H \rightarrow 0$ slope, for any given member of this topological insulator family is universal. An example of this is shown in Figure 3-7a, where we compare two Bi$_2$Te$_3$ crystals with carrier concentrations differing by two orders of magnitude. The cusp is independent of frequency (Figure 3-9a), as expected for such a low-frequency response ($2 \sim 10$ kHz). This is further confirmed by the singular signal in the differential susceptibility obtained from the DC magnetization measurements using a superconducting quantum interference device (SQUID) magnetometer (Figure 3-10). Finally, compelling evidence that the cusp originates from the surface states is illustrated in Figure 3-7b, which shows that for the same crystal area, when the sample thickness is reduced the height of the cusp remains unchanged, whereas the diamagnetic background closely scales with the volume. We note that a similar, albeit weaker, response is detected with the sample rotated by $90^\circ$, consistent with the signal originating from non-cleaving surfaces where the Dirac dispersion is more complex.
Figure 3-5 Susceptibility cusp for a Bi$_2$Se$_3$ crystal at several temperatures measured (a) an hour after the crystal growth and (b) two weeks later after the crystal was stored in flowing nitrogen. While the temperature robustness is intact, the overall cusp height has been reduced with time likely due to surface reconstruction and the formation of two-dimensional electron gas (2DEG) associated with the band bending of the bulk states at the surfaces$^{34,55}$.

Figure 3-6 a, Shubnikov-de Haas (SdH) oscillations for a Bi$_2$Te$_3$ crystal with carrier density $n \sim 10^{17}/$cc; b, The oscillations are fitted to Lifshitz-Kosevich formula using Monte Carlo technique. Below 8 Tesla field, the fit gives cyclotron mass $m_c = 0.0767m_e$, where $m_e$ is bare electron mass. The obtained value of the $g$-factor is $g \sim 2m_e/m_c \approx 30$, of the order of $g$-factors $\sim 60$ reported in other experiments$^{23}$. 
CHAPTER 3 SINGULAR ROBUST DIRAC SPIN RESPONSE

Figure 3-7 Signature of the surface origin of the cusp. a, Susceptibility cusp for two Bi$_2$Te$_3$ crystals with carrier densities differing by two orders of magnitude. The slope of the cusp is independent of the bulk carrier density $n$. Here the diamagnetic background was subtracted and the height of the cusp was normalized to $\frac{\Delta \chi}{\Delta \chi_{B=0}}(B=0)$, which for the $n \sim 10^{19}$ cm$^{-3}$ crystal was $3 \times 10^{-5}$ emu/cc, and for the $n \sim 10^{17}$ cm$^{-3}$ crystal was $3.5 \times 10^{-5}$ emu/cc. b, Left: The susceptibility cusp before and after cutting the crystal thickness by factors of 0.63, 0.29 and 0.15 seems to be independent of the thickness $t$. The diamagnetic background scales with thickness. Right: The data for all thicknesses shown on the left shifted to match the diamagnetic background.

Our finding of a prominent singular magnetic response that survives high temperatures, huge variations in carrier density, and does not scale with sample volume is quite surprising and as far as we know unprecedented. In the absence of any paramagnetic impurities or signs of itinerant ferromagnetism, the origin of this particular low field anomaly may be traced most naturally to the ungapped Dirac point. The simplest description of the Dirac fermions is captured by a non-interacting Rashba-type Hamiltonian that effectively locks electron spin to its momentum—that is, parallel to the sample’s surface. The effect of a magnetic field applied transverse to the surface enters through a Zeeman coupling which we treat explicitly and via orbital quantization which
we ignore (this approximation is justified by the absence of oscillatory effects at low fields in our experiments, and, a posteriori, we can also confirm that the Dirac Landau level spacing is essentially negligible compared to the Zeeman gap in the parameter range relevant to our experiments). The equilibrium susceptibility is obtained by taking the second derivative of the total free energy with respect to the magnetic field \( B \). With both

![Figure 3-8 a. The simple Dirac model of equation (3.1) produces a very good match to the data, as illustrated for the case of \( Sb_2Te_3 \). Here \( \chi = x\chi_A/L_z \) and \( \chi_A \) is the 2D susceptibility of the Dirac state, \( L_z \approx 10^{-3} m \), thickness of our samples, and \( x < 1 \) the effective areal fraction occupied by the ungapped Dirac state (\( x \) is used as a fitting parameter). Other parameter values used to generate this plot are \( \mu = k_B T = 0 \), \( g = 60 \), \( v_F = 2 \times 10^3 \) m s\(^{-1} \), which are known from our own studies (Supplementary Information) and those of others\(^21\). Both \( x \) and \( \Lambda \) (effective radius of k-space contributing to singular response) were adjusted to match the data, producing \( x \approx 0.002 \) and \( \Lambda = 5 \times 10^8 \) m\(^{-1} \). The cusp is preserved even when hexagonal warping (inset in a) is taken into account—it is merely subsumed into \( \Lambda \). b. Rare regions of chemical potential \( \mu \approx 0 \) (grey) can exist between electron (blue) and hole (yellow) droplets owing in part to the electrostatic potential established by the charged defects in the bulk. Such fluctuations of the local surface charge are probably 'healing' in the course of the ageing process as the mean chemical potential steadily floats away from the Dirac point towards bulk conduction or valence bands, as has been documented in ARPES studies. This is qualitatively consistent with the observed decrease in the amplitude of the paramagnetic anomaly over time.
the chemical potential $\mu$ and temperature set to zero, the low-field areal (sheet) susceptibility $\chi_A$ (see next section Discussion) reduces to

$$\chi_A \cong \frac{\mu_0}{4\pi^2} \left[ \frac{(g\mu_B)^2 \Lambda}{\hbar v_F} - \frac{2(g\mu_B)^3}{\hbar^2 v_F^2} |B| + \cdots \right]$$  \hspace{1cm} (3-1)$$

where $g$ is the Landé $g$-factor and $v_F$ is the Fermi velocity. This paramagnetic Dirac susceptibility has the form of a cusp with a linear-in-field decay at low fields, just as the cusp observed in our experiments (Figure 3-8). The maximum of $\chi_A$ depends on the effective size of the momentum space $\Lambda$ contributing to the singular part of the free energy, and thus may be controlled in part by hexagonal warping of the Dirac cone$^{59}$ and by the details of the bulk bands. However, the singular field dependence depends only on universal (low-energy) parameters through the slope $2 (g\mu_B)^3 / (\hbar^2 v_F^2)$ of $\chi$ in the limit $B \to 0$. To compare with the experiment we write the total susceptibility as a sum of the background contribution $\chi_0$ and surface contribution $\chi = \chi_0 + \chi_A x/L_z$, where $x$ is the fraction of the surface contributing and $L_z \approx 1$mm is the sample thickness. We obtain a good match to the shape and the magnitude of the cusp (Figure 3-8a) by using parameter values consistent with the reported velocity $v_F$ in Bi$_2$Te$_3$ from Landau level spectroscopy$^{60}$ and the large effective $g$-factor$^{23}$, broadly consistent with the overall scale of $g$-factors expected for topological insulators and obtained from our SdH measurements (Figure 3-8b and Figure 3-6a). The participating surface fraction that emerges from this analysis is remarkably small, $x \approx 0.002$, meaning that these states are very rare.

The existence of the sharp non-analytic paramagnetic cusp at zero temperature requires the surface Fermi level to be at the Dirac point, $\mu = 0$. Otherwise, for $\mu \neq 0$, we expect a smooth dependence (rounding) near $B = 0$ with sharp jump singularities in $\chi$ on a field scale $\delta B = \mu / (g\mu_B)$ where the Fermi level enters the valence or conduction band.
Further phenomenological description can be facilitated by recasting the low-field paramagnetic response in Equation (3-1) in terms of the effective Dirac bandwidth $W = \hbar v_F \Lambda$ and field energy $E_B = g\mu_B B$ as $\chi_A(B) = ((g\mu_B)^2\Lambda^2)/(W)(1 - (2E_B)/(W)) + \cdots$, so the characteristic width of the cusp is set by the condition $W \approx E_B$. The observed temperature insensitivity requires that the thermal energy $E_T = k_B T \ll E_B < W$, or $T \lesssim 10\text{K}$—which may be relaxed slightly on the level of this simple phenomenology if both the $g$-factor and Fermi velocity are temperature dependent.

In our AC experiments, no appreciable rounding of the cusp is observed—this finding is profoundly unexpected in view of the location of the Fermi level gleaned from ARPES or scanning tunneling microscopy (STM). Separate experimental work will be required to obtain a clear and detailed understanding of the microscopic origins of the electronic states giving rise to the singular response. From the established surface nature and the observed ageing effects we infer that renormalization of the effective potential near the sample’s surface in the course of ageing is important. Also, the remarkable robustness to the variation in bulk carrier density, and therefore bulk screening length, suggests that electrostatic models invoking bulk dopants as the dominant source of disorder at the surface may not be adequate to capture these states. Such models do readily produce the large-scale inhomogeneities of the chemical potential, $\mu$, which have been observed, for example, in graphene and recently directly mapped in several topological insulators via scanning tunneling microscopy. The typical amplitude of inhomogeneity in the latter study, 10~20 meV, seems too small to couple to the electronic states near the Dirac point. However, rare states, that based on our analysis occupy only approximately 0.2% of sample’s surface, may not be readily observed in STM. Moreover, the role played by
unavoidable differences in surface preparation among different experiments remains to be established.

**Figure 3-9 a.** The in-phase component of the susceptibility containing the singular cusp is independent of frequency (shown here for Sb$_2$Te$_3$). However, the diamagnetic susceptibility is slightly dependent on frequency. **b.** The nonlinearity of the surface–bulk connection is witnessed by the observed second harmonic of $\chi$. It is consistent with the existence of ‘rectifying’ paths in the putative thermoelectric cooling elements required for the cooling of a small fraction of the sample’s surface and thus suppressing thermalization of Dirac surfaces with the bulk, as explained in the text. The effective cooling of the surface is naturally achieved by the electron and hole puddles in the subsurface region, forming a Peltier element (inset) owing its cooling efficiency partly to nanoconstriction and partly to frequency-dependent transport coefficients.

Yet another intriguing finding in our experiments is the apparent thermal stability of the singular AC response. This is certainly not within our simple Dirac phenomenology, which has in it scales on the order of only 10 K. In fact, we may argue that any equilibrium theory of the singular response in these narrowband semiconductors must show thermal effects near room temperature, as the band gap is only a few times larger, at best. Indeed,
in DC magnetization measurements using SQUID the singular response at higher

\[ \frac{dM}{dH} \]

Figure 3-10 DC magnetization \( M \) measured in the Superconducting Quantum Interference Device (SQUID) magnetometer shows clear nonlinearity near zero field. Numerical derivative \( \frac{dM}{dH} \) of \( M \) gives a ‘cusp’, as shown in the inset. Consistent with our model in Section II the dc cusp singularity is rounded and diminished with decreasing temperature. We note that taking numerical derivatives of magnetization in the vicinity of zero DC magnetic field expectedly produces a spurious numerical noise. Hence, not surprisingly, data taken at a finite frequency, where the derivative is taking ‘in situ’ using a small AC oscillation gives a much more accurate record of magnetic susceptibility \( \chi = \frac{dM}{dH} \) near \( H \sim 0 \). The derivative noise is partly controlled by the lock-in amplifier in the ac detection circuit and it decreases at higher excitation frequencies.

temperatures is rounded (Figure 3-10). We propose, therefore, that the local temperature at the location of electronic states responsible for the cusp is, in fact, strongly affected by the AC probe itself—that is, these patches are kept at very low, possibly cryogenic effective temperature even though the cryostat and the rest of the sample are ‘warm’. One plausible, albeit still speculative, scenario (Figure 3-9) for this invokes disorder as the origin of local Peltier elements. The most natural source of power for the putative Peltier cooler is the
rather large eddy current, which does not contribute to $\chi$ itself but rather to the imaginary, out-of-phase part of $\chi(\omega)$ (Figure 3-12). To suppress Peltier heating (unavoidable owing to AC excitation), this would require a rectifying element as well (Figure 3-9 and Figure 3-14). From general considerations of the rectification process there should then be second harmonic generation, which we clearly observe (Figure 3-9b Figure 3-15). The above scenario implies strong enhancement of the effective (local) thermoelectric figure of merit as compared to known bulk values for these materials\textsuperscript{62}, which would be natural, based on the existing work on improved thermoelectricity in nano-constrictions\textsuperscript{63,64}, and on the strong frequency dependence of the transport coefficients under geometric confinement, as in the case of phonon heat conductivity\textsuperscript{65}. We also note that strong (local) variations of material properties, for example, due to the presence of disorder, can give rise to a novel variant of thermoelectric cooling, a ‘Thompson cooler’, which has been predicted to show significant improvement of performance and, in principle, enable cooling to very low, even cryogenic temperatures\textsuperscript{66}. Detailed theory of the mechanism of thermal stability is beyond the scope of this work and should be further explored.

Our experiments reveal a singularity in the low-field response in a whole family of materials with topological surface states which does not arise from either strong correlations or fine tuning the chemical potential to the Dirac point. They are profoundly counterintuitive, as they suggest the controlling role of rare states (patches) near the Dirac point realized under generic surface conditions in these samples. With this assumption we are able to reproduce the overall shape and magnitude of the response. One of the surprising quantitative insights that emerged was that a minority ($\approx 0.2\%$) of the surface is responsible for the singular signal. This simple phenomenology is a step forward to a
precise theoretical understanding and improved experimental control of these phenomena that will be crucial for manipulating the robust polarization of protected surface states at room temperature.

3.3 Landé \( g \)-factor from Shubnikov de Haas

oscillations

Topological insulators are expected to have strong Zeeman effects based on previously reported values of Landé \( g \)-factor of about 60\textsuperscript{23}. We have measured magneto-oscillations of the bulk conductivity and used it to deduce \( g \approx 30 \), fitting the oscillations with the Lifshitz-Kosevich equation

\[
\frac{\Delta \sigma_{xx}(T)}{\Delta \sigma_{xx}(0)} = \frac{\lambda(T)}{\sinh \lambda(T)}.
\]

Here \( \sigma_{xx} \) is the in-plane conductivity for magnetic field applied normal to the cleavage plane, \( \lambda(T) = \frac{2\pi k_B T}{\hbar eB} m_c \), and \( m_c \) is the cyclotron mass. This is illustrated in Figure 3-6 for Bi\(_2\)Te\(_3\) crystal with carrier density \( n \sim 10^{17}/\text{cc} \). Determination of the differences between the surface and bulk \( g \)-values remains a challenge, as has been found in other studies.

3.4 The low frequency limit of the AC magnetic susceptibility

Frequency dependent magnetic fields are screened by mobile charges on the scales set by the skin depth, estimated as \( \sqrt{\frac{2}{\sigma \omega \mu}} \), where \( \sigma \) is the sample’s conductivity and \( \omega \) is frequency. For our samples, and in the up to 10 kHz frequency range used, this value is on
the order of a few millimeters, i.e., the field may be considered uniform inside the sample. Some residual frequency dependence of the diamagnetic background may be due partly to frequency dependence of the skin depth. However, the cusp is frequency independent in the same frequency range. This is illustrated in Figure 3-11b where all the data curves were shifted down to coincide with the 10 KHz data. Finite frequency magnetic response is necessarily complex, \( \chi(\omega) = \chi_R(\omega) + i\chi_I(\omega) \), with an in-phase (real) and an out-of-phase (imaginary) components, \( \chi_R \) and \( \chi_I \), respectively. Finite \( \chi_I \) signals dissipation (by eddy currents) and therefore must vanish, usually linearly in frequency. This is indeed the case as we checked explicitly. Its slope, \( \chi_I(\omega)/\omega \), is quantitatively consistent (also as a function of temperature) with eddy current heating, see Figure 3-12b. This dissipative component shows no sign of nonanalytic behavior as a function of magnetic field, see Figure 3-12a.

Figure 3-11 (a) Frequency dependence of the background and (b) independence of the singular contributions.
Figure 3.12 (a) The out-of-phase susceptibility component of $\chi(\omega)$ is routinely recorded simultaneously with the in-phase component (here in Sb2Te3). It is purely dissipative and regular in the vicinity of $H = 0$; it does not display the cuspy behavior. (b) Assuming the standard eddy current mechanism is responsible for dissipation, the out-of-phase component of $\chi(\omega)$ is proportional to conductivity, which is confirmed in our measurements: the observed value is consistent, up to geometric factors and closely follows the temperature dependence in-plane resistivity $\rho_{xx}$, according to the standard formula for power $P = \pi^2(h_0^2d^2f^2/2\rho_{xx}(T))$ dissipated during the AC excitation cycles. Here $f = \omega/2\pi$ and $d$ is the sample thickness.

The in-phase component is expected to provide a good estimate to the thermodynamic susceptibility, up to a small $\omega^2$ correction. This is indeed the case, see Figure 3.11a. Moreover, rather weak magnetoresistance implies negligible field dependence of the $\omega^2$ correction, i.e. only a simple vertical offset is sufficient to compensate for the $\omega^2$ correction, see Figure 3.11b. The singular cusp, however, persists and is frequency independent.
3.5 Phenomenological theory

Our experimental findings are notable not only in that singular response itself is detected but also that the same kind of response (with a remarkably consistent large magnitude) persists across a broad swath of samples and three distinct families of topological materials. The physical phenomenon that underlies the data must transcend the unavoidable variations in screening, chemistry, preparation and such, making it particularly unlikely that the bulk of the samples makes any significant contribution to the low field singularity (see also Figure 3-7b). Thus, our interpretation of the data aims squarely at the samples’ surfaces, where interplay of universal physics of helical surfaces and large scale disorder appear to capture some of the more salient aspects of physics.

Surface states of ideal topological insulators are described by a band of helical Dirac fermions, minimally characterized by a simple Hamiltonian which assumes a particularly symmetric form for the (00\bar{1}) cleavage surface

\[ H = \sum_{k,s,s'} (\hbar v_F \hat{n} \cdot \mathbf{k} \times \sigma_{ss'} - \mu \delta_{ss'}) c_{k,s}^\dagger c_{k,s'} \]  

(3-2)

here \( v_F \) is the Fermi velocity, \( \mu \) is the chemical potential, \( \sigma \)'s are the Pauli spin matrices, \( \hat{n} \) is the surface normal vector, and \( c (c^\dagger) \) are the creation and annihilation operators. For this particular surface the helicity parameter, \( |\mathbf{k} \times \sigma| \), is uniform in \( \mathbf{k} \)-space (which we take to be a disk with radius \( \Lambda \)) and the dispersion is circularly symmetric near the Dirac point (\( \mathbf{k} = 0 \)).

For states sufficiently far away from the Dirac point the hybridization with bulk bands and various warping effects become pronounced\(^{57,59,13} \), e.g. hexagonal warping, while the electronic structure on non-cleaving surfaces is likely to be characterized by
Figure 3-13 Temperature dependence predicted in Equation (3-6) is explored here by plotting traces at $T = 0$ (black), $T = 1$K (red) and $T = 10$K (green and red, see below). We compare the first plot to choose experimental parameters, i.e. $g = 60$, $\Lambda = 5 \cdot 10^8$ $^{-1}$, $v_F = 2000$ m/s, $x = 0.18 \cdot 1^{-3}$, also $L_z \approx 1$ mm is the measured thickness of these samples. These parameters are sample dependent—they can vary with sample growth and preparation, as well as with the level of surface reconstruction and/or ‘aging’\textsuperscript{55}. The value of $\Lambda$, the characteristic size of the momentum space of the surface states, is a few percent of the corresponding bulk quantity, its microscopic meaning remains to be established. Here, we pick it to be some fraction of the bulk Brillouin zone. Lastly, the Fermi velocity we use is consistent with that obtained from a low energy probe\textsuperscript{60}. If these parameters are rescaled upon raising temperature (from 1K to 10K), $\Lambda \rightarrow \Lambda$, $v_F \rightarrow 10v_F$, $x \rightarrow x/10$, $g \rightarrow 10g$, $x$ is invariant (red trace). Otherwise, in the green trace, we explore whether approximate invariance may be maintained under less fine-tuned rescaling, e.g. with $\Lambda \rightarrow \Lambda/2$, $v_F \rightarrow 2v_F$, $x \rightarrow 1.5x$, $g \rightarrow 3g$. This discussion is not intended as analysis of the experiments, but rather to explore potential thermal effects theoretically.
inhomogeneous helicity \(^{57}\). In this simple model these effects will enter implicitly as phenomenological parameters, such as the size of the \( \vec{k} \)-space unit cell \( \Lambda \). In what follows we will not be including orbital quantization effects \(^{52}\)—we expect these to be unimportant on general grounds, namely owing to the presence of disorder and to large \( g \)-factors in our samples, and they are indeed absent at low fields in our experiments. The Zeeman coupling is introduced via

\[
H_Z = \frac{g \mu_B B}{2} \sum_{k,s,s'} \hat{n} \cdot \sigma_{s,s'} c_{k,s}^\dagger c_{k,s'}
\]

It opens a gap near the Dirac point between conduction and valence bands, \( s = \pm 1 \), respectively

\[
\epsilon_{k,s} = s \sqrt{(g \mu_B B)^2 + (\hbar v_F k)^2} \equiv s \epsilon_k
\]

Quantitatively, we may gauge the relative importance of the orbital vs. Zeeman contribution by comparing gaps (Landau vs. Zeeman). While Landau quantization is expected to dominate at sufficiently low fields, the relevant field scale

\[
\hbar v_F \sqrt{\frac{eB}{h}} > g \mu_B B \Rightarrow B < \frac{e v_F^2 h}{(g \mu_B)^2} \lesssim 10^{-4}T
\]

is much lower than our experimental resolution (we have used large \( g \approx 60 \) and small \( v_F \approx 2000 \text{ m/s} \), which is an appropriate ballpark for our samples in the \( \mu = 0 \) patch regions).

The exact expression for the areal (sheet) susceptibility of a single 2D Dirac state

\[
\frac{\partial M}{\partial H} = -\mu_0 \frac{\partial^2 F}{\partial B^2}
\]

can be obtained, where
\[ M = -\frac{\partial}{\partial B} (E - k_B T S) = -(g\mu_B)^2B \sum_{s=\pm} \int \frac{d^2k}{(2\pi)^2} \frac{s}{\epsilon_k} \tanh \frac{\beta}{2}(s\epsilon_k - \mu) \]

\[ = -\frac{(g\mu_B)^2B}{\pi\beta h^2v_F^2} \sum_{s=\pm} \log \cosh \frac{\beta}{2}(sy - \mu) \bigg|_{y=\epsilon_0}^{y=\epsilon_A} \]

\[ \beta \to \infty, \mu = 0 \to \frac{(g\mu_B)^2B}{\pi\hbar^2v_F^2} \left( \sqrt{(g\mu_B B)^2 + (h\nu_F \Lambda)^2} - g\mu_B |B| \right) \] (3-6)

Weak magnetic field acts perturbatively as long as \( \mu \neq 0 \) in that spin-orbit locked electrons at the Fermi level polarize only slightly (far less than in the spin degenerate Fermi gas), hence \( \chi(B) \) is analytic as \( B \to 0 \). For very large \( B \) the response follows \(-1/|B|^3\) (as typical of Van-Vleck paramagnetism\(^5\)). The transition between these two behaviors takes place at \( B_C = \pm \mu/(g\mu_B) \) via a jump singularity in \( \chi_A \). The singular response at \( \mu = 0 \) descends from these singularities.

At \( k_B T = \mu = 0 \) the susceptibility reduces to

\[ \chi_A(B) = \frac{\mu_0(g\mu_B)^2[-2g\mu_B|B|\sqrt{(g\mu_B B)^2 + (h\nu_F \Lambda)^2} + 2(g\mu_B B)^2 + (h\nu_F \Lambda)^2]}{4\pi^2h^2v_F^2\sqrt{(g\mu_B B)^2 + (h\nu_F \Lambda)^2}} \]

which has the form of susceptibility data shown in Figure 3-2, Figure 3-4, Figure 3-7. In particular, the hallmark of Dirac physics is the universality of the slope of the \( \sim |B| \) term, which only depends on the \( g \)-factor and the Fermi velocity and not on the size of Brillouin zone, while the maximum of susceptibility \( \chi_A(0) \) at \( B = 0 \) depends on the details of warping and hybridization with the bulk through \( \Lambda \).

We now turn to a more quantitative exploration of this phenomenology, to establish the existence of reasonable choice of parameters that can reproduce the experimental results. As already discussed in the main text, we postulate the existence of regions with \( \mu \approx 0 \) that are sufficiently large so that this (nominally translationally invariant) theory
applies and the net response can be approximated as arithmetic average over various contributions from the bulk and surfaces, \( \chi(B) = \chi_0 + \chi_A(B)x/L_Z \). Most of the surface is significantly detuned from the Dirac point and only contributes to the (non-singular) background, as discussed above. Thus, we introduce one additional parameter, the corresponding surface fraction, \( x < 1 \), of \( \mu \cong 0 \) regions. Prevalence of singular response then points to ubiquity of such regions; however to actually locate these regions and elucidate the physics responsible for their formation is outside the scope of our Dirac phenomenology and will require further theoretical and experimental work.

The apparent susceptibility, without all non-singular background contributions, \( \delta\chi(B) = \chi_A(B)x/L_Z \), where \( L_Z \) is sample’s thickness.

### 3.6 Efficient cooling of Dirac fermions

Room temperature stability of the singular cusp is difficult to achieve within the confines of theory outlined in the previous section. In fact, the challenge is considerably more general—the thermal energy at \( T = 300 \) K is an appreciable fraction of the bulk gap of these narrowband materials. This results in significant temperature dependence of ‘background’ diamagnetic susceptibility, conductivity and other transport properties. We surmise that (Dirac) surface states responsible for the singularity are maintained at a different (significantly colder) temperature than the rest of the sample. We note a very recent study\(^{44}\) where uniquely slow, power-law in time, energy relaxation out of excited surface states in TIs was detected and attributed to acoustic-phonon-dominated coupling to the bulk. Such a weak coupling is a prerequisite for our proposal.
The basic idea is that a local fluctuation in doping, e.g. of the kind that may create a $\mu \sim 0$ patch at the surface (see above), will also create adjacent p and n doped regions. With the ac excitation field oriented along the c-axis we expect induced currents primarily flowing in the ab plane. For the experimental parameters, such as material’s conductivity and probe’s frequency range, the dissipative out-of-phase (eddy current) component dominates the in-phase component by about 3 orders of magnitude. These finite frequency currents are expected to exhibit inhomogeneities induced by variations in electronic structure. In particular, it is natural to find current loops traversing through the bulk, the surface and the said p-n regions. Under favorable conditions such current loops will act as mesoscopic Peltier coolers. These favorable conditions include relatively low local resistance (and smooth disorder) along the current path that helps focus the current flow through thermoelectrically asymmetric region and also direct contact between p and n regions that result in formation of a depletion layer and rectification path (diode shunt in the figure).

We now sketch out a simple and plausible, albeit speculative, scenario by which the AC nature of the probe itself, and, more specifically, eddy currents as observed in out-of-phase component of AC susceptibility, combined with subsurface disorder which essentially provides for proximate p and n regions, act to cool the surface electronic excitations responsible for the cusp well below the sample’s bulk temperature. While systematic studies of surface mesoscopics are needed to flesh out and test the various
aspects of this scenario, it is certain that some kind of a (non-equilibrium) cooling process is operative based on the energy scales argument above and but also on our additional experimental observations of slow equilibration, dissipative eddy current response and sizeable harmonic generation. The picture below is the simplest example of how the combination of known good thermoelectric properties of these materials, disorder morphology, simple semiconductor facts and ac nature of our probe may produce the sought after cooling behavior.

Unlike conventional (macroscopic) Peltier coolers, where \( p \) and \( n \) regions are well separated and no depletion layer forms, here we need direct contact if this ‘device’ is to operate on alternating currents source provided by eddy response (conventional Peltier cooler require DC power source). Such direct contact will form an effective rectifying element—a diode shunt—which will redirect the electric current away from the Peltier cooling path during the ‘wrong’ half of the cycle (when it would otherwise act as a heater). The detected 2\(^{\text{nd}}\) harmonic generation (shown, e.g., for Sb\(_2\)Te\(_3\) in panels (a) and (b) of Figure 3-15 and in Figure 3-9) is consistent with this scenario. Generation of second harmonic through rectification is illustrated by a simple calculation that shows that (Figure 3-15c) the presence of the 2\(^{\text{nd}}\) harmonic can originate from signal rectification (Figure 3-15d). Dissipative out-of-phase (eddy current) response, shown for Sb\(_2\)Te\(_3\) in Figure 3-12, is fully consistent with both the values and the temperature dependence of the in-plane resistivity \( \rho_{xx} \), giving a temperature dependent power dissipation. More elaborate configurations of \( p \) and \( n \) regions capable of simultaneous rectification and cooling may be imagined, of course. However, one particularly appealing aspect of this simplest ‘two-blob’
Figure 3-15 The nonlinearity of the surface-bulk connection is witnessed by the observed 2nd harmonic $\chi_2$.

(b) As expected, $\chi_2$ is quadratic in frequency. It is consistent with the existence of ‘rectifying’ paths in the putative thermoelectric cooling elements required for the cooling of small fraction of sample’s surface and thus suppressing thermalization of Dirac surfaces with the bulk. The generation of 2nd harmonic (c) is associated with (d) the rectified signal.

Maintaining the singular cusp response implies the need to keep relevant electrons at very low cryogenic temperature. This requires a highly effective cooling ‘device’, much more than what’s presently available on macroscales. Cooling efficiency can be enhanced
by a number of other effects, such as mesoscopic self-compatibility\textsuperscript{57} and nanoconstrictions\textsuperscript{54}. However, the observed unusually strong harmonic generation suggests that larger values of thermoelectric parameters may originate from strong frequency dependence of the transport coefficients under geometric confinement. For phonons such resonances are natural and well known\textsuperscript{56}. Recent considerations of the spin Seebeck effect show that in contrast with bulk Seebeck effect, the figure of merit of nanoscale thermal-spin conversion can be infinite, leading to the ideal Carnot efficiency\textsuperscript{55} (in the nonlinear spin Seebeck transport regime the system acts as a nanoscale thermal spin rectifier). Finally we note that thermopower is strongly affected by the spin-orbit coupling\textsuperscript{54}, with asymmetry provided by the nondegenerate spin channels, leading to much larger cooling enhancements on mesoscale.
Chapter 4

Surface superconductivity of Dirac puddles

4.1 Introduction

Unconventional superconductivity involving electron bands with nontrivial topological character has been predicted\cite{43,60-63} to arise on the surfaces of three dimensional (3D) topological insulators (TIs), where conduction channels host helical Dirac fermions\cite{37,38,64,65,20} that cannot be destroyed by non-magnetic scattering processes. Such topological superconductors are expected to host unusual Majorana modes\cite{43} with special non-Abelian statistics\cite{66} and lead to a paradigm change in approaches to quantum device engineering and materials science. The superconducting phases in topological insulators reported thus far, obtained e.g. by Cu doping\cite{67-69} or under very high (GPa range) pressures\cite{70,71}, were found to have increased bulk carrier densities\cite{68} and thus appear to be of bulk origin.

In this chapter we show that surface superconductivity in the topological material Sb$_2$Te$_3$ with hole-like bulk conduction (p-type) can be induced at a remarkably high temperature by a very minor change in Te vapor pressure during the crystal growth. Tellurium overpressure in a very narrow pressure range, while making no detectable
structural changes, supplies net free electrons to the system (see Section 4.4) so that in the superconducting state the carrier density relative to that in the non-superconducting state is strongly reduced, upshifting the Fermi energy from the bulk valence bands towards the vicinity of the Dirac point. Concurrently, the low-frequency diamagnetic screening in this new state is hugely and unprecedentedly enhanced.

In this nonmonotonically ‘doped’ state we observe a superconducting resistive transition at a relatively high temperature $T_{CR} \sim 9K$. Surprisingly, however, the onset of enhanced diamagnetic response, as probed by AC susceptometry and by SQUID magnetometry, begins well above $T_{CR}$. Furthermore, scanning tunneling spectroscopy (STS) indicates a patchy landscape of superconducting regions with gaps in the differential conductance that can be locally as large as $\sim 25$ meV. While the 9 K superconductivity discovered at ambient pressure in a topological material with reduced carrier density is unusual in its own right, observation of strong precursor effects at much higher temperatures points to an inchoate superconducting instability of topological surface states; it raises a possibility of even higher transition temperatures through controlled tuning of electron doping during materials’ growth.

4.2 Experimental results

Figure 4-1a shows resistivity of Sb$_2$Te$_3$ synthesized under $\sim 1.4$ MPa Te vapor pressure during the high-temperature step of the crystal growth cycle and measured at ambient pressure. The system undergoes a transition to zero resistance at the onset temperature $T_{CR} \cong 8.6$ K. We note that this is the highest $T_C$ observed in a topological material at ambient pressure after synthesis. In the narrow Te pressure range $1.2 < P < 1.5$ MPa, where
CHAPTER 4 SURFACE SUPERCONDUCTIVITY OF DIRAC PUDDLES

Figure 4-1 a, Resistivity of an exfoliated 100 nm thin Sb$_2$Te$_3$ crystal synthesized under ~ 1.4 MPa Te vapor pressure shows onset of transition to zero resistance at $T_{CR} \approx 8.6$ K. b, Diamagnetic susceptibility (left) at 1.9 K measured in a 0.2 T field shows huge diamagnetism (Meissner effect) only in the narrow Te vapor pressure range. It is anticorrelated with the measured hole density (right) which in the superconducting state is decreased by a factor of nearly 300. Outside this range samples are nonsuperconducting. c, Longitudinal sheet resistance $R_{xx}$ (red) showing a discernible downstep on decreasing temperature. This downstep is localized within ~ 5 K in temperature, as indicated by the derivative $dR_{xx}/dT$ (blue). d, Illustration of superconductive puddles (blue, size 2R) of Dirac bands, where pairing occurs at high temperature ($T_{CD}$), connected to the 2DEG metallic matrix (grey) which establishes a percolative path (dashed red line) at $T_{CR}$. The electronic Dirac dispersion and spin-orbit split dispersion of 2DEG are also sketched, see text. d,e,f, STM and STS scans were performed at 4.8 K in areas ~ 25 nm each, spaced ~ 200 nm apart. e, A typical topograph of a scanned area shows well-ordered hexagonal lattice, with differential conductance $dI/dV$ (from the average of 500 scans) shown in f. Depending on the scan area, the gaps 2$\Delta$ vary from zero to ~ 20 meV.
superconductivity is found. The bulk hole density obtained from Hall resistivity (see Section 4.7 and Table 4-2) is reduced by over two orders of magnitude (Figure 4-1b and Figure 4-7) to \( \lesssim 10^{18} \) cm\(^{-3}\)—a finding which should be contrasted with the increased electron density recorded in, e.g., superconducting Bi\(_2\)Se\(_3\) doped with Cu\(^{67}\).

In the same narrow pressure range the diamagnetism is enhanced by two orders of magnitude (Figure 4-1b) relative to the signal in samples outside the critical pressure range\(^{73}\) and displays a rich structure, of which the most revealing is the pronounced and sharp diamagnetic transition (Figure 4-3) in the vicinity of 50 K mirrored by a downstep in resistivity (Figure 4-1c)—both consistent with the onset of non-percolating superconductivity\(^{74,75}\). Scanning tunneling spectroscopy (STS) (Figure 4-1d,e and Section 4.6) shows very flat hexagonally ordered surfaces and a clear presence of gaps in differential conductance \( dI/dV \) with clear coherence peaks (Figure 4-1f) as well as more complex gap structures that can vary locally from 0 to \( \gtrsim 20 \) meV (Figure 4-6). This patchy gap landscape indicates a laterally inhomogeneous superconducting state which within the BCS theory (using the gap equation \( 2\Delta(0) = 3.5k_B T_C \), where \( k_B \) is the Boltzmann constant) has local transition temperatures that can be greater than 60 K, inline with the temperature where the onset of the enhanced diamagnetic signal is observed.

The resistive transition downshifts and broadens with increasing external magnetic field (Figure 4-2a). By tracking onset temperature \( T_{CR} \) we map the field-temperature \( H - T \) phase diagram for the two field orientations: applied along the \( c \)-axis of the crystal, \( H \parallel c \), and parallel to the \( ab \)-plane, \( H \parallel ab \) (upper and lower panels of Figure 4-2a, respectively). The low temperature critical field anisotropy \( \sim 1.4 \) at first glance is surprisingly small, given the large structural anisotropy of Sb\(_2\)Te\(_3\) (Figure 4-5). This may
Figure 4.2 a, Resistive transition temperature downshifts with increasing magnetic field applied (top) along the c-axis of the crystal, $H \parallel c$, and (bottom) parallel to the ab-plane, $H \parallel ab$. The onset of superconductivity is indicated by the arrows. b, $H \times T$ phase diagram of the superconducting state has relatively small zero-temperature anisotropy of ~ 1.3, the same as that of the square root of g-factor. The critical field values at $T \to 0$ agree with the paramagnetic (Zeeman) depairing field $H_p$. Inset: de Haas van Alfen quantum oscillations (dHvA) shown for $H \parallel c$ at 1.9 K persist below $H_p$ down to the temperature dependent onset field $H^*$. 

indicate that the limiting field obtained from the resistive onset may be of spin rather than orbital origin. Indeed, the anisotropy $H_p^\parallel /H_p^\perp \propto g_\perp /g_\parallel \equiv 1.4$ is consistent with the anisotropy of the g-factor detected by microwave spectroscopy as expected when depairing is of paramagnetic (Zeeman) origin. If depairing is of paramagnetic origin, a
simple relationship between the paramagnetic field (so called, Chandrasekhar-Clogston limit), \( H_p \), and the superconducting transition temperature \( T_C \) can be conventionally obtained\(^{76} \) by comparing magnetic energy \( \frac{1}{2} N(0) g^2 \mu_B^2 H_p^2(0) \) with the superconducting condensation energy at the transition \( \frac{1}{2} N(0) \Delta^2(0) \). Here \( N(0) \) is the density of states at the Fermi energy, \( g \) is the Landé \( g \)-factor, and \( \Delta(0) \) and \( H_p(0) \) are zero-temperature superconducting gap and upper critical field, respectively. Equating the observed low temperature values of \( H_{CR} \sim 3 - 4 \text{T} \) with \( H_p(0) \) and taking the large measured value\(^{23} \) of \( g \)-factor, \( g \approx 50 \), gives \( H_p(0) = 2.6 \frac{T_{CR}}{g} \approx 3.2 \text{T} = H_{CR}(0) \), from which the superconducting transition temperature \( T_C \) can be estimated. Surprisingly, this estimate does not correspond to the \( T_{CR} \) value observed in the resistivity, but to a much larger value of \( \sim 60 \text{K} \). This apparent inconsistency, as we will discuss later, can be resolved by considering superconducting correlations emerging at a temperature \( T_{CD} \gg T_{CR} \).

Another prominent feature in the \( H - T \) phase diagram is the large high-field low-temperature region in the superconducting state below the critical field \( H_p \) where we observe de Haas-van Aphen (dHvA) quantum oscillations (inset in Figure 4-2b and Figure 4-3a,b)—one of the most direct probes of quasiparticle excitations in metals\(^{78} \). This region, bound by the field \( H^*(T) \) at which oscillations first appear is obtained from the temperature dependence of the low-field onset of dHvA. Occurrence of dHvA oscillations is known in extreme type II superconductors\(^{79} \), although in conventional superconductors there is a considerable amplitude damping effect below upper critical field \( H_{c2} \). The superconducting \( \text{Sb}_2\text{Te}_3 \) shows no additional damping at the critical field (Figure 4-8). This suggests two possible reasons for robust quantum oscillations in the superconducting state: one is
Figure 4.3 a, dHvA oscillations show beats arising from two very close oscillation periods. The beats are a signature of spin-splitting by a strong spin-orbit interaction in 2DEG surface regions. b, dHvA oscillations vs. inverse transverse component of magnetic field $H_{\perp} = H \cos \theta$ shown for three values of magnetic tilt angle $\theta$. The observed constancy of beat nodes scaling with $H_{\perp}$ is a characteristic signature of 2D. c, Hole carrier density $n$ vs. Te pressure in the superconducting Sb$_2$Te$_3$ (in the narrow $\Delta P$ vicinity of ~ 1.4 MPa of Te pressure) and non-superconducting Sb$_2$Te$_3$ (outside $\Delta P$) states. In the superconducting region $n$ is reduced, bringing the chemical potential from the deep inside the bulk valence band to just below the Dirac point, as illustrated in sketches of the band structures in all three regions. The conductivity of the superconducting Sb$_2$Te$_3$ remains hole-like ($p$-type), with Fermi level crossing both, the Dirac bands and a 2DEG sliver of the valence band (upper middle sketch).
inhomogeneous superconductivity with some residual unpaired fermion quasiparticles\textsuperscript{72}, another is the superconducting gap developing nodes in momentum space\textsuperscript{79}.

Quantum dHvA oscillations (Figure 4-2b, Figure 4-3a,b and Section 4.7 ) are quite remarkable in their own right as they clearly display a beat structure found in two-dimensional quantum well states (2DEG) in semiconductors\textsuperscript{80}. Such states have been predicted in topological insulators (TIs) in ab initio calculations\textsuperscript{76} and detected in angle-resolved photoemission spectroscopy (ARPES)\textsuperscript{46} but, to the best of our knowledge, have never been observed in magneto-oscillations. Lifshitz-Kosevich\textsuperscript{78} analysis of these oscillations (see Section 4.8 , and Table 4-2) yields $v_F \approx 5.3 \times 10^5 \text{m/s}$, $l \approx 95\text{nm}$, $m = 0.065m_e$, $k_{F+} = 3.7 \times 10^8 \text{m}^{-1}$, $k_{F-} = 3.3 \times 10^8 \text{m}^{-1}$ for Fermi velocity, mean-free path, effective mass, and the two close Fermi wavevectors that induce the beats, respectively. We note that the overall beat structure closely scales with the magnetic field component transverse to the surface as shown in Figure 4-3b for three field orientations and therefore implies a two-dimensional origin of the signal. The peculiar dispersion of the bulk valence bands\textsuperscript{64} of Sb\textsubscript{2}Te\textsubscript{3} suggests the mechanism by which Te overpressure promotes the visibility of 2DEG (see Figure 4-3c)—in the narrow region of Te overpressure the hole density is severely reduced, bringing the Fermi level up to just below the Dirac point where it intersects the surface state and six small pockets of bulk bands which are expected to reconstruct to form 2DEG. In addition to Hall data (Section 4.7 , Figure 4-7 and Table 4-2), these dHvA oscillations provide an independent estimate of exceptionally high carrier mobility $\mu \sim 25,000 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$ in these samples compared to our non-superconducting samples (where $\mu$ is a factor of ~ 165 smaller and mean free path $l$ is a factor of 4 smaller, see Table 4-2), but also to the values in other TIs\textsuperscript{81-83}. 
Figure 4.4 a, Differential magnetic susceptibility vs. temperature for several values of magnetic field $\mu_0 || c$-axis. Meissner-like signal is ‘flat’ below $T_{CR}$ and monotonically vanishes at much higher (> 100 K) temperatures. $\chi$ was measured by applying a small ac excitation field $h_{ac} = 10^{-5}T$ at $f = 10$kHz. b, Full temperature and field dependence of $\chi$. c, $\chi$ vs. magnetic field $\mu_0 || c$-axis for a series of temperatures. The pronounced dHvA oscillations are apparent at 1.9 K. d, strongly depends on frequency. e, Full frequency and temperature dependence of $\chi$. Inset in g: Analysis of the frequency dependence of $\chi$ shows it to be quadratic in $\omega = 2\pi f$. The prefactor $b(T)$ in the $\omega^2$ term dominates the total variation of $\chi$ at finite frequencies. This variation is consistent with kinetic inductance of the patchy distributed 2DEG network discussed in the text.
Main panel: The zero frequency $\chi_0(T)$ obtained from the fits to $\chi(T) = \chi_0(T) + b(T)\omega^2$ shows a sharp onset at $T_{CD}$. d.c. magnetization $M(T)$ measured at 0.5 Oe using SQUID shows a sharp diamagnetic onset at the same $T_{CD} \approx 50$ K $> T_{CR}$. The data taken under zero-field-cooled (ZFC) and field-cooled (FC) conditions are essentially identical, consistent with negligible vortex pinning in 2D. We note that $M(T) \propto T_{CD} - T$ as well the onset of $\chi_0$ variation near $T_{CD}$ shown in (g) are both approximately linear in temperature, which within Ginzburg-Landau phenomenology (valid at high temperatures) would be related to magnetic penetration depth $\lambda$, $M(T) \propto 1/\lambda^2(T)$.

Differential diamagnetic magnetic susceptibility (Figs. 4a, b) measured using a relatively low (10 kHz) frequency is strongly field and temperature dependent. Large diamagnetism persists to high (> 100 K) temperatures as it smoothly and monotonically reduces to nonsuperconducting values. This unusual differential response is found to vary strongly with frequency of AC excitation even for the low (~ kHz) frequencies used (Figure 4-4d, e). General analyticity considerations at low external field dictate that low frequency corrections enter quadratically, $\chi(\omega, T) = \chi_0(T) + b(T)\omega^2$, and such behavior is found in resistively shunted Josephson networks and other cases of two fluid (normal + super fluid) dynamics. With this in mind we analyzed the frequency dependence of $\chi_0(T)$ by performing frequency scans at different temperatures, focusing on the low field regime. The data (inset in Figure 4-4g) clearly follows a parabolic frequency dependence.

The fit to this simple form uncovers a spectacular dichotomy between temperature variation of the zero frequency value $\chi_0(T)$ (Figure 4-4g and Section 4.10 ) and the dispersion coefficient $b(T)$ (Figure 4-4f): $\chi_0(T)$ shows a sharp diamagnetic onset in the vicinity of $\sim 55$ K, while the prefactor $b(T)$ smoothly marches toward the near null value at very high temperatures (see also Figure 4-11 and Section 4.10 ). This sharp variation of
\( \chi_0 \) is closely mirrored both in its saturation magnitude, \( \sim 10^{-4} \text{emu/cc} \) (Figure 4-4g), and in the onset temperature of 55 K in DC magnetization measured using SQUID (Figure 4-4h). The absolute value of diamagnetic response obtained in these samples clearly exceeds those of typical topological insulators (and semiconductors generally) by at least two orders of magnitude\(^{73,77} \) but also the largest known diamagnetic susceptibilities of semimetals\(^{85-87} \). Thus, we identify the pronounced variation (with temperature, field and/or frequency) and the magnitude of the diamagnetic response with a mesoscopic Meissner transition near 50 K detected in SQUID (Figure 4-4g). We observe that the finite frequency response coefficient, \( b(T) \), varies smoothly and persists to much higher temperatures, well above 100 K—this indicates a distribution of superconducting regions (‘puddles’). Significantly, both \( \chi_0 \) and \( b \) in relatively thick (~ mm size) samples exhibit a large field anisotropy (Figure 4-4g) consistent with the 2D nature of diamagnetic response and with the surface origin of the signal. Notably, the mesoscopic Meissner transition near 50 – 55 K is also mirrored in transport, whereby resistivity exhibits a clear downward step consistent with the appearance of finite-sized superconducting “shorts” inside the sample.

### 4.3 Discussion

The large separation between \( T_{CR} \) and \( T_{CD} \) signifies two separate physical processes at work. The patchy network of large local superconducting gaps detected by STS, strong diamagnetic response above \( T_{CR} \), and a characteristic downstep in resistivity are most naturally ascribed to isolated and well separated ‘puddles’ of local superconductivity, which we take to be congregating near the surface of the sample and very thin\(^{76} \), to account for the observed large orbital anisotropy of the response (Figure 4-4g, f), and also the
overall depletion of bulk conducting channels (Figure 4-1d, Figure 4-3c). It should be noted that puddles have been known to form at low carrier densities in other two-dimensional Dirac systems\(^{30}\) owing to nonlinear screening effects\(^{88,89}\). The relatively sharply defined \(T_{CD}\) implies that sufficiently many of these puddles are larger than the superconducting coherence length \(\xi\) (roughly estimated\(^{90}\) at below \(\sim 10\) nm) so that we may ignore size effects on the local order parameter. Such mesoscopic transition is expected to be accompanied by \(\sim 1/\lambda^2(T)\) growth of diamagnetic susceptibility\(^{84}\) (Figure 4-4g, h) and a downward step in \(R(T)\) (see Figure 4-1c).

The resistive transition at \(T_{CR}\) requires a mechanism for generation of Josephson (phase) coupling and establishment of global coherence of a sufficient fraction of these puddles. We propose that long-lived quasiparticles that give rise to the observed magneto oscillations also mediate Josephson coupling among puddles\(^{72,91}\). Such mediated couplings do exhibit spatial decay, which is dominated by exponential envelope with a temperature dependent length\(^{72}\). As long as the separation between puddles is larger than the quasiparticle mean-free path global coherence is attained once the typical puddle separation, \(a\), is comparable to metallic quasiparticles’ diffusion length \(a \approx L_T = \sqrt{\hbar D/k_BT} \approx \sqrt{\hbar v_F l/k_BT}\), where we have used the semiclassical expression for the diffusion constant \(D\) in terms of Fermi velocity \(v_F\) and mean free path \(l\), \(D \approx v_F l\). In this limit \(T_{CR} \approx hD/a^2k_B\) is largely controlled by the properties of the metallic matrix and the spatial arrangement of puddles, with the sizes of the puddles, the strength of local order parameter and other details (such as imperfect screening by the metallic matrix, also dimensionality of the metallic matrix) only modifying the criterion above through small (logarithmic) corrections\(^{72}\) (usually reductions) to \(T_{CR}\). Taking puddle distribution to be uniform, the
interpuddle spacing can be then estimated from the observed $T_{CR}$ provided the diffusion constant, $D$, of the intervening metal is known. Based on results of dHvA analysis we obtain the diffusion constant $D = v_F l/2 = 0.025 \text{ m}^2/\text{s}$, and the typical interpuddle separation $a \approx 140 \text{ nm}$. Using the actual sample area $A$ to estimate the total number of puddles, and relating the absolute value of the diamagnetic response to the single puddle’s response by assuming simply additive contributions of individual monodispersed puddles, we obtain the typical puddle size of about $R \approx 37 \text{ nm}$ (see Section 4.9). It is inline with the size scales of surface Dirac puddles reported in STS studies of nonsuperconducting TIs$^{31}$. 

The quantitative estimates above based on our transport and magnetic data point to a percolative puddle superconductivity$^{72,91,92}$ originating from Sb$_2$Te$_3$ surfaces. Together with the observed patchy distribution of superconducting gaps, we associate the superconductivity in Sb$_2$Te$_3$ at $T_{CD} \approx 50 \text{ K}$ with the assembly of Dirac puddles and identify the nature of the metallic matrix that mediates global superconductivity at $T_{CR}$ as 2DEG, the two dimensional electron gas$^{46}$. The identification of the matrix is consistent with (a) the 2D scaling nature of the beats in dHvA quantum oscillations (Figure 4-3b), and (b) with the estimate of spin-orbit splitting $\delta \approx 1.34 \text{ meV}$ obtained from the beats (see Section 4.8 and Figure 4-9) which is comparable to spin splitting observed in e.g., a 2DEG InGaAs/InAlAs heterostructures, also with strong spin orbit coupling$^{80}$. This suggests that the emergent superconducting state is potentially tunable through material’s control, e.g., of the quasiparticle mean free path, as well as the system’s Fermi velocity. Another reason is low bulk carrier density, which brings the Fermi level up to just below the Dirac point. Owing to a peculiar dispersion of the bulk valence bands$^{64}$ of Sb$_2$Te$_3$, the finite $k$
superconducting pairing spans the δk sliver that includes the combined system of puddles connected via coherent diffusion in the metallic 2DEG as visualized in Figure 4-1d.

The observed huge carrier mobility relative to nonsuperconducting Sb$_2$Te$_3$ and strong anisotropy of diamagnetism are indicative of spatially anisotropic, two-dimensional superconductivity. However, the nature of superconductivity in the puddles remains to be elucidated. For one, the weak anisotropy of the resistive limiting field implies that spin physics and spin-orbit coupling play a significant role in establishing the resistive $H$ vs. $T$ phase diagram via paramagnetic depairing mechanism$^{93}$. Also mesoscopic fluctuations in the effective Josephson network governing the resistive transition$^{72}$ must play a role in the observed variations in transport and in significant superconducting fluctuations at high temperatures. We remark that two-dimensional superconductivity that has been found in 2D electron gas at the interfaces between two band insulators, LaAlO$_3$ and SrTiO$_3$ occurs below 200 millikelvin$^{94}$, a much lower temperature than we observe; the much higher $T_{CR}$ in Sb$_2$Te$_3$ is a likely spillover from the superconducting puddles supporting pairing of helical Dirac holes.

4.4 Structural and elemental analysis

Transmission Electron Microscopy (TEM), X-ray diffraction and micro-Raman spectra of superconducting Sb$_2$Te$_3$ indicate absence of any structural changes induced by Te overpressure. We confirmed that the relatively low Te vapor pressure during the synthesis did not alter the layered rhombohedral van der Waals structure or lattice parameters of Sb$_2$Te$_3$ as determined from the X-ray diffraction (XRD) spectra (Figure 4-5).
Our XRD studies of the bulk lattice parameters in superconducting and nonsuperconducting Sb$_2$Te$_3$ did not detect any structural changes. We note that mesoscopic strain effects were reported in thin MBE-grown films of TIs$^{95,96}$, where complex (island-like) and rough surface morphology, on the scale of a nm or about 1 QL is often observed and we observe it as well in our MBE films$^{97}$. From the STM topography studies (see Section 4.5) we know that our exfoliated surface terraces are very flat (Figure 4-6), with

![Figure 4-5](image)

**Figure 4-5 (a)** X-ray diffraction spectrum for Sb$_2$Te$_3$ grown in a sealed quartz tube under Te vapor pressure $P = 1.42$ MPa collected in Panalytical diffractometer using Cu K$_\alpha$ ($\lambda = 1.5405\text{Å}$) line from Philips high intensity ceramic sealed tube (3 kW) X-ray source with a Soller slit (0.04 rad) incident and diffracted beam optics. Rietveld refinement lines shown in red are in full correspondence with the measured spectra. The obtained lattice parameters $a = b = 4.26$ Å and $c = 30.446$ Å remain unchanged in the 0.5 - 2 MPa pressure range, indicating that in this pressure range there is no structural change. (b) The robustness of the structure is also apparent in the identical micro-Raman spectra for superconducting (red) a non-superconducting (black) Sb$_2$Te$_3$. The E phonon modes are doubly degenerate modes in the ab-plane and the A modes are nondegenerate vibrations with atomic motion along the c-axis. The spectra were taken in ambient conditions in a backscattering geometry with linearly polarized excitation in the ab plane and normalized to the out of plane vibration at 70 cm$^{-1}$. 
vertical variation of about 20 pm, and with the lateral surface lattice parameter unaltered in the areas with and without the superconducting gap. We therefore conclude that intrinsic post-growth surface strains are not a factor.

Glow discharge mass spectrometry (GDMS) analysis of superconducting Sb$_2$Te$_3$ lists the impurity content in these crystals, see Table 4-1. The impurity content is the same

<table>
<thead>
<tr>
<th>Element</th>
<th>Concentration (ppm weight)</th>
<th>Concentration (weight ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li</td>
<td>&lt; 0.005</td>
<td>&lt; 5×10$^{-9}$</td>
</tr>
<tr>
<td>Al</td>
<td>&lt; 0.01</td>
<td>&lt; 1×10$^{-8}$</td>
</tr>
<tr>
<td>Ca</td>
<td>0.08</td>
<td>8×10$^{-8}$</td>
</tr>
<tr>
<td>Ti</td>
<td>&lt; 0.005</td>
<td>&lt; 5×10$^{-9}$</td>
</tr>
<tr>
<td>V</td>
<td>&lt; 0.005</td>
<td>&lt; 5×10$^{-9}$</td>
</tr>
<tr>
<td>Cu</td>
<td>0.4</td>
<td>4×10$^{-7}$</td>
</tr>
<tr>
<td>Zn</td>
<td>0.12</td>
<td>1.2×10$^{-7}$</td>
</tr>
<tr>
<td>Se</td>
<td>2.6</td>
<td>2.6×10$^{-6}$</td>
</tr>
<tr>
<td>Nb</td>
<td>&lt; 0.01</td>
<td>&lt; 1×10$^{-8}$</td>
</tr>
<tr>
<td>Mo</td>
<td>&lt; 0.05</td>
<td>&lt; 5×10$^{-8}$</td>
</tr>
<tr>
<td>Ru</td>
<td>&lt; 0.05</td>
<td>&lt; 5×10$^{-8}$</td>
</tr>
<tr>
<td>Pd</td>
<td>&lt; 0.01</td>
<td>&lt; 1×10$^{-8}$</td>
</tr>
<tr>
<td>In</td>
<td>Binder</td>
<td>Binder</td>
</tr>
<tr>
<td>Sn</td>
<td>&lt; 0.5</td>
<td>&lt; 5×10$^{-7}$</td>
</tr>
<tr>
<td>Sb</td>
<td>Matrix</td>
<td>Matrix</td>
</tr>
<tr>
<td>Te</td>
<td>Matrix</td>
<td>Matrix</td>
</tr>
<tr>
<td>I</td>
<td>&lt; 0.5</td>
<td>&lt; 5×10$^{-7}$</td>
</tr>
<tr>
<td>Ta</td>
<td>Source</td>
<td>Source</td>
</tr>
<tr>
<td>Hg</td>
<td>&lt; 0.1</td>
<td>&lt; 1×10$^{-7}$</td>
</tr>
<tr>
<td>Pb</td>
<td>0.63</td>
<td>&lt; 6.3×10$^{-7}$</td>
</tr>
</tbody>
</table>

*Table 4-1*
in the nonsuperconducting Sb$_2$Te$_3$. GDMS was performed by Evans Analytical Group (EAG). It has detection limits on the sub-ppm range for most elements that are nearly matrix-independent. In this technique, collisions between the gas-phase sample atoms and the plasma gas pass energy to the sample atoms, exciting the atoms. The atoms then lose their energy by emitting light with the atom specific wavelength. From the intensity of emitted light the atomic concentration can be determined. Sample atoms are also ionized through collisions that then are detected by mass spectrometry. The impurity content of all elements detected in Sb$_2$Te$_3$ matrix is less that a small fraction of ppm.

4.5 Te pressure during crystal growth and Te doping

Chemistry of crystal growth from the stoichiometric melt in a fixed volume at high (1000 °C) temperature sets the stoichiometry of the final crystal product. To finetune Te pressure for each crystal growth run we were changing the sealed tube volume while keeping the starting materials in stoichiometric ratio and the melt temperature fixed. The key physical consequence of Te overpressure in the 1.2 – 1.5 MPa pressure range was the reduced hole density and enhanced mobility. Since the topographic and structural integrity and robustness of crystalline surfaces are unaltered, these huge changes in transport give us important information regarding the contribution of surface conduction channels relative to the bulk. In the bulk there are charge carriers donated by intrinsic crystalline lattice defects such as vacancies and antisites. In Sb$_2$Te$_3$ vacancies on cation (Sb) or anion (Te) sites respectively act as electron donors and acceptors. From the detailed studies of these materials$^{98,99}$ we know that growth conditions using stoichiometric ratio starting materials result in slightly Te poor crystals, Sb$_2$Te$_{3-x}$ with $x \approx 0.03 – 0.05$. Also, native
concentrations of Sb antisites (contributing one hole each) and Te vacancies (contributing two electrons each) are about $10^{20}$ cm$^{-1}$ and $2 \times 10^{19}$ cm$^{-1}$ respectively, consistently accounting for the observed Hall number and other material properties$^{99,100}$. As Te pressure during the growth is increased, a fully stoichiometric (2:3) crystalline material is obtained in which bulk carrier density is the lowest. Under the growth conditions with Te pressure exceeding $\sim 1.5$ MPa, the process is reversed since as the system becomes slightly Te rich (or Sb poor), the antisite defects delivering holes to the bulk are less likely to form and the hole density is increased again. This change in stoichiometry is very slight and transport is the best way to detect it. We note that such fine-tuned non-monotonic doping is not unprecedented; it has been also observed in Sb$_2$Te$_3$ doped with Cr$^{99}$, TI films grown by MBE$^{101}$, and upon stoichiometry variation in Bi-Sb-Te system grown by zone melting$^{102}$.

The superconductivity is observed in the narrow range of Te vapor pressure in sealed quartz tubes during the crystal growth. The pressure range is quantified from the ideal gas law and is consistent with equilibrium binary phase diagram of Sb-Te$^{103}$. We made over 20 crystal growth runs under similar growth conditions and screened all samples, usually $\sim$ mm size, for their overall diamagnetic response using AC susceptibility first. We uncovered that the huge diamagnetic enhancement for the samples grown in the 1.4 – 1.5 MPa range was 100% robust and reproducible. SQUID measurements were subsequently performed on two samples measured directly. Seven samples were exfoliated and configured with van der Paw contacts for transport measurements, three of these exhibited zero field resistive transition with $T_c \sim 9$ K. Two were used for mapping the $H$ vs. $T$ phase diagram in Figure 4-2, two were used for a careful measurement of the high temperature feature near 50 K. STM studies were performed on five samples (see Section 4.6 ).
4.6 Superconducting gap mapping using scanning tunneling spectroscopy

To measure the local density of states of the samples, STS spectra were obtained in the clean areas identified topographically. A set of 500 spectra were obtained over each topographic area spaced approximately 200 nm apart using standard lock-in techniques. It was found that the STS spectrum was reasonably reproducible over each topographic area measured, but displayed dramatically different behaviors from region to region across various samples. In our measurements across all samples, we find these three types of spectra appearing with roughly equal probability. Unlike the case of the cuprates where sharp variations in the gap are seen over a short length scale (~ tens of angstroms), we do not find sharp variations in the spectra in space for these samples. The smooth variation in the spectra is likely due to the proximity effect between neighboring regions of the surface, especially since the surface is coupled to the bulk metallic states in the sample. Proximity coupling should become strongly attenuated on scales of diffusion length at higher temperatures but at 4.8 K where all these spectra were obtained our estimates of diffusion length based on dHvA data consistently predict negligible attenuation. The smooth and long-scale variations in the spectra make it challenging to obtain images of the evolution of the spectral lineshape (i.e. “gapmap”) in the sample due to field-of-view limitations in the microscope. Our best estimate of the length scale of the variation in the spectra is 100 – 200 nm, which is obtained from combining spectroscopic imaging in a given location with coarse motion of the microscope motors.

The slope in $dI/dV$ is also seen in good metals such as copper$^{104}$ or gold$^{105}$ where it
Figure 4-6 Scanning tunneling spectroscopy obtained on cleaved surfaces of superconducting Sb$_2$Te$_3$. (a) Typical STM topography (sample 1) showing atomically flat terraces of individual Sb$_2$Te$_3$ layers. Topographs of different areas (sample 2) taken approximately 200 nm away from each other display the triangular lattice of tellurium atoms on the surface, and sub-surface defects and dopants. While the topography in (c) and (d) is indistinguishable, the area in (c) displays the superconducting gap while the area in (d) does not. Differential conductance $dI/dV$ (V) was obtained in $25 \times 25 \text{ nm}^2$ areas of sample 2, at the locations where topographs were obtained. Spectra were obtained in 500 locations in each area. Individual spectra are shown in grey in each panel, and the average spectrum of each area is shown in red. (e) Spectra showing well-articulated coherence peaks and the gap $2\Delta \approx 10 \text{ meV}$ corresponding to $T_C \sim 30 \text{ K}$ estimated from BCS gap equation. (f) Spectra showing a more complex behavior with two gaps: a smaller one $2\Delta \approx 8 \text{ meV}$ and a larger one $2\Delta \approx 20 \text{ meV}$ corresponding to $T_C \sim 70 \text{ K}$. (g) Spectra showing gapless metallic behavior.
is due to band structure effects. The three types of spectra in (e) – (g) are observed with roughly equal probability across samples. Other samples grown in the Te pressure range where superconductivity is found show similar distribution of gap energies, consistent with superconducting puddles with average $T_{CD} \sim 50$ K embedded in the gapless matrix seen in area 3 in Figure 4-6. The variation in the articulation of coherence peaks as well as finite density of states (DOS) at low energies is commonly seen in strongly correlated systems such as high-$T_c$ cuprates$^{106}$ and heavy fermion superconductors$^{107}$ where it is a consequence of electronic inhomogeneity and strong Coulomb repulsion. It can also be detected when the Fermi energy is small as in the case of superconductivity at the LAO/STO interfaces$^{108}$.

4.7 Determination of carrier density from Hall resistivity and dHvA oscillations

Variation of carrier density with Te pressure (Figure 4-1b) is obtained by Hall effect and dHvA oscillations as shown here for two samples in Figure 4-7. We note that all Sb$_2$Te$_3$ crystals are $p$-type (the charge carriers are holes). The non-superconducting crystals synthesized at Te vapor pressures $P < 1.2$ MPa and $P > 1.55$ MPa have metallic-like temperature dependence ($dR/dT > 0$) of resistance and carrier densities $n \sim 2 \times 10^{19} - 10^{20}$ cm$^{-3}$. In the superconducting Sb$_2$Te$_3$ crystals ($P \sim 1.4$ MPa) $n$ is over an order of magnitude lower.
4.8 Lifshitz-Kosevich analysis and determination of spin-orbit splitting

The beating effect in dHvA oscillations implies the existence of two closely spaced frequency components with similar amplitudes (see Figure 4-3a). This has been seen in In$_x$Ga$_{1-x}$As/In$_{0.52}$Al$_{0.48}$As hetero-structures$^{80}$ where a single subband is spin-split by strong spin-orbit coupling. A spin-split Landau level gives rise to two closely spaced frequencies with similar amplitudes leading to a modulation of the dHvA amplitude given by $A \sim \cos \pi \nu$, where $\nu = \frac{\delta}{\hbar \omega_c}$, and $\delta$ is the energy separation between the spin-split Landau levels. Nodes in the beat pattern in dHvA will occur at half-integer values of $\nu$ ($\pm 0.5$, $\pm 1.5$, ...) where $A$ is zero. The total spin splitting $\delta$ can be expressed as $\delta = \delta_0 + \delta_1 \hbar \omega_c + \delta_2 (\hbar \omega_c)^2 + \cdots$, where $\delta_0$ is the zero field splitting, $\delta_1 \hbar \omega_c$ is the linear in field splitting, and $\omega_c = eB/\hbar m_c$. The higher order terms become significant at high fields. Figure 4-9a shows a plot of $\delta$ vs. $\hbar \omega_c$ for the fields corresponding to the nodes in dHvA oscillations in Figure 4-9b, using cyclotron mass $m_c = 0.065m_e$ obtained from the fit of oscillations to Lifshitz-Kosevich theory$^{58}$ $\frac{\Delta \sigma_{xx}(T)}{\Delta \sigma_{xx}(0)} = \frac{\lambda(T)}{\sinh(T)}$ in (c) using Monte Carlo technique. Here $\sigma_{xx}$ is the in-plane conductivity for magnetic field applied normal to the cleavage plane and $\lambda(T) = \frac{2\pi k_B T}{\hbar e B} m_c$. The extrapolation to zero field yields zero-field spin splitting $\delta_0 \approx 1.34$ meV, comparable to spin splitting found in other 2DEG heterostructures.
Figure 4-7 (a) Main panel: Hall resistivity of non-superconducting Sb$_2$Te$_3$ shows it to be p-type. Carrier density determined from Hall is $n = 3.9 \times 10^{20}$/cc. However, as determined from dHvA $n = 2.56 \times 10^{19}$/cc, is over an order of magnitude lower. Similar differences are found in other TIs when chemical potential is located deeply inside either valence or conduction bands. This is reconciled when taking into account the incoherent addition of contributions to Hall conductivity from hexagonal “pockets” (factor of 6), while the Fermi cross-sections in dHvA are sampled coherently. The additional factor of 2 comes from spin-splitting in the bulk bands, fully accounting for the differences in n. (b) Main panel: Hall resistivity of superconducting Sb$_2$Te$_3$ shows it also to be p-type. Here, however, carrier densities determined from Hall and dHvA (top out of) are identical $n = 1.4 \times 10^{18}$/cc, and over an order of magnitude lower than in the non-superconducting samples. This is consistent with the location of the Fermi level just below the Dirac point, and nearly on top of the valence band (Figure 4-3c).
Figure 4-8 Comparison of de Haas van Alfen (dHvA) oscillation amplitude damping for V$_3$Si and Sb$_2$Te$_3$. Plot shows field dependence of $D = \ln[\alpha \sinh(X) B^{1/2} T^{-1}]$ (known as Dingle plot) that shows the change in the dHvA oscillation amplitude upon crossing the superconducting limiting field. Here $\alpha(T, B)$ is the amplitude of quantum oscillations, $X = \frac{2\pi^2 k_B T}{\hbar \omega_c}$ and $\omega_c = eB/m_e$ is the cyclotron frequency. The estimated error of $D$ is 10%. The data shown for V$_3$Si are from Ref.$^{109}$ where the field scale was normalized to upper critical field $B_{c2}$. In conventional superconductor V$_3$Si there is an additional attenuation of the oscillation amplitude at the transition into the superconducting state observed in many extreme type II superconductors$^{110}$. Here we shifted the Dingle scale (in the same units) to overlay the data for both systems in their normal states (red line). The observed modulation of the Dingle factor $D$ in Sb$_2$Te$_3$ is a result of beats in dHvA.
Table 4-2

<table>
<thead>
<tr>
<th>S.C.</th>
<th>$n_{\text{LH}}$ (cm$^{-3}$)</th>
<th>$n_{\text{DNA}}$ (cm$^{-3}$)</th>
<th>$\mu_{\text{trans}}$ (cm$^2$V$^{-1}$s$^{-1}$)</th>
<th>$\mu_{\text{DNA}}$ (cm$^2$V$^{-1}$s$^{-1}$)</th>
<th>$l$ (nm)</th>
<th>$k_F$ (Å$^{-1}$)</th>
<th>$v_F$ (m/s)</th>
<th>$\tau_0$ (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4x10^{18}</td>
<td>1.4x10^{18}</td>
<td>24,966</td>
<td>26,600</td>
<td>95</td>
<td>0.0347</td>
<td>5.3x10^{5}</td>
<td>1.56x10^{-15}</td>
<td></td>
</tr>
<tr>
<td>Non S.C.</td>
<td>3.9x10^{19}</td>
<td>2.56x10^{19}</td>
<td>152</td>
<td>--</td>
<td>22.8</td>
<td>0.2279</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

4.9 Estimating interpuddle separation

Electronic parameters of Sb$_2$Te$_3$, such as cyclotron mass, carrier densities, or Fermi velocities are obtained from de Haas van Alfen (dHvA) oscillations using Lifshitz-Kosevich-theory, see Figure 4-9 and Table 4-2. The mobility was calculated both from Hall data and in the standard way from Dingle analysis of dHvA quantum oscillations from the Dingle plot in Figure 4-8.

Based on the material parameters in Table 4-2, we obtain the diffusion constant $D = v_F l/2 = 0.025$ m$^2$/s, and the typical interpuddle separation $a \approx 140$ nm. From the area $A$ of the sample $A = 12.6$ mm$^2$ we estimate the total number of puddles to be on the order $n \approx A/a^2 \approx 6 \times 10^8$. Next we compare the absolute value of the diamagnetic response and relate it to the single puddle’s response by assuming simply additive contributions of individual monodispersed puddles, $V \chi_0 \approx n \chi_1$, where $V$ is the sample volume and $\chi_1$ is the average extensive susceptibility of one typical puddle. Using $V \approx 2.5$ mm$^3$ we obtain $\chi_1 \approx 2 \times 10^{-22}$ at low temperatures. Relating this value to puddles’ dimensions, e.g., radius $R$ and thickness $t$, is complicated at high temperatures without
an independent measurement of the local penetration depth, $\lambda$. However, at low temperatures and for sufficiently large puddles we assume field exclusion which significantly simplifies the analysis, yielding $\chi_1 = -4R^3$. Independence of this expression of the thickness $t$ is due to large demagnetization correction.$^{84}$ The estimated puddle size of about $R \approx 37$ nm is comparable to the size scales of surface Dirac puddles observed by scanning tunneling microscopy.$^{31}$

Figure 4-9 (a) Determination of spin-splitting due to s pin-orbit coupling from (b) the beats in de Haas van Alfen quantum oscillations in superconducting $\text{Sb}_2\text{Te}_3$. (c) Determination of cyclotron mass from the temperature dependence of dHvA oscillations in (b). (d) The dHvA beats in two different $\text{Sb}_2\text{Te}_3$. While some oscillation amplitude variations are present, the beat pattern is robust.
4.10 Frequency and temperature dependence in the inductive linear response

Complex susceptibility, $\chi(\omega, T) = \chi'(\omega, T) + i\chi''(\omega, T)$, is characterized by reactive inphase response, $\chi'$ (denoted $\chi$ in previous sections), but also has a dissipative out-of-phase component $\chi''$. We can readily identify the dissipative component with eddy conductivity (Figure 4-10). Classical current fluctuations can only screen magnetic fields.

![Figure 4-10](image)

Figure 4-10 (a) The in-phase component of $\chi$ is quadratic in frequency $\chi(T) = \chi_0(T) + b(T)\omega^2$. (b) The dissipative (out-of-phase) component of $\chi(\omega)$ is strictly frequency linear, as expected. (c) The standard eddy current mechanism is responsible for dissipation dominated by the bulk, the out-of-phase component $\chi''$ is proportional to conductivity: the observed value is consistent, up to geometric factors and closely follows the temperature dependence in-plane resistivity $\rho_{xx}$ of the bulk, according to the standard formula for power $P = \pi(h/2\rho_{xx})^2 f^2/2\rho_{xx}(T)$ dissipated during the AC excitation cycles. Here $f = \omega/2\pi$ and $d$ is the sample thickness.
Figure 4-11 (a) A sharp diamagnetic transition at \( \sim 50K \) in the zero frequency response \( \chi_0(T) \) obtained from the fits to \( \chi(T) = \chi_0(T) + b(T)\omega^2 \) of another \( \text{Sb}_2\text{Te}_3 \) crystal. (b) The prefactor \( b(T) \) in the \( \omega^2 \) term is varying smoothly, dominating the total variation of \( \chi \) at finite frequencies. This variation is consistent with kinetic inductance of the patchy distributed 2DEG network.

at high frequencies, with \( \chi'(\omega \to \infty) \to -1/4\pi \) and \( \chi'(\omega \to 0) \approx - (\omega \tau)^2 \), where \( \tau \) is the characteristic relaxation time. For example, modeling a conductor as a simple LR-circuit we find \( \tau = L/R \). Proper interpretation of finite value of \( \chi'(0) < 0 \) requires quantum mechanics. One may, however, use purely classical phenomenology to capturediamagnetism by positing existence of macroscopic “perfect inductor” paths,
Figure 4-12 Longitudinal sheet resistance $R_{xx}$ (red) on decreasing temperature in (a) 0 T field, (b) 5 T, and (c) 14 T fields. At zero field this downstep is localized within $\sim 5$ K in temperature, as indicated by the derivative $dR_{xx}/dT$ (blue). Consistent with superconductivity, this downstep is smaller and broader at 5 T and is not detected at 14 T.

with $\tau \to \infty$. For superconductors we may simply think of perfect inductors as linearized Josephson elements and there exists abundant literature on phases and phase transitions of resistively shunted Josephson networks$^{111}$.

Roughly speaking, the normal phase can be thought of in a coarse grained fashion as a macroscopic $LR$ circuit, while the superconductor has $R \to 0$. While this ‘1-loop’ phenomenology correctly captures asymptotic $\omega \to 0$ limit of the two phases, it misses the low frequency correction in the superconducting phase, both $\sim \omega^2$ in $\chi'$, and, importantly
$\chi'' \sim \omega$, which may be thought of as the response from a finite normal fluid fraction. Various simple ‘2-loop’ improvements are possible to rectify this situation, e.g., two inductors in parallel (and only one perfect) have a simply additive response, i.e. 

$$\chi'(\omega \to 0) = \chi_0 + b\omega^2 + \cdots,$$

with $\chi_0$ coming from perfect inductors and $b$ representative of resistively shunted elements. We have used this phenomenology to extract $\chi_0$ in two different samples, see Figure 4-4g, f and Figure 4-11.

### 4.11 Resistive signature at 50 K consistent with superconducting puddles

Simple effective medium treatment of electrical conduction combined with mean field description of superconductivity in finite sized puddles predicts a discontinuous drop (reduction) in resistance of the sample, reflective of the shortening of resistive paths. Our zero-field data in Figure 4-12 is consistent with that expectation. The overall broadening ($\sim 50$ K) is consistent with some residual distribution in the $T_{CD}$, which continues to broaden with increased field. The feature is entirely absent by 14T.
Chapter 5

Using electron irradiation to achieve surface transport in topological insulators

5.1 Introduction

Unconventional quantum matter can be easily hidden within the rich existing library of condensed matter and to recognize its existence time and again novel concepts have to be invoked. It also takes a profound understanding of the real material constraints that can prevent the unconventional properties from being detected. Three-dimensional (3D) topological insulators are a spectacular example of this—narrow-band semiconductors well known for their high performance as thermoelectrics they were discovered to support unusual gapless robust two-dimensional (2D) surface states that are fully spin-polarized with Dirac-type linear electronic energy-momentum dispersion, which makes them protected against backscattering by disorder. These materials have narrow bulk gaps (~200-300 meV) and have carriers donated by intrinsic crystalline lattice defects such as vacancies and antisites. As a result, the conduction through the bulk and its intermixing with surface channels is what largely denies direct access to surface charge
transport required for the implementation in spin-based nanoelectronics\textsuperscript{114} or fault-tolerant topological quantum computing\textsuperscript{43}.

Extensive attempts to reduce the contribution of bulk carriers, involving techniques such as nanostructured synthesis/growth\textsuperscript{24,28,115}, chemical doping\textsuperscript{22} or compositional tuning\textsuperscript{116}, have relied on electrostatic gating of micro- or nano-structures comprising tens of nanometer thin films\textsuperscript{27} or similarly thin exfoliated crystals\textsuperscript{83} to gain less ambiguous access to the surface states.

\section*{5.2 Achieving surface quantum transport using high energy electron irradiation}

In this chapter we demonstrate that bulk conductivity in topological insulators (TIs) can be decreased by orders of magnitude to charge neutrality point on a large (depth) scale by the controlled use of electron beams which for energies below \(~3\) MeV are known to produce a well-defined, stable and uniform spread of Frenkel (vacancy-interstitial) pairs\textsuperscript{117} within their penetration range of hundreds of microns (Figure 5-1a,b). The combined effect of these pairs is to compensate for the intrinsic charged defects responsible for the conductivity of the bulk while crystal lattice integrity is maintained (see Figure 5-1c) and robust topological surfaces are unaffected. Stable high mobility surface conduction channels are achieved when a sufficient irradiation dose is followed by the optimally engineered annealing protocol; it is a proof-of-principle demonstration of practiced semiconductor processing applied to solving one of the key limitations of bulk TIs.

Under exposure to light particles such as electrons the production of vacancy-interstitial point defects\textsuperscript{118} in solids known as Frenkel pairs is well established. The pair
Figure 5-1  a, Energetic electron beams can penetrate solids to a depth of many tens of microns. Electron irradiation affects the bulk but not the robust topological surfaces. b, Impinging electrons induce formation of Frenkel vacancy-interstitial pairs (inset), which act to compensate the intrinsic bulk defects. Main panel: Calculated crosssections $\sigma$ for Frenkel pair production in Bi, Te and Se sublattices as a function of electron energy $E$, assuming displacement energy $\sim 25$ eV. Energy thresholds, $E_{th}$ in $\sigma$ are set by atomic weight; choosing $E < E_{th}^{Bi}$ or $E > E_{th}^{Bi}$ allows to tune Fermi level in both $p$- and $n$-type TIs. c, Transmission electron microscopy image of Bi$_2$Te$_3$ with electron dose $\phi = 1$ C/cm$^2$; the atomic displacements of $\sim 1$ per 5,000 are not seen.

formation has an energy threshold which scales with atomic weight—it is high for heavy Bi and is lower for lighter Te and Se, and the effective cross-section $\sigma$ for pair creation on different sublattices in materials containing these elements depends on the projectile energy (Figure 5-1b). This defines a pair-production window: below these thresholds the given sub-lattice becomes immune to irradiation while at energies above $\sim 3$ MeV clustering and
even nuclear reactions\textsuperscript{119} may occur. The process of Frenkel pair formation is straightforward and can be well controlled. The charge is primarily delivered by introduced vacancies, since at room temperature interstitials do not contribute owing to their much lower (compared to vacancies) migration barriers. For electron beams energies above \( \sim 1.5 \) MeV the effective cross-section \( \sigma \) on a Bi sublattice is the highest and we will show that donor type defects on this sublattice do prevail. For energies below Bi threshold (\( \sim 1.2 \) MeV) the creation of Frenkel pairs will be mostly on Se or Te sublattices and a correspondingly different net charge doping is expected.

Generally, crystal growth of tetradyminate crystals such as \( \text{Bi}_2\text{Te}_3 \) and \( \text{Bi}_2\text{Se}_3 \) results in equilibrium defect configurations comprising vacancies and antisites on both sublattices\textsuperscript{113,120}. The net free charge balance that delivers carriers to bulk conduction or valence bands from these defects can be varied by growth conditions or doping\textsuperscript{22}. In undoped \( \text{Bi}_2\text{Se}_3 \), where Se vacancies are presumed to dominate, the net conduction is by electron carriers or \( n \)-type. In \( \text{Bi}_2\text{Te}_3 \), where antisites are prevalent\textsuperscript{98} the conductivity is usually \( p \)-type, namely by hole carriers, although by varying stoichiometry it has been grown of either conductivity type. Without a priori knowledge of the net donor or acceptor flavor of the pairs on different sublattices we have chosen to irradiate \( p \)-type topological materials \( \text{Bi}_2\text{Te}_3 \) and Ca-doped\textsuperscript{22} \( \text{Bi}_2\text{Se}_3 \) with 2.5 MeV electron beams which create Frenkel pairs on all sublattices according to \( \sigma(E) \) in Figure 5-1b.

Longitudinal resistivity \( \rho_{xx} \) of the initially \( p \)-type \( \text{Bi}_2\text{Te}_3 \) measured at 20 K \textit{in situ} in the irradiation chamber as a function of electron irradiation dose is shown in Figure 5-2. All irradiations were performed with samples kept at 20 K, the temperature of liquid hydrogen, below the mobility threshold of the interstitials which tend to be more mobile
Figure 5-2 Resistivity of p-type Bi$_2$Te$_3$ irradiated with 2.5 MeV electrons vs. dose $\phi$ (red squares) measured in situ at 20 K shows about three orders of magnitude increase at the charge neutrality point (CNP) where the conduction is converted from p- to n-type, moving the Fermi level $E_F$ across the Dirac point (see cartoon). Cycling to room temperature reverses the process, which can be recovered by further irradiation (blue circles) and stabilized.

than vacancies\textsuperscript{118}—this ensures the stability of all charges introduced by the irradiation process. Most immediately notable features in the figure are (i) a nearly three orders of magnitude resistivity increase to a maximum $\rho_{xx}^{\text{max}}$ and (ii) the observed ambipolar conduction as a function of irradiation dose with well-distinguished $p$ (hole) and $n$ (electron) conduction regions. The resistivity maximum $\rho_{xx}^{\text{max}}$ (or conversely, the conductivity minimum $\sigma_{xx}^{\text{min}}$) is at the charge neutrality point where conduction is converted from $p$- to $n$-type, as determined from Hall resistivity (see Figure 5-3). The same type conversion is observed in Ca-doped Bi$_2$Se$_3$ (Figure 5-8) which\textsuperscript{22} is also $p$-type. As long as the terminal
irradiation dose $\phi$ is relatively low, below $\sim 0.1$ C/cm$^2$, $\rho_{xx}(\phi)$ traces its shape upon temperature cycling to room temperature and back to 20 K, with $\rho_{xx}^{\text{max}}$ reproduced by the next irradiation cycle. Here it is apparent that the value of $\phi_{\text{max}}$ at the charge neutrality point is not universal; it depends on the starting free carrier concentration $n_b$ but can be straightforwardly scaled using universal slope $\approx \partial n_b / \partial \phi$ of quasi-linear variation of $n_b$ vs. dose (Figure 5-3a).

*Ex situ* measured resistivities of Bi$_2$Te$_3$ crystals irradiated to different terminal doses and taken to room temperature before the chill-down to 4.2 K are in full correspondence with the *in situ* results. Figure 5-3a shows three orders of magnitude increase in $\rho_{xx}$ to the maximum, $\rho_{xx}^{\text{max}}$, at $\phi_{\text{max}} \approx 90$ mC/cm$^2$ higher than the in situ $\phi_{\text{max}}$ (Figure 5-2) likely owing in part to some defect migration (mostly interstitials$^{118}$) above 100 K. The $p$- to $n$-type conversion is clearly seen in Hall resistance $R_{xy}$ (Figure 5-2b and Figure 5-3 f-h) flipping its slope $d R_{xy} / d H$ and Hall confident $R_H = -1 / n_b e$ changing sign in the conversion region. Near charge neutrality point (CNP) the net residual bulk carrier density $n_b = n_D - n_A$ is very low ($n_D$ and $n_A$ are concentrations of donors and acceptors respectively). In this region local charge fluctuations can be very large creating inhomogeneity akin to a network of puddle-like $p$-$n$ junctions that nonlinearly screen random potential on very long length scales$^{89}$ and a simple estimate of free carrier density from $R_H$ is no longer appropriate$^{83}$. Remarkably, outside the immediate vicinity of CNP effective carrier mobilities $\mu$ estimated from $\sigma_{xx} = e n_b \mu$ in the Drude model are not much affected by the irradiation process, confirming that scattering events by Frenkel-pair point defects are scarce and the main effect is charge compensation. Mobilities in our crystals
are significantly higher ($\mu \approx 7,000 - 11,000 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$, see Table 5-1 on page 95) than a commonly observed $\leq 1,000 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$ range near the CNP region.

Figure 5-3  a. Bi$_2$Te$_3$ crystals are irradiated to different terminal doses and measured at 4.2 K.  b, Hall resistance.  c,d,e. The SdH oscillations in longitudinal resistance and Hall resistance.  f,g,h, low-field $R_{xy}$ and the Fermi level for the samples in c, d, and e.  i. The Landau level index plot vs. filed minima in the SdH oscillations.  j. A cartoon of a TI with conducting bulk before irradiation (top) and after, with ideally only topological surface conducting (bottom).
Figure 5-4 a. Evolution of longitudinal resistivity $\rho_{xx}$ measured at 4.2 K after cycling to room temperature (RT) for a crystal irradiated with dose $\phi = 90$ mC/cm$^2$. Each RT dwell time is coded with a different color. Resistivity is seen to cross charge neutrality point (CNP) in reverse from $n$-type back to $p$-type. It is consistent with slow migration (hundreds of hours at RT) of vacancies (in accord with $\sim 0.8$ eV migration barriers, see text) and shows that CNP can be reached by designing a suitable thermal protocol. Insets show $\rho_{xx}(T)$ for $n$-type region (upper left) and $p$-type region (lower right).

b. Change in magnetoconductance (MC) at different RT dwell times; here MC evolves from a quadratic field dependence of a typical bulk metal at short RT dwell times, through a complex region dominated by the charge-inhomogeneous bulk, to a weak antilocalization (WAL) region showing the characteristic low-field cusp near CNP. Here the data were normalized to the value at zero field. A fit to 2D localization theory is shown as red line.

The change in net carrier density induced in the irradiation-tuned process is reflected in the observed Shubnikov-de Haas (SdH) quantum oscillations (Figure 5-3 c-e) of sheet and Hall resistances, $\Delta R_{xx}$ and $\Delta R_{xy}$, from which an estimate of the Fermi surface size can be obtained (Table 5-1). The Fermi vector in pristine crystals $k_F \approx 0.025\text{Å}^{-1}$ is consistently larger than after irradiation and, for example, the dose $\phi \approx 89$ mC/cm$^2$ which
tunes the crystal close to CNP results in $k_F \approx 0.014 \text{Å}^{-1}$. The corresponding net carrier densities are $n_b = 5.06 \times 10^{17} \text{cm}^{-3}$ and $n_b = 1.08 \times 10^{17} \text{cm}^{-3}$ respectively. At higher irradiation dose $\phi \approx 99 \text{mC/cm}^2$, with $k_F \approx 0.022 \text{Å}^{-1}$ Fermi surface size becomes comparable to the pristine material. A comparison of these numbers with carrier densities obtained independently from Hall data (Table 5-1) reveals a good agreement between the two techniques near CNP but about an order of magnitude higher estimate of $n_b$ by the latter well outside the CNP region. Similar differences are found in other TIs when chemical potential is located deeply inside either valence or conduction bands\textsuperscript{121}.

The stability of net carrier density crucially depends on the terminal irradiation dose. First we illustrate that for low terminal electron doses ($\phi \lesssim 0.1 \text{C/cm}^2$) the resistivity is not stable as it evolves from the $n$-region after irradiation through CNP back into the $p$-region. The experiment was to simply change the dwelling time at room temperature (RT) to allow for vacancies to diffuse. The result clearly shows (Figure 5-4a) the reverse conversion from metallic-like resistivity on the $n$-side just after irradiation, through insulating resistivity at CNP (after nearly 250 hours at RT), and back to a weakly semiconducting-like resistivity on the $p$-side. As the system evolves across CNP a weak antilocalization quantum interference correction\textsuperscript{122} to classical magnetoresistance (MR) emerges in the non-equilibrated conversion region (Figure 5-4b), with a complex in-field MR structure possibly reflecting two carrier types present. From isochronal annealing experiments (Figure 5-9, Section 5.7) we estimate energy barriers controlling defect migration to be $\sim 0.8 \text{ eV}$ corresponding to $\sim 9000 \text{ K}$, in line with literature values for vacancy migration in solids\textsuperscript{118}. We conclude that at low terminal doses while defect migration is sluggish, the equilibrium is not attained.
Figure 5-5 a, Sheet resistance $R_{xx}$ vs. temperature of Bi$_2$Te$_3$ crystal at the charge neutrality point (red line) exhibits a plateau at low temperatures - a thumbprint of 2D surface conduction, see right panels in (c) and (d). Magnetic field breaks time reversal symmetry and gaps Dirac bands, resulting in localizing behavior shown as dash. Inset: Optical image of the crystal showing van der Pauw contact configuration used. b, Annealing protocol with time steps $\Delta t = 30$ min implemented to tune Bi$_2$Te$_3$ crystal with dose 1 C/cm$^2$ back to stable CNP. Inset: Magnetoresistance at 1.9 K after each annealing step, with colors matched to indicate different annealing temperatures. c, Left: WAL low-field quantum interference correction to the linear-in-field magnetotransport (in the right panel) at CNP in a Bi$_2$Te$_3$ crystal at 1.9 K with its characteristic low-field cusp. The 2D character of WAL is evident in its scaling with transverse field $H_{\perp} = H \cos \theta$, where $\theta$ is the tilt angle of the field measured from sample’s c-axis. A fit to 2D localization (HLN) theory (solid line) confirms that the contribution is only from two surfaces and yields a dephasing $\Phi = \hbar/(4e\ell^2) = 0.01$ T. Right: Linear magnetoresistance at CNP shows 2D scaling with $H_{\perp}$. d, Left: WAL contribution at CNP in a Bi$_3$Se$_2$:Ca(0.09%) crystal at 1.9 K also scales with $H_{\perp}$. At high fields outside the cusp, the scaling is seen to fail for $\theta \geq 60^\circ$. A fit to HLN theory (solid line) again confirms the contribution only from two surfaces and
CHAPTER 5 USING ELECTRON IRRADIATION

yields a smaller dephasing field $B_\phi \sim 0.004$ T (and almost twice as long dephasing length $l_\phi \sim 220$ nm). Right: Linear magnetoresistance at CNP also scales with $H_\perp$.

Next we demonstrate that stability can be controlled and optimized in a ‘reverse’ conversion process by the annealing protocol when the terminal electron dose is sufficiently high. This is the key part of the material modification process by electron irradiation; without it surface transport via electron irradiation could not be obtained. We demonstrate this for the samples exposed to the dose ten times higher, $\phi \cong 1$ C/cm$^2$. Figure 5-5a shows temperature dependence of sheet resistance $R_{xx}$ and conductance $G_{xx}$ of Bi$_2$Te$_3$ tuned back to CNP through a thermal protocol shown in Figure 5-5b, where it remained for months of testing. $R_{xx}(T)$ increases exponentially as the temperature decreases, turning below $\sim 200$ K into variable range hopping (Section 5.8, Figure 5-5). This bulk behavior is cut off at low temperatures when the contribution from the surfaces becomes comparable and a temperature-independent surface transport with minimum conductance$^{83} G_{xx}^{\text{min}} \cong 20$ $e^2/h$ reveals its quantum nature. The 2D character of this region is witnessed by MR that depends only on the transverse component of magnetic field $H_\perp = H \cos \theta$ (right panels in Figure 5-5 c and d). Unlike the quadratic-in-field MR of a conventional metal, the large ($\gtrsim 10\%$) magnetoresistance at CNP (Figure 5-5a) is found to be linear in field (Figure 5-5 c, d).

The topological protection of the surface states near CNP can be tested in at least three ways. One, by breaking time reversal symmetry (TRS) which should gap out Dirac states. Another, by the appearance of topological Berry phase$^{43}$ of $\pi$. And yet another, by detecting two-dimensional (2D) weak antilocalization (WAL) of Dirac particles traveling
through time reversed paths\textsuperscript{122} associated with this Berry phase. Figure 5-5a shows that applying TRS-breaking magnetic field indeed causes $R_{xx}(T)$ at the lowest temperatures to upturn, showing localizing behavior consistent with opening of the Dirac gap\textsuperscript{123}. Berry phase $\varphi_B = 2\pi \beta$ can be estimated from SdH oscillations using a semiclassical description\textsuperscript{124} $R_{xx} = A_{SdH} \cos \left[ 2\pi \left( \frac{\hbar \nu}{H} + \frac{1}{2} + \beta \right) \right]$, where $A_{SdH}$ is the oscillation amplitude, $H$ is the frequency in $1/H$, and $\beta$ is the Berry factor. Near CNP the obtained Berry factor is $\beta = 0.5 \pm 0.06$ as expected for the topological Dirac particles, while outside the CNP region a trivial value of $\beta = 0$ is obtained (Figure 5-3i). From this we conclude that near CNP, even with the complex Bi$_2$Te$_3$ band structure, topological conduction channels dominate.

Weak antilocalization (WAL) quantum correction to ‘classical’ conductivity is small and is not detected in TIs where bulk conduction is appreciable and Fermi level is pinned within the bulk bands. As we approach CNP, however, WAL emerges and clearly articulates at low magnetic fields as a characteristic positive magnetoresistance cusp (inset in Figure 5-5b). In close proximity to CNP the corresponding cusp in negative magnetoconductance scales with the transverse field $H_\perp = H \cos \theta$ (Figure 5-5 c, d), confirming its 2D character. In this case, the number $n_Q$ of quantum conduction channels contributing to WAL can be estimated from 2D localization theory\textsuperscript{34}

\[
\Delta G \simeq \alpha \frac{e^2}{2\pi^2 \hbar} f \left( \frac{B_\phi}{B} \right)
\]  \hspace{1cm} (5-1)

Where $\Delta G(B)$ is the low-field quantum correction to 2D magnetconductance, coefficient $\alpha = n_Q/2$ equals to $1/2$ for a single 2D channel, $f(x) \equiv \ln x - \psi(\frac{1}{2} + x)$, $\psi$ is the digamma function, and field $B_\phi = \frac{\hbar}{4e l_\phi}$ is related to the dephasing length $l_\phi$ of interfering
CHAPTER 5 USING ELECTRON IRRADIATION

electron paths. In Bi\textsubscript{2}Te\textsubscript{3} at CNP the fie (in Figure 5-5c) yields \( \alpha \approx 1.26 \pm 0.1 \) corresponding to \( n_Q \approx 2 \), smaller than \( G_{xx}^{\text{min}} \). We note that in Bi\textsubscript{2}Te\textsubscript{3}, CNP and Dirac point do not coincide, with the Dirac point situated within a valley in the bulk valence bands\textsuperscript{9}, and \( G_{xx}^{\text{min}} \) is expected to reflect that\textsuperscript{122}. The obtained value of \( \sim 2 \) is quite remarkable since ‘universality’ of \( n_Q \) has been questioned\textsuperscript{125} given a likely formation of subsurface two-dimensional electron gas (2DEG) states of bulk origin. In Bi\textsubscript{2}Se\textsubscript{3}, where the Dirac point is expected to coincide with CNP, the fit similarly yields \( \alpha \approx 1.12 \pm 0.1 \) (Figure 5-5d), corresponding to two 2D quantum channels we associate with two independent topological surfaces.

Finally, we remark that using thermal protocol illustrated in Figure 5-5b we found that once CNP is reached the system remains there for months on cycling, which was the duration of our experiments. This robustness suggests that at higher electron doses vacancies are likely correlated and more stable complex defects (such as di-vacancies\textsuperscript{118,126}) may form, the details of which ought to be further understood. With the choice of electron beam energy and terminal electron dose controlling the stability of pairs, and with a suitably designed thermal tuning to charge neutrality, the high-energy electron irradiation offers a path to large scale access to topological states.

5.3 Penetration range of electron beams in Bi\textsubscript{2}Te\textsubscript{3} and Bi\textsubscript{2}Se\textsubscript{3}

We note that particle beams such as Ne\textsuperscript{+} and He\textsuperscript{2+} (alpha particles) have a much shorter penetration depth and are known to produce defect cascades and extended defects. For
example, 150 keV Ne\(^+\) ion irradiation has a nonuniform depth profile and shows a pileup of defects at about 0.2 \(\mu\)m. Such irradiations were shown to increase electron concentration away from the surface states and to stabilize Fermi level high up in the bulk conduction bands\(^{127}\) of n-type Bi\(_2\)Te\(_3\) and Bi\(_2\)Se\(_3\). Using 2.5 MeV electron irradiation on 15 \(\mu\)m thick Bi\(_2\)Te\(_3\) and Bi\(_2\)Se\(_3\) crystals create uniform defects due to the large penetration depth of electron in these materials (see Figure 5-6).

![Figure 5-6 Penetration depth of electrons in (a) Bi\(_2\)Te\(_3\) and (b) Bi\(_2\)Se\(_3\) calculated using NIST ESTAR simulator (http://physics.nist.gov/PhysRefData/Star/Text/ESTAR.html). The penetration depth at 2.5 MeV (red dots) used in our experiments is > 2, 000 \(\mu\)m and the resulting depth profile of vacancies is uniform over hundreds of microns.](http://physics.nist.gov/PhysRefData/Star/Text/ESTAR.html)

In our experiments, irradiation were performed in liquid hydrogen. This allowed us to control pair formation during the energetic irradiation process and develop a scheme under which stable vacancy content that compensated native charged defects was established.
5.4 Transmission Electron Microscopy of Bi$_2$Te$_3$ exposed to 2.5 MeV e-beams

Transmission Electron Microscopy (TEM) images of Bi$_2$Te$_3$ irradiated to a dose $\Phi = 1$ C/cm$^2$ are shown in Figure 5-7. At this dose and after 40 hours at room temperature, we estimate the number of the atomic displacements as follows. For 2.5 MeV electrons the effective cross-section is $\sigma \sim 300$ barns (1 barn = $1 \times 10^{-24}$ cm$^2$). With the dose of 0.1 C/cm$^2 = 6.25 \times 10^{17}$ electron/cm$^2$ we have 0.000187 d.p.a on Bi sublattice, corresponding to a roughly $\sim 1$ per 5,000 atoms ejected, which are not easily discerned in the image. The images of irradiated and pristine samples are indistinguishable.

![TEM images and diffraction spots of (a) pristine and (b) irradiated Bi$_2$Te$_3$ showing the same hexagonal lattice in the ab-plane. (c) Layered van der Walls structure along the c-axis (normal to the cleavage plane) after irradiation. Inset: Rhombohedral layered structure of Bi$_2$Te$_3$ constructed using lattice parameters ($a = 4.38$ Å and $c = 30.45$ Å) from the X-ray diffraction (XRD). The van der Walls structure has three quintuple layers per unit cell.](image-url)
We note that electron irradiation induced vacancy - interstitial (Frenkel) pairs\textsuperscript{117,128} can be well controlled by the integrated irradiation dose (beam fluence) to a very low carrier concentration level ($\lesssim 10^{13} \text{ cm}^{-3}$). Electron irradiation produces donor-type doping in both Bi$_2$Te$_3$ and Ca-doped Bi$_2$Se$_3$ compounds (with a slightly higher rate $\approx \partial n_b / \partial \phi$ for Bi$_2$Se$_3$). At the present stage we cannot differentiate between the doping action of vacancies on Bi and Te/Se sites, however, we clearly demonstrate bulk charge compensation of $p$-type TIs across charge neutrality point.

5.5 Tuning through charge neutrality point (CNP) in Bi$_2$Se$_3$:Ca

Longitudinal resistivity $\rho_{xx}$ of the initially $p$-type Ca-doped Bi$_2$Se$_3$\textsuperscript{22,129} measured in situ in the irradiation chamber kept at 20 K as a function of 2.5 MeV electron irradiation dose shows behavior identical to Bi$_2$Te$_3$ under the same irradiation conditions (see Figure 5-1). Figure 5-8 shows two orders of magnitude resistivity increase to a maximum $\rho_{max}$ where the conversion from $p$- to $n$-type takes place defining charge neutrality point (CNP). Irradiation doses to achieve CNP are lower in Bi$_2$Se$_3$:Ca than in Bi$_2$Te$_3$ since the initial bulk conductivity is lower. As in Bi$_2$Te$_3$, for low irradiation doses $\rho_{xx}(\phi)$ is reversible upon temperature cycling to room and back to 20 K, with $\rho_{xx}(\phi)$ and $\rho_{max}$ traced exactly by the next irradiation cycle.
Figure 5-8 (a) Longitudinal resistivity $\rho_{xx}$ of the initially p-type Bi$_2$Se$_3$ doped with 0.09% Ca measured in situ in the irradiation chamber maintained at 20 K as a function of electron irradiation dose $\phi$. Initial (1st) irradiation is shown as red squares. The conversion of conductance from p- to n-type takes place at $\phi_{\text{max}} \approx 18 \text{ mC/cm}^2$. Warming up to room temperature partially reverses the compensation process, which can be fully recovered with repeated irradiation (blue circles). The accumulated dose on the 2nd irradiation at $\rho_{xx}^{\text{max}}$ is $\phi_{\text{max}} \approx 24 \text{ mC/cm}^2$. Inset: The change of sign of the slope $\partial R_{xy}/\partial H$ of Hall resistance $R_{xy}$ from below $\phi_{\text{max}}$ to above. (b) Conductivity type can be changed by annealing (see main text) after irradiating to a very low dose $\phi_{\text{max}} \approx 12.7 \text{ mC/cm}^2$ as shown here in another Bi$_2$Se$_3$:Ca crystal by the changing slope of Hall resistivity.

<table>
<thead>
<tr>
<th>$k_f$ (Å$^{-1}$)</th>
<th>$n$ (cm$^{-3}$)</th>
<th>$1/eR_H$ (cm$^{-1}$)</th>
<th>$\mu$ (cm$^2$V$^{-1}$s$^{-1}$)</th>
<th>$l$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mC/cm$^2$</td>
<td>0.0247</td>
<td>5.06 $\times$ 10$^{17}$</td>
<td>6.8 $\times$ 10$^{18}$</td>
<td>11205.11</td>
</tr>
<tr>
<td>89 mC/cm$^2$</td>
<td>0.0143</td>
<td>1.03 $\times$ 10$^{17}$</td>
<td>1.0 $\times$ 10$^{17}$</td>
<td>--</td>
</tr>
<tr>
<td>99 mC/cm$^2$</td>
<td>0.0226</td>
<td>3.95 $\times$ 10$^{17}$</td>
<td>4.5 $\times$ 10$^{17}$</td>
<td>6892.39</td>
</tr>
</tbody>
</table>

Table 5-1
5.6 Material’s parameters from longitudinal and Hall resistivities and from SdH

Transport parameters for a pristine and irradiated Bi$_2$Te$_3$ is shown in Table 5-1. 1st column shows irradiation doses $\phi$, Fermi wavevectors $k_F$ and the corresponding carrier densities $n$ obtained from Shubnikov-de Haas (SdH) oscillations are in the 2nd and 3rd columns respectively. The inverse Hall coefficient which gives an estimate of carrier density at low magnetic fields is in the 4th column. The last two columns show carrier mobilities $\mu$ and mean free paths $l$.

5.7 Isochronal annealing experiments and energy barriers for defect migration

To evaluate the stability of compensating defects and to estimate energy barriers for defect migration, we performed isochronal anneals of a Bi$_2$Te$_3$ crystal irradiated with 2.5 MeV electrons at 20 K to a relatively low dose $\phi = 76$ mC/cm$^2$ (below type conversion) and traced how the hole concentration $p$ evolved relative to the initial value, positing that the incremental change $\Delta p$ was proportional to the concentration of remaining defects. The initial decrease in hole concentration (here measured at 1 T) was from $3.36 \times 10^{18}$ cm$^{-3}$ to $7.5 \times 10^{17}$ cm$^{-3}$ obtained after cycling to room temperature (RT) and back to 4.2 K. Annealing was performed in 30 min steps each with the temperature increment of 5 °C. After each annealing step, the sample was transferred into the cryostat for the measurements of resistivity and Hall effect at 4.2 K. Above $\sim 70$ °C a measurable change
in carrier density (shown in Figure 5-9a) becomes apparent, and above \( \sim 150 \, ^\circ\text{C} \) the recovery is complete. For the determination of migration energy we used a standard method\(^{118}\): we calculated the increment of the carrier concentration change \( \Delta p_i = p_i - p_{i-1} \) at each annealing step and normalized it to the remaining concentration of holes after the preceding step \( p_{i-1} \), namely \( \Delta p_i / p_{i-1} \). We assumed that the annealing process in each step obeys the rate equation \( \frac{dp}{dt} = -Kp \), corresponding to exponential decay law, \( p = p_0 e^{-Kt} \), where \( K \) is the diffusion coefficient \( K \propto e^{-E_b/kT} \) with an energy barrier \( E_b \) impeding the migration of defects. The rate of annealing in step \( i \), \( \Delta p_i / p_{i-1} \) (\( \approx \frac{d\ln p}{dt} \)) is proportional to \( K \), thus the slope of \( \ln(\Delta p_i / p_{i-1}) \propto \ln(e^{-E_b/kT}) \) vs. \( 1/k_B T \) gives an estimate of the energy barrier controlling defect migration, \( E_b \approx 0.794 \, \text{eV} \) (Figure 5-9b). This is a typical value for the migration energy of vacancies and over an order of magnitude above \( \lesssim 0.08 \, \text{eV} \) expected for interstitials\(^{118}\). The slow annealing process observed in a
crystal irradiated to just above neutrality point in Figure 5-3 is consistent with this large value of $E_b$; this explains why it takes hundreds of hours at RT to get back to the charge neutrality point. We note that at room temperature interstitials do not contribute since their concentration is about six orders of magnitude below that of vacancies\textsuperscript{118}.

### 5.8 Variable range hopping bulk charge transport at low carrier density

Sheet resistance $R_{xx}$ vs. temperature of Bi$_2$Te$_3$ crystal at the charge neutrality point (shown in Figure 5-10a) at low temperatures is temperature independent, consistent with minimum conductance $G \approx 20G_0$ ($G_0 = e^2/h$ is the conductance quantum) of surface conduction channels, see main text. At higher temperatures two types of activated bulk behavior are
observed. With a note of caution, since the temperature range is small here, at the highest temperatures (near room temperature) a simple activation law $R_{xx} \propto e^{E_a/k_B T}$ with a very small activation barrier $E_a \sim 7$ meV fits best (see inset in Figure 5-10a). We note that the barrier $E_a$ appears much smaller than the bulk gap $E_g \approx 200 - 300$ meV, and well below room temperature equivalent to $\sim 30$ meV.

This simple activated behavior changes below $\sim 220$ K into variable range hopping (VRH) of the Efros-Shklovskii type: $R_{xx} = R_0 \exp\left[\left(T_0/T\right)^{1/2}\right]$. A fit of $R_{xx}$ to the VRH law (Figure 5-10b) gives $E_{VRH} = k_B T_0 \sim 160$ meV equivalent to $T_0 \sim 1883$ K. Variable range hopping is a tunneling transport between separated charge puddles created by large charge fluctuations at CNP which are poorly screened, leading to unreliable estimates of the true carrier densities from Hall transport. Such behavior is expected in theory\(^{89}\), where VRH temperature scale is given by $T_0 \propto 1/\xi$, and $\xi$ is the localization length of states in the vicinity of the Fermi level measured in the units of cube root of carrier density $n^{1/3}$. Using our value of $T_0$ from the fit to VRH law and carrier densities $n \sim 10^{18}$ cm\(^{-3}\), we obtain a rough estimate\(^{89}\) of $\xi \sim 0.4n^{-1/3} \approx 4$ nm. Owing to this remarkably long localization length VRH can dominate transport over a large temperature range, from $T \approx 220$ K down to the low-temperature plateau in $R_{xx}$.

5.9 Stable charge neutrality point in Bi2Te3 achieved through annealing

We remark that the achieved high bulk resistivity values at CNP can be particularly useful in spintronic TI-based devices. For example, a simple estimate based on spin-torque
transfer effect induced by spin-polarized currents through topological surfaces indicates that in a typical thin layer structure spin-torque transfer (STT) will be hugely enhanced; indeed in our irradiated TIs even at room temperature it will be only about 20-30% below the maximum STT expected from the spin-polarized surfaces alone. This is a large improvement over the unirradiated TIs\textsuperscript{130} where STT is 70% below maximum owing to the significant shunting by the bulk.

Figure 5-11 Sheet resistance $R_{xx}$ vs. temperature of another Bi$_2$Te$_3$ crystal at the charge neutrality point reached through the annealing schedule shown in Figure 5-4. Inset: Magnetoresistance at 1.9 K in a pristine Bi$_2$Te$_3$ crystal (blue) and of Bi$_2$Te$_3$ crystal irradiated with electron dose of 1 C/cm$^2$ after annealing step at 120ºC, which returned the crystal to CNP (red), see also Figure 5-4. At CNP a sharp weak antilocalization (WAL) cusp appears, which is not seen in either pristine crystals or after irradiation.
Chapter 6 Open questions

The work done in this thesis focused mainly on the disorder effects in charge transport and spin response of topological insulators. Disorder enabled us to uncover a singular paramagnetic response from the Dirac spins, and promoted an exotic surface superconductivity that emerges at extraordinarily high temperatures. We also developed a new technique to counteract disorder using high energy electron irradiation, and achieved stabilizing topological insulators at the charge neutrality point (CNP).

This work has also opened many new questions yet to be investigated. Thermal robustness of topological surfaces all the way up to room temperature observed in the AC susceptibility measurements hints the existence of a non-equilibrium connection between these surfaces and the bulk. It may promote an extreme surface cooling, which could have profound consequences in science and technology, and is yet to be understood. In the irradiated topological insulators, we observe anomalous Hall effect which appears at CNP in the absence of magnetic impurities. This clearly indicates new physics waiting to be uncovered. The fundamental origins and pairing mechanism of exotic superconductivity we have discovered in Sb$_2$Te$_3$ at relatively high temperatures are still unknown. Is this new state chiral? Does it support Majorana fermions? What is the nature of superconducting vortices residing in superconducting Dirac puddles? It is clear that further experimental
explorations are needed to check if this new superconducting state is truly two-dimensional and has a $p$-wave symmetry of the ground state. And more generally, perhaps single-particle physics – a common theoretical description of these materials – is not adequate to explain the unusual high-temperature phenomena we have observed and particle-particle correlations cannot be ignored.

Finally, our understanding of the effects of disorder on topological phases is still at its infancy, and so is the role of charged bulk defects in the self-organization of surface Dirac charge puddles. We expect that striving to achieve the control of the material’s growth and charge doping processes will eventually bring much needed clarity and scientific progress to this field.
List of publications – Lukas Zhonghua Zhao


Robust topological interfaces and charge transfer in epitaxial Bi$_2$Sw$_3$ II-VI semiconductor superlattices, Chen, Zhiyi; Zhao, Lukas; Park, Kyungwha; Garcia, Thor; Tamargo, Maria; Krusin-Elbaum, Lia. in review in Nano Letters (2015).


Hollandites as a new class of multiferroics, S. Liu, A.R. Akbashev, X. Yang, X. Liu, W. Li, L. Zhao, X. Li, A. Couzis, M.-G. Han, Y. Zhu, Lia Krusin-Elbaum, J. Li, L. Huang,
LIST OF PUBLICATIONS – LUKAS ZHONGHUA ZHAO


Conference talks

Compensation of intrinsic charge carriers in topological insulators using high energy electron beams. Lukas Zhao, Haiming Deng, Jeff Secor, Marcin Konczykowski, Andrzej Hruban, Lia Krusin-Elbaum. APS March Meeting 2014

Superconductivity in a topological insulator Sb$_2$Te$_3$. Lukas Zhao, Haiming Deng, Milan Begliarbekov, Inna korzhovska, Zhiyi Chen, Jeff Secor, Lia Krusin-Elbaum. APS March Meeting 2013.


18. Qi, X. & Zhang, S. Topological insulators and superconductors. 1–54


36. Fu, L. & Kane, C. L. Superconducting proximity effect and Majorana fermions at the surface of a topological insulator. 1–4


