Quantum transport study in three-dimensional topological insulator BiSbTeSe₂

Yang Xu⁷ and Yong P. Chen⁷,⁶,⁵,⁴,³,²,¹,*,⁶

¹Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing, China
²Department of Physics and Astronomy, Purdue University, West Lafayette, IN, United States
³Birck Nanotechnology Center, Purdue University, West Lafayette, IN, United States
⁴Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, IN, United States
⁵Quantum Science Center, Oak Ridge, TN, United States
⁶School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, United States
⁷Institute of Physics and Astronomy and Villum Centers for Dirac Materials and for Hybrid Quantum Materials, Aarhus University, Aarhus-C, Denmark
⁸WPI-AIMR International Research Center for Materials Sciences, Tohoku University, Sendai, Japan

*Corresponding author: e-mail address: yong.chen@purdue.edu

Contents

1. Introductions 74
   1.1 Brief introduction to topological states of matter 74
   1.2 Introduction to 3D topological insulators 75
   1.3 Introduction to an intrinsic 3D TI BiSbTeSe₂ 76

2. Electrical properties of BiSbTeSe₂ 77
   2.1 Sample preparation 77
   2.2 Extracting conduction for the surface states 79
   2.3 Hall measurements of bulk samples 82
   2.4 Ambipolar field effect 83
   2.5 Weak antilocalization effect 85
   2.6 "Half-integer" quantum Hall effect in BSTS 86
   2.7 Extracting π Berry's phase for the TSS 89
   2.8 Temperature dependence of the QHE 91
   2.9 Discussion 92

3. Quantum transport in dual-gated 3D topological insulators 93
   3.1 Dual-gated device fabrication 93
   3.2 Independent gating of top and bottom surfaces 93
   3.3 Gate efficiency and mobility of each surface 95
   3.4 Minimum conductivity at double Dirac point 96
   3.5 Dual-gate tunable QHE 97
   3.6 Non-local measurements 99
   3.7 Temperature dependence study 101
   3.8 Dissipationless and dissipative edge states 103

4. Inter-surface hybridization in thin BSTS 107
1. Introductions

1.1 Brief introduction to topological states of matter

For a long period, principles of symmetry breaking have been successful in classifying most of the states of matter. However, the discovery of quantum Hall effect (QHE) in the 1980s brought a new paradigm for studying topology and symmetry protected topological states in condensed matter systems (Klitzing et al., 1980; Thouless et al., 1982). In the QHE, the Hall conductance of a two-dimensional electron gas (2DEG) is precisely quantized into integer multiples of $e^2/h$ ($\sim 1/25812.8 \Omega^{-1}$), insensitive to the sample geometry or microscopic details of the material. The QHE has been used as a resistance metrology and the determination of fundamental constants. The precise quantization of the Hall signal can be understood by the fact that it is topological states of matter characterized by a topological invariant, so-called Chern number or Thouless–Kohmoto–Nightingale–Nijs (TKNN) number (Thouless et al., 1982). In a topological phase, certain fundamental properties, such as the quantized Hall conductance or the number of edge modes (equal to Chern number) in a QHE, are independent of material details and insensitive to smooth deformation of the system (such as small density change or disorders) unless the system passes through a quantum phase transition.

Since the proposals of quantum spin Hall effect (Bernevig and Zhang, 2006; Kane and Mele, 2005), many new materials belonging to different topological classifications have been theoretically predicted and experimentally verified. Typical examples are 2D and 3D topological insulators (TIs), topological crystalline insulators, and topological Dirac/Weyl semimetals (Armitage et al., 2018; Dziawa et al., 2012; Fu, 2011; Hasan and Kane, 2010; Qi and Zhang, 2011). Recent high-throughput searches give a catalogue of more than a thousand of topological materials (Tang et al., 2019; Vergniory et al., 2019). Different from the QH states that need magnetic field to explicitly break the time-reversal symmetry, the 2D and 3D TIs
belong to a class that is invariant under time reversal, where spin-orbit coupling (SOC) and band inversion play essential roles. While the bulk remains gapped, helical and non-trivial edge/surface states emerge near the boundary of a 2D/3D TI.

The revelation of topological Dirac and Weyl semimetals have broadened the classification of topological phases of matter beyond insulators (Armitage et al., 2018). The Dirac semimetal is a natural 3D counterpart of graphene, possessing 3D Dirac fermions in the bulk. In other words, the energy bands disperse linearly along all (three) momentum direction, with the conduction bands touching the valence bands only at the Dirac point. The 3D Dirac fermion in a Dirac semimetal is composed of two overlapping Weyl fermions (chiral massless particles) that has opposite chirality and can be separated in momentum space if either time-reversal or inversion symmetry is broken, resulting in a Weyl semimetal. Interestingly, the surface states of a Dirac/Weyl semimetal take the form of Fermi arcs connecting the projection of bulk Dirac/Weyl nodes in the surface Brillouin zone. Dirac semimetals include materials such as Na$_3$Bi and Cd$_3$As$_2$ (Liu et al., 2014a,b). Weyl semimetals are such as TaAs (Lv et al., 2015; Xu et al., 2015).

The Dirac/Weyl semimetal phase could also exist near the topological phase transition from a band insulator to a 3D TI by breaking symmetries at the critical point (Liu et al., 2014a). Or a Weyl semimetal can be constructed by stacking thin films of magnetically doped 3D TIs (Burkov and Balents, 2011). This provides a way in realizing Weyl semimetals from the tuning phase transitions in a 3D TI. A 2D topological semimetal (TSM) phase has been predicted to be induced by an in-plane magnetic field in thin 3D TIs with hybridization gap opening (Burkov and Balents, 2011). The in-plane magnetic field plays a role to split and separate the top and bottom surface Dirac points in the momentum space and prohibit their mutual annihilation, resulting in the 2D TSM phase that provides a 2D analogue of Weyl semimetals (even though strictly speaking Weyl fermions cannot be defined in even spatial dimensions (Armitage et al., 2018)).

1.2 Introduction to 3D topological insulators

A 3D TI is characterized by an insulating bulk band gap and gapless conducting topological surface states (TSS) of spin-helical massless 2D Dirac fermions. The conducting edge in a 3D TI is a closed 2D surface (while in a 2D TI it is a 1D edge channel). Simple versions of 3D TI materials have been first found in Bi$_x$Sb$_{1-x}$ (x in certain range), Bi$_2$Se$_3$, Bi$_2$Te$_3$ and
Sb\textsubscript{2}Te\textsubscript{3}, all of which are considered to be prototypical 3D TIs (Hasan and Kane, 2010; Qi and Zhang, 2011). Historically those bismuth/antimony based chalcogenides have been studied for a long time but mostly from the perspective of thermoelectric properties (Heremans et al., 2017; Nolas and Jeffrey Sharp, 2013). Only until recent years has their profile as a 3D TI been discovered. 3D TI has also been reported in strained HgTe (Brüne et al., 2011). There are also some other types of TI candidates such as SmB\textsubscript{6} and ternary Heusler compounds (Chadov et al., 2010; Wolgast et al., 2013). The 3D TI provide a novel playground for realizing potential applications and intriguing physics. For example, the spin–momentum locking of the TSS promises for current-driven spintronics. A magnetic gap on the surface state can lead to an exotic quantum anomalous Hall state and topological magnetoelectric effect (Chang et al., 2013; Qi et al., 2008; Yu et al., 2010). The TSS in proximity with s-wave superconductors would lead to a state that supports Majorana fermions, which have the potential for building topological quantum computers (Fu and Kane, 2008).

The experimental demonstration of 3D TIs was first done by angle-resolved photoemission spectroscopy (ARPES) method (Hasan and Kane, 2010). The ARPES data clearly resolved the spin-polarized Dirac cone. However, the prototypical 3D TIs (such as Bi\textsubscript{2}Se\textsubscript{3}, Bi\textsubscript{2}Te\textsubscript{3}) are usually suffering from unintentional impurity doping, in which bulk carriers unavoidably present and overwhelm the topological surface states (TSS) in transport studies. There are a lot of efforts trying to lower the bulk contribution to observe truly TSS dominated transport signatures. Several strategies are commonly applied. (1) Chemical doping, such as adding Ca substitution for Bi in Bi\textsubscript{2}Se\textsubscript{3} to compensate the electrons created by Se vacancies (Hor et al., 2009). (2) Lowering the surface to bulk ratio, such as fabricating thin film and nanostructures. (3) Electrical gating (Chen et al., 2010). Through these efforts, people have been able to observe for example, weak-antilocalization (Chen et al., 2010), Shubnikov-de Haas (SdH) oscillations (Analytis et al., 2010; Qu et al., 2010) and an ambipolar electric field effect (Kong et al., 2011) arising from the TSS. An example of developing QHE from TSS has only been reported in strained HgTe (Brüne et al., 2011).

1.3 Introduction to an intrinsic 3D TI BiSbTeSe\textsubscript{2}

After initial attempts in reducing bulk carriers of binary compounds, ternary and quaternary tetradymitite TI materials such as Bi\textsubscript{2}Te\textsubscript{2}Se and
Bi$_{2-x}$Sb$_x$Te$_{3-y}$Se$_y$ have been found to exhibit large bulk resistivity and more intrinsic behaviors (Arakane et al., 2012; Ren et al., 2010; Taskin et al., 2011; Xia et al., 2013; Xiong et al., 2012b). Here, “intrinsic” means “compensated” (a terminology commonly used for describing undoped semiconductors) and the bulk carriers should freeze out at low temperatures. High-quality single crystals of BiSbTeSe$_2$ (BSTS) have been grown by vertical Bridgman method. Various material characterizations are performed on the BSTS crystals, including X-ray diffraction spectrum (Fig. 1A), elemental energy-dispersive X-ray spectroscopy (Fig. 1B), ARPES (Fig. 2), and scanning tunneling spectroscopy (see the results elsewhere (Xu et al., 2014)), all in good agreement with previous measurements in this material system. They indicate that the BSTS is promising for observing surface-dominant transport properties, which will be presented in the following sections.

2. Electrical properties of BiSbTeSe$_2$

2.1 Sample preparation

In order to perform electrical transport measurements on BSTS, we fabricated samples into two forms, bulk samples and relatively thin devices. The bulk samples (thickness $>1\mu$m) are cleaved from the bulk crystal with razor blades and then cut into rectangular shapes, followed by making electrical connections with multiple indium contacts and gold wires (Fig. 3A). The thicknesses of bulk samples are measured by taking an optical image of the cross section under a microscope. Thinner flakes ($<\sim500\text{nm}$)
are mechanically exfoliated onto Si substrates (with 300-nm-thick SiO₂ coating) by a typical “Scotch tape” method (exploited for the discovery of monolayer graphene) (Novoselov et al., 2004). The thickness of a nanoflake is measured by atomic force microscopy (AFM). The Cr/Au electrodes for the device with exfoliated flakes are defined by standard electron-beam lithography and metal evaporation (see Fig. 3B).
Both types of devices have a six terminal quasi-Hall-bar geometry. Four-terminal longitudinal \(R_{xx}\) and Hall \(R_{xy}\) resistances were measured with the standard lock-in technique and low-frequency (<20 Hz) excitation current of 1 µA (bulk samples) or 100 nA (thin flake devices). Then we can extract the corresponding 2D resistivities \(\rho_{xx} = R_{xx} \frac{W}{L} = R_{sh}\) (with \(R_{sh}\) being the sheet resistance, \(W\) and \(L\) the channel width and length for the quasi-Hall-bar devices), \(\rho_{xy} = R_{xy}\), and conductivities \(\sigma_{xx} = \rho_{xx}/(\rho_{xx}^2 + \rho_{xy}^2)\), \(\sigma_{xy} = \rho_{xy}/(\rho_{xx}^2 + \rho_{xy}^2)\) from tensor inversion.

2.2 Extracting conduction for the surface states

We first discuss transport measurements performed at zero magnetic field \((B=0\,\text{T})\) on BSTS samples with various thicknesses \((t)\). Fig. 4A shows the temperature \((T)\) dependence of the sheet resistance \(R_{sh}\) in five selected samples. The thicker samples (e.g., \(t = 37 \,\mu\text{m}\)) exhibit an insulating behavior (attributed to the bulk) at high \(T\) and a saturating and weakly metallic behavior (attributed to the residual metallic surface conduction after bulk carriers freeze out) at low \(T\). As \(t\) (thus the bulk portion) is reduced (e.g., \(t = 160\) and 480 nm), the bulk-insulating behavior weakens and moves to higher \(T\), while the \(T\) range for the low-\(T\) metallic conduction expands. For thinnest samples \((t = 20 \,\text{nm})\), \(R_{sh}\) is relatively \(T\)-insensitive and metallic over the entire \(T\) range measured. We also note that while \(R_{sh}\) (attributed to the bulk) at high \(T\) (>250 K) is strongly thickness-dependent (from <10 to >2 kΩ), \(R_{sh}\) (attributed to the surface) at low \(T\) (<50 K) reaches comparable values (a few kΩ) for all the samples despite the three orders of magnitude variation in their thickness (Fig. 4B).
In order to more quantitatively extract the surface contribution to the total conductance \( G_{\text{tot}} \), we fit our \( R_{\text{sh}}(T) \) data to a simple model (Gao et al., 2012), where the total (sheet) conductance \( G_{\text{tot}}(T) = 1/R_{\text{sh}}(T) \) is the parallel sum of a thermally activated bulk conductance \( G_b(T) = \rho_b e^{\Delta/k_BT} \) (where \( k_B \) is the Boltzmann constant and the fitting parameters \( \rho_b \) is the high temperature bulk resistivity and \( \Delta \) is the activation energy) and a metallic surface conductance \( G_{\text{sur}}(T) = (R_{\text{sh0}} + A^*T)^{-1} \) (where fitting parameters \( R_{\text{sh0}} \) represents a low-\( T \) residual resistance due to impurity scattering and \( A \) reflects the electron-phonon scattering) (Giraud et al., 2012; Novoselov et al., 2004). This model gives reasonably good fits to most samples over the full temperature range, and extract the surface conductance \( G_{\text{sur}} \) for samples with different thickness (Fig. 5). In few samples (e.g., see Fig. 5A, \( t=80\text{nm} \)), the fitting is excellent from 300 K down to 50 K, but would not account for a small resistance peak at lower \( T \) (\( \sim 30\text{K} \), where the fit underestimates the data by up to 10%). This peak might be due to a small part of the sample insulating with a thermal activation gap smaller than the main bulk activation gap. Each curve was fitted multiple times over different ranges of temperatures to calculate approximate confidence intervals and error bars (Fig. 5B and C). For each fitting parameter, the values for all samples (with widely-varying thickness) are found to be on the similar order of magnitude. For example, \( A \) ranges mostly from 1 to \( 8\Omega/\text{K} \) (average \( 6\Omega/\text{K} \)), comparable with a previous measured \( 3\Omega/\text{K} \) in Bi\(_2\)Se\(_3\) (Kim et al., 2012b).

Fig. 4  (A) Sheet resistance \( (R_{\text{sh}}) \) measured at zero magnetic field vs temperature \( (T) \) in five devices with different thicknesses. (B) \( R_{\text{sh}} \) and 3D resistivity \( \rho_{3D} \) as functions of sample thickness at two typical temperatures. At \( T=2\text{K} \), the sheet resistance exhibits a 2D behavior (relatively insensitive to the thickness, highlighted by the blue horizontal band). The dashed line and red horizontal band represent expected behaviors for \( R_{\text{sh}} \) and \( \rho_{3D} \) of a 3D bulk conductor (our data at 290K exhibited such 3D behavior for samples thicker than \( \sim 00\text{nm} \)).
Fig. 6A shows the ratio \( G_{\text{sur}}/G_{\text{tot}} \) between the fitted \( G_{\text{sur}} \) and the measured total conductance \( G_{\text{tot}} \) as a function of thickness (t) at three representative temperatures. At low T (50K or below), the surface contribution to the total conductance is nearly 100% for all the samples (with thickness from <100nm to >10μm). Remarkably, in thinner samples (t<100nm) the surface conduction can exceed the bulk conduction with \( G_{\text{sur}}/G_{\text{tot}} > 60\% \) even at room temperature (T=290K).

The averaged fitting parameters (between all samples) \( \rho_{b0} = (3.5 \pm 0.8) \times 10^{-3} \Omega \cdot \text{cm}, \Delta = 57 \pm 4 \text{meV}, R_{\text{sh0}} = 2.0 \pm 0.3 k\Omega, A = 6 \pm 2 \Omega/\text{K} \) can be taken as representative material parameters for BSTS to predict \( G_{\text{sur}}/G_{\text{tot}} \) for any given thickness and temperature, as shown in the 2D color plot in Fig. 6B. We see that surface-dominant transport with \( (G_{\text{sur}}/G_{\text{tot}} > 80\% \) can be achieved in moderate low T (<50K) for samples up to
macroscopic thickness of 1 mm, or in thin (tens of nm) samples up to room temperature (300 K) in this material system. The extracted bulk thermal activation energy $\Delta$ is larger than $2k_B T$ for $T = 300K$, which enables surface-dominant transport at room temperature. We also note $\Delta$ is much smaller than the bulk band gap of 300 meV. This has been noted in previous measurements in similar materials (Taskin et al., 2011), and may be a result of impurity states in the bulk gap in the compensation-doped sample (while spatial variation in the doping can further reduce the observed $\Delta$) (Skinner et al., 2013).

2.3 Hall measurements of bulk samples

Hall measurement of bulk samples with different thicknesses are performed to ensure the carriers mostly coming from the surface. See Fig. 7, the Hall resistance ($R_{xy}$) at low magnetic fields in three different bulk samples measured at 2 K. The linear Hall slope was used to extract the 2D carrier densities to be $6.6 \sim 7.8 \times 10^{12} \text{cm}^{-2}$ (p type), nearly independent of thickness (varying from 20 to 52 μm), indicating surface origin of the carriers (shown in Fig. 7B). In contrast, the converted 3D carrier density $n_{3D} (= n_{2D}/t)$ nearly scales as $t^{-1}$ and is as low as $1.4 \times 10^{15} \text{cm}^{-3}$ for the 52-μm-thick sample. Furthermore, according to the ARPES measured band structure (Arakane et al., 2012), the maximum carrier density that can be accommodated in the surface bands before occupying the bulk bands is at
least $10^{13}$ cm$^{-2}$ (both surfaces combined). Therefore, the measured Hall density comes mostly from the surface. This is also consistent with the surface-dominated conduction shown in Fig. 6 and the Fermi level residing inside the bulk band gap discussed above. Almost all the samples measured in this work (down to 20 nm thick) give densities on the order of $10^{12}$ cm$^{-2}$ before gating. The true bulk density should be even lower than $1.4 \times 10^{15}$ cm$^{-3}$, which is more than one order of magnitude lower than the lowest values from bulk-insulating 3D TIs previously reported (Analytis et al., 2010; Taskin et al., 2011; Xiong et al., 2012a).

### 2.4 Ambipolar field effect

Almost all the thin-flake device on Si/SiO$_2$ substrates measured have shown nice ambipolar field effect, with a single resistance peak by sweeping back-gate voltages $V_{bg}$, similar to that of graphene (Novoselov et al., 2005). Need to note that for samples thicker than $\sim 100$ nm, they are thick enough such that our Si/SiO$_2$ backgate can be used to independently tune the bottom surface carrier density $n_b$ without affecting the top surface density $n_t$ (see Fig. 8A for device schematic). Fig. 8B shows field effect measurement of a 160-nm-thick device (Sample A) at different temperatures. A single resistance peak maintains up to $\sim 150$ K, indicating high quality of samples.

The peak in $R_{xx}$ is identified as the charge neutrality point (or Dirac point, DP) of the bottom surface, whose carriers change from electrons at higher $V_{bg}$ to holes at lower $V_{bg}$. Repeated gate sweeps can have a slight hysteresis (of several volts). The ungated surfaces (both bottom surface, as seen by the field effect, and top surface) are generally found to be n-type (with electron carriers) in the exfoliated thin flakes.
When both surfaces have electron carriers ($V_{bg} > -60\,\text{V}$), the $R_{xy}(B)$ shows linear behavior in low-$B$ regime, which can be used to extract the total 2D carrier density $n_{2D}$ and mobility $\mu$ (see Fig. 9). The linear fit of $n_{2D}$ vs $V_{bg}$ gives a gate efficiency of $2.7 \times 10^{10}\,\text{cm}^{-2}/\text{V}$ and the extrapolation to $V_{bg} \approx -60\,\text{V}$ (Dirac point of bottom surface) gives an approximate top surface density $\sim 0.4 \times 10^{12}\,\text{cm}^{-2}$. The highest mobility $\sim 3000\,\text{cm}^2/\text{Vs}$ is extracted when $V_{bg}$ is close to bottom surface DP. At more negative $V_{bg}$ ($<-60\,\text{V}$), the holes from the bottom surface will compensate for the electrons from the top surface and start to reduce $R_{xy}$, and $R_{xy}(B)$ start to show non-linear behavior (Fig. 9 inset).
2.5 Weak antilocalization effect

The low-field magnetotransport measurements of our BSTS flakes show weak antilocalization (WAL) behaviors, which is a characteristic feature of a system with spin-orbit coupling. It comes from the destructive quantum interference of coherently back-scattered spin-1/2 electrons. Such a WAL can be suppressed by a magnetic field (which breaks time-reversal symmetry) and higher temperature (killing coherence), giving rise to the observed positive magnetoresistance at low temperature. The WAL in TI can be fit by the Hikami-Larkin-Nagaoka (HLN) formula (Hikami et al., 1980):

\[
\Delta \sigma(B) = -\alpha \frac{e^2}{2\pi^2\hbar} \left[ \ln \left( \frac{\hbar}{4el_p^2B} \right) - \Psi \left( \frac{\hbar}{4el_p^2B} + \frac{1}{2} \right) \right]
\]

where \(\Delta \sigma(B)\) is two-dimensional conductivity change due to magnetic field, \(\Psi\) is the digamma function, \(l_p\) is the phase coherent length and \(\alpha\) is the prefactor. For an ideal 2D TI surface which has strong spin-orbit coupling, \(\alpha\) should be \(-0.5\). However, in a real TI sample, two parallel surfaces (top and bottom) are involved in the conduction, giving rise to \(\alpha = -1\). In BSTS flakes, \(\alpha\) is close to \(-1\) when both surfaces are n doped \((V_{bg} > -40\text{ V})\) in Fig. 10. When bottom surface is p doped \((V_{bg} < -40\text{ V})\) or...
close to Dirac point ($V_{bg} \sim -40$ V) and top surface is n doped, the fit gives $\alpha$ value deviates from $-1$. The reason is not clear so far, and it has not been reported before in this bipolar situation (Chen et al., 2010; Kim et al., 2013).

2.6 “Half-integer” quantum Hall effect in BSTS

When a back-gate tunable BSTS device (such as Sample A) is brought into a perpendicular high magnetic field $B$, well-developed QHE shows up.
Fig. 11A presents the longitudinal resistance \( R_{xx} \) and Hall resistance \( R_{xy} \) as functions of the backgate voltage \( V_{bg} \), measured at \( T=0.35 \) K. At the electron side \( (V_{bg} > -60 \) V) of the bottom surface (same carrier type as the ungated top surface), we observe well-defined quantized plateaus in \( R_{xy} \) with values \( \sim h/e^2 \sim 25.8 \) k\( \Omega \) and \( \sim h/2e^2 \sim 12.9 \) k\( \Omega \), accompanied by vanishing \( R_{xx} \), over broad ranges in \( V_{bg} \) around \(-45 \) and \(-20 \) V, respectively. These are the hallmarks of well-developed QHE, associated with total (top + bottom surfaces) Landau filling factors \( \nu = 1 \) and 2, respectively (see more discussions below). In addition, a developing \( R_{xy} \) plateau \( \sim h/3e^2 \) is observed, accompanied by a minimum in \( R_{xx} \). To gain more insights on the QHE, we perform tensor inversion to extract the 2D longitudinal and Hall conductivities \( \sigma_{xx} \) and \( \sigma_{xy} \), shown in Fig. 11B in units of \( e^2/h \). We again observe quantized (and developing) plateaus in \( \sigma_{xy} \) at \( \nu e^2/h \) with integer \( \nu = 1, 2 \) (and 3), concomitant with vanishing \( \sigma_{xx} \) (or minimum in \( \sigma_{xx} \)).

Such integer quantized Hall conductivities (QHC) measured can be understood as the sum of a half-integer QHC from the top surface (fixed at \( e^2/2h \) at \( B=31 \) T) and another half-integer QHC from the bottom surface (at \( \nu_b e^2/h \) with \( \nu_b = 1/2, 3/2, \) and \( 5/2 \), etc, depending on the \( n_b \) tuned by the backgate). In other words, the observed QHC \( \sigma_{xy}^{\text{total}} = \nu e^2/h = \sigma_{xy}^{\text{top}} + \sigma_{xy}^{\text{bottom}} = (\nu_t + \nu_b)e^2/h = (N_t + N_b + 1)e^2/h \), with top (bottom) surface QHC \( \sigma_{xy}^{\text{top(bottom)}} = \nu_{t(b)}e^2/h = (N_{t(b)} + 1/2)e^2/h \), where \( \nu_{t(b)} = N_{t(b)} + 1/2 \) and \( N_{t(b)} \) are the Landau filling factor and Landau level index of top
(bottom) surface corresponding to the QH state (in our case $\nu_t=1/2$ and $N_t=0$, thus $\sigma_{xy}^{\text{top}}=(1/2)e^2/h$, even though the individual (half-integer) QHC of each surface cannot be directly measured in the experiment (where we always measure the total Hall conductance of the two parallel conducting surfaces). Such a half-integer QHE reflecting the contribution in units of half quantum conductance ($e^2/h$) from each surface is the hallmark of the topological surface state QHE unique to 3D TI.

We further studied $\sigma_{xx}(V_{bg})$ at a series of magnetic fields ($B$), and found quantum Hall features can be observed down to $B=13$ T. Fig. 12 plots $\sigma_{xx}$ (color scale) as functions of both $B$ and $V_{bg}$. Dashed lines trace the gate voltage ($V_{bg}$) positions for the observed QH minima in $\sigma_{xx}$ and are labeled by the corresponding half-integer fillings ($\nu_b$) of the bottom surface LL. When extrapolated to $B=0$ T, these lines converge to $\sim -58$ V, the bottom surface DP (noted as $V_D=-58$ V). Dotted lines trace the maxima in $\sigma_{xx}$ labeled by corresponding integer $\nu_b$, also extrapolate and converge to the same DP. The odd integer ratio $(-3:-1:1:3:5)$ in the $V_{bg}$ positions for the QH minima in $\sigma_{xx}$ as measured away from DP ($V_{bg}-V_D$, proportional

---

![Fig. 12](image.png)

**Fig. 12** A 2D color plot showing $\sigma_{xx}$ (color scale) as function of $B$ field and backgate voltage ($V_{bg}$) measured at 0.35 K in sample A. Dashed and dotted lines correspond to half-integer and integer Landau filling factors (marked) for the bottom surface, respectively. Inset: the $V_{bg}$ positions for minima in $\sigma_{xx}$ (left axis) and $(V_{bg}-V_D)/\Delta V$ (right axis) as a function of $\nu$ at 31 T, where $\Delta V$ is half the averaged $V_{bg}$ separation between two consecutive Landau levels.
to $m_b$), exemplified in the inset for the 5 QH minima in $\sigma_{xx}$ at $B=31$ T (corresponding to the intercepts of dashed lines on the top axis in Fig. 12), is in contrast to the QHE for conventional 2DES, where the density (filling factor) ratio for successive QHE states ($\sigma_{xx}$ minima) would be consecutive integers. This half-integer shift in our data (integer/half-integer fillings for maxima/minima) is a consequence of the unique 0th LL shared half-half between Dirac electrons and holes, and a manifestation of the Berry’s phase $\pi$ associated with spin-helical carriers of TSS. It is known that $\sigma_{xx}$ peaks when the gate tunable chemical potential ($\mu_b$) passes through the center of last-filled LLs where the density of state (DOS) peaks and underlying state may become delocalized. The prominent central peak (the central line with $\nu_b=0$, non-dispersive with $B$) at DP in our $\sigma_{xx}$ data represents the 0th LL, that is fixed at zero energy and non-dispersive with $B$, and a hallmark of Dirac fermions.

2.7 Extracting $\pi$ Berry’s phase for the TSS

The QHE is also measured by sweeping the magnetic field ($B$) at fixed carrier densities ($V_{bg}$). Fig. 13A upper (lower) shows $R_{xy}$ ($R_{xx}$) measured vs $B$ at $T=0.35$ K in sample A for various $V_{bg}$’s. The $R_{xy}(B)$ trace for $V_{bg}=-52$ V (close to bottom surface DP) exhibits a wide quantized plateau $\sim h/e^2$ for $B>\sim 13$ T, indicating both surfaces are in the $1/2(e^2/h)$ QH state (0th LL, $N_t=N_b=0$, $\nu=1/2+1/2=1$). When $V_{bg}$ is increased to $-41$ V, the higher $n_b$ shifts the $\nu_b=1/2$ ($N_b=0$) QH state of bottom surface, thus the observed $h/e^2$ ($\nu=1$) plateau to higher $B$. Meanwhile, another QH plateau develops at $h/2e^2$ at lower $B \sim 15$ T (but still $>\sim 13$ T with $N_t=0$), corresponding to $N_b=1$, $\nu=\nu_t+\nu_b=1/2+3/2=2$. This QH state also shifts to higher $B$ as $V_{bg}$ increases ($-35$, $-30$ and $-20$ V). Further increasing $V_{bg}$ to $-14$, $-5$ and $2$ V, yet another QH plateau develops at $h/3e^2$, now corresponding to $N_t=0$, $N_b=2$, $\nu=\nu_t+\nu_b=1/2+5/2=3$. All these QH plateaus are accompanied by vanishing $R_{xx}$ or dips in $R_{xx}$, with each QH state generally better developed at higher $B$ (reflecting the larger LL energy gap).

From the evolution of the QH states seen in Fig. 13A for $B>\sim 13$ T, we can extract the bottom surface LL index $N_b$ (note $N_t$ is fixed at 0) and perform the standard fan diagram analysis (Fig. 13B, showing $N_b$ vs $1/B$), where integer/half- integer $N_b$ is assigned the minima/maxima in $R_{xx}$ (or alternatively in $dR_{xy}/dB$, with an example shown in Fig. 13B inset). We find $N_b$ vs $1/B$ (exhibiting good linear fit) extrapolates (in the limit
of $1/B\rightarrow 0$) to an $N$-axis intercept of $\sim-1/2$ for all $V_{bg}$’s studied. Such a $1/2$-intercept in LL index is again a hallmark of $\pi$ Berry’s phase for Dirac fermions and underlies the half-integer QHE.

As the bottom surface density can be extracted as $n_b = eB_F/h$ (where $B_F$ is the SdH oscillation frequency), the gate efficiency can be extracted from the linear fit of $B_F$ vs $V_{bg}$ to be $2.7 \times 10^{10}$ cm$^{-2}$/V (Fig. 13C), consistent with the result in Fig. 9. The matching of density extracted from QH and low-field Hall data also indicates that the carriers contributing to the transport is completely from the TSS in our BSTS. There are no other parallel conducting channels in our BSTS. However, the value $2.7 \times 10^{10}$ cm$^{-2}$/V is notably smaller than $7.3 \times 10^{10}$ cm$^{-2}$/V given by the simple capacitance of 300-nm thick SiO$_2$, possibly due to trapped charged impurities in the oxide or other screening effects. The extrapolated linear fit to $V_{bg} = -60$ V gives $B_F$ close to zero, consistent with expected vanishing carrier density of bottom surface near its DP.
2.8 Temperature dependence of the QHE

Fig. 14 shows backgate-tuned QHE ($R_{xx}$ and $R_{xy}$ vs $V_{bg}$) measured a fixed $B=31$ T in another BSTS device Sample B (thickness $t=80$ nm) at different temperatures ($T$). At low $T$ (4.5 K), we observe QH states corresponding to $\nu = 1$ (well-developed) and $\nu = 2$ (developing), interpreted (similarly to Sample A, Fig. 11) to $\nu_t + \nu_b = 1/2 + 1/2$ and $1/2 + 3/2$, respectively. The insensitivity of the quantization values to the sample thickness (changing by factor of 2 between samples A and B) confirm that the observed QHE arises from the surface of the intrinsic TI, and is not related to the “bulk” QHE observed previously (in highly-bulk-doped Bi$_2$Se$_3$, attributed to electronic decoupling between bulk quintuple layers with high Se vacancies) (Cao et al., 2012). As seen in Fig. 14A, increasing $T$ does not weaken substantially the QH states till at least ~10 K, and QH states are still observable up to 35 K, before disappearing above ~50 K. Fig. 14B and C show more quantitative analysis of the $T$ dependence of the QHE at $\nu = 1$ (both surface

![Fig. 14](image-url) Measurements were performed in Sample B, an 80-nm-thick exfoliated BSTS flake. (A) $V_{bg}$ dependent $R_{xx}$ and $R_{xy}$ measured at $B=31$ T and 4 different temperatures (0.35, 10, 35, and 50 K). Vertical dot-dashed lines indicate the position of QH state at total filling factor $\nu = 1$ (the shifts in $V_{bg}$ positions may reflect thermally induced carrier density change). (B) $\sigma_{xx}$ (log scale) as function of inverse temperature ($1/T$) at $\nu = 1$. Blue dashed line indicates a fit in higher temperature with $\sigma_{xx} \propto \exp(-E_{F}/kT)$. (c) $\Delta R_{xy}$ vs $R_{xx}$ at $\nu = 1$. Blue dashed line indicates a linear fit.
filling 0th LL). The $\sigma_{xx}$ (QH minimum) exhibits a thermally activated behavior $(\propto e^{-E_0/k_BT})$ at higher $T$ (Fig. 14B) with an extracted activation gap $E_0 \approx 5.6 \pm 0.3$ meV, but drops more slowly with decreasing $T$ at lower $T$. The $E_0$ value is notably smaller than the theoretical energy separation $v_F(2eB\hbar)^{1/2} \approx 60$ meV between the 0th and 1st LLs of ideal 2D Dirac fermions, for a typical Fermi velocity $v_F \approx 3 \times 10^5$ m/s estimated from ARPES measurements. This could stem from both disorder-induced Landau level broadening (such an effect was also seen in graphene with similar mobility) (Jabakhanji et al., 2014) and a finite thickness (side surface) effect. At even higher temperature ($>50K$), the bulk conduction starts to become significant and possibly suppress the QH states. The thermally induced change $\Delta R_{xy} = R_{xy}(T) - R_{xy}(4.5K)$ vs $R_{xx}$ at $\nu = 1$ follows an approximately linear relationship (with a slope of $-0.18$, Fig. 14C). The $T$ dependence behaviors of the QHE observed in our BSTS are quantitatively similar to those of the previously studied integer QHE in 2DEG.

2.9 Discussion

Well-developed “half-integer” QHE was previously observed for Dirac fermions (fourfold degenerate with 2 spins and 2 valleys) in graphene (Novoselov et al., 2005; Zhang et al., 2005), giving QHC $\sigma_{xy} = 4(N + 1/2)e^2/h$. A zero-gap HgTe quantum well was found to give a 2DES with single-valley Dirac fermions (thus considered “1/2-graphene,” with two spins), and exhibit well-defined QHE (Büttner et al., 2011). In contrast, a TI surface contains only 1 species of Dirac fermions (a single non-degenerate Dirac cone, sometimes referred to a “1/4- graphene”), resulting in a QHC of $(N + 1/2)e^2/h$. Two such degenerate surfaces contribute in parallel to give the observed $\sigma_{xy} = 2(N + 1/2)e^2/h$ (odd integer plateaus). The surface (top or bottom) index in our sample can be viewed as a pseudospin, which is not affected by external $B$ field. One can use a gate voltage (controlling the chemical potential difference between two surfaces, thus acting as a fictitious Zeeman field) to make the pseudospins degenerate (giving odd integers of QHC, Fig. 13) or non-degenerate (giving consecutive integers of QHC, Fig. 11), realizing a more tunable system. However, a complete understanding of some of the QH states (such as in the bipolar situation) and underlying mechanism would favor an independent gating of the densities and carrier types of both top and bottom surfaces. Further work on dual-gated BSTS devices would be presented in the next few sections.
3. Quantum transport in dual-gated 3D topological insulators

3.1 Dual-gated device fabrication

Similarly, as discussed in the last section, BSTS flakes (typical thickness \( \sim 50–200 \text{ nm} \)) are exfoliated (Scotch tape method) onto highly doped Si (p+) substrates (with 300 nm-thick SiO2 coating), and deposited with Cr/Au contacts after e-beam lithography. The BSTS flakes are then etched into standard Hall-bar geometry by Argon plasma (dry etching) to eliminate uncertainties in evaluating conductivities due to the influence of electrical contacts or the irregular shape of the samples. A thin flake of hexagonal boron nitride (h-BN, typical thickness \( \sim 10–40 \text{ nm} \)) is mechanically transferred on the BSTS flake to serve as a top-gate dielectric and a top-gate metal (Cr/Au) is deposited afterwards. The h-BN as a substrate or gate dielectric is known to preserve good electronic properties for graphene, resulting from the atomic flatness and relatively low density of impurities in h-BN (Dean et al., 2010). Temperature above 130°C is avoided during the fabrication process to prevent any heating induced doping effect. See Fig. 15 for a schematic of a dual-gated BSTS device and the fabrication procedures.

3.2 Independent gating of top and bottom surfaces

Qualitatively similar data are measured in multiple samples, while results from a typical sample C (channel length \( L = 9.4 \mu\text{m} \), width \( W = 4.0 \mu\text{m} \), with \( \sim 100 \text{ nm}-\) thick BSTS and 40 nm-thick h-BN as top-gate dielectric) are presented below unless otherwise noted. The carrier densities of the

Fig. 15 (A) A dual-gated BSTS device schematic. (B) Optical images and device fabrication procedure for a typical dual-gated device Sample C.
top and bottom surface of the BSTS flake are tuned by top-gate voltage $V_{tg}$ (applied to top-gate metal) and back-gate voltage $V_{bg}$ (applied to Si substrate), respectively. Fig. 16A and B show the double-gated electric field effect measured at $T=0.3$ K. The longitudinal resistivity $\rho_{xx}$ at magnetic field $B=0$ T (Fig. 16A) and Hall resistivity $\rho_{xy}$ at $B=1$ T (Fig. 16B) are plotted in color scale as functions of both top and bottom gate voltages ($V_{tg}$ and $V_{bg}$).

Fig. 16  (A) and (B) show 2D map of $\rho_{xx}$ at $B=0$ T and $\rho_{xy}$ at $B=1$ T as functions of $V_{tg}$ and $V_{bg}$ on sample C. The blue (red) dashed lines in (A) are guides to the eye for the top (bottom) surface DP. The 2D map is generated by data measured from $V_{tg}$ sweeps at a series of $V_{bg}$’s, with one example at $V_{bg}=3$ V shown in the inset of (B).
The extracted field effect and Hall mobilities are typically several thousands of cm$^2$/Vs. A minimum carrier density $n^* \sim 9 \times 10^{10} \text{cm}^{-2}$ per surface can be extracted from the maximum Hall coefficient (absolute value) $\sim 3.5 \text{k} \Omega/\text{T}$ (when both surfaces are slightly n-type or p-type) measured in Fig. 16B. A set of exemplary $V_{tg}$-sweeps with $V_{bg} = 3 \text{V}$ is shown in Fig. 16B inset. By adjusting $V_{tg}$ (or $V_{bg}$), the device can be gated through a $\rho_{xx}$ peak, identified as the charge neutrality Dirac point (DP) of the top (or bottom) surface, marked by the blue (or red) dashed lines in Fig. 16A. Gating through the DP, the carriers in the corresponding surface change from hole-like to electron-like (i.e., ambipolar), as evidenced by Hall measurements (Fig. 16B). The slight deviation of the two lines from being perfectly vertical and horizontal arises from the weak capacitive coupling between the top (bottom) surface and the back (top) gate (Fatemi et al., 2014). The crossing of these two lines corresponds to the double-DP (both surfaces tuned to DP) where $\rho_{xx}$ ($\sigma_{xx} = 1/\rho_{xx}$) reaches a global maximum (minimum). Within the gate voltage range used, the carriers predominantly come from the TSS and we observe relatively good particle-hole symmetry in the transport properties (e.g., the symmetrical appearance of $\rho_{xx}$ on both sides of DP in each surface in Fig. 16A and the similar absolute values of the positive and negative maximum Hall coefficient in Fig. 16B).

### 3.3 Gate efficiency and mobility of each surface

Now we present how to extract the gate efficiency and mobility for each surface of Device C (inter-surface capacitance coupling is neglected here). From the color map of $\rho_{xx}$ and $\rho_{xy}$, we identify the double Dirac point (at which both surfaces are tuned to Dirac point) located at $(V_{tg}, V_{bg}) \sim (-3.13 \text{V}, 2.5 \text{V})$. We performed line cuts for the longitudinal conductivity $\sigma_{xx} (= 1/\rho_{xx}$ at $B = 0$) and calculated density $n (= 1/e \rho_{xy} = 1/e R_{xy}$ at $B = 1 \text{T}$) color maps at fixed $V_{bg} = 2.5 \text{V}$ or $V_{tg} = -3.13 \text{V}$ (shown in Fig. 17). The gate efficiency for top and backgates are extracted to be $1.7 \times 10^{11}$ and $2.9 \times 10^{10} \text{cm}^{-2}/\text{V}$, respectively, from the linear fitting of $n$ vs the corresponding gate voltage on the electron doped side (some nonlinearity presents in the negative doping side due to the existence of both charge carrier types). The mobility for each surface can be extracted through Drude model

$$\mu = \frac{1}{e} \frac{d\sigma_{xx}}{dn} = \frac{1}{e} \frac{d\sigma_{xx}}{dV} / \frac{dn}{dV},$$

where we use $\frac{d\sigma_{xx}}{dV}$ obtained from the linear fitting near the Dirac point in Fig. 17A and C. The mobility for the top and bottom surfaces are then calculated to be of comparable values $\sim 2500$ and $\sim 2000 \text{cm}^2/\text{Vs}$, respectively.
3.4 Minimum conductivity at double Dirac point

We have studied six dual-gated BSTS devices with different thicknesses ($t$) and aspect ratios ($L/W$). These devices are measured at low temperatures ($T<2\,$K) and the results are repeatable after multiple thermal cycles. When both surfaces are tuned to DP, the minimum 2D conductivity $\sigma_{\text{min}}$ at $B=0\,$T exhibits relatively constant value $(3.8 \pm 0.1) e^2/h$ for all the devices measured (with the uncertainty representing 90% confidence interval), whose thicknesses range from 50 to 195 nm and $L/W$ range from 1.3 to 3.5 (Fig. 18). Our observation indicates that the conductivity at the DP for each major surface (top or bottom) is $\sim 2 e^2/h$, within the range of values $(2–5 \, e^2/h)$ reported by Kim et al. (2012a) on thin flakes of Bi$_2$Se$_3$ ($\sim 10\,$nm). The better consistency over multiple samples in our dual-gated BSTS devices may be attributed to the more insulating bulk (whose conduction is immeasurably small at low temperature) and uniformity of the exfoliated BSTS flakes, which are sandwiched between SiO$_2$ and h-BN to achieve
better device stability. The minimum conductivity at DP has also been discussed in graphene with considerable interest (Chen et al., 2008; Das Sarma et al., 2011; Martin et al., 2008; Miao et al., 2007; Nilsson et al., 2006; Tan et al., 2007).

For pristine graphene, a minimum conductivity $\sigma_{\text{min}}$ of $4e^2/\pi h$ is predicted and experimentally verified in ballistic graphene samples (Das Sarma et al., 2011; Miao et al., 2007). For graphene samples with more scattering, $\sigma_{\text{min}}$ was reported to have a universal value of $4e^2/h$ (Novoselov et al., 2005) while another experiment observed $\sigma_{\text{min}}$ that falls into a wide range $2–12e^2/h$ varying from sample to sample (Tan et al., 2007). Further experiments revealed that the minimum conductivity is strongly affected by carrier density inhomogeneities (puddles) induced by disorder on or near graphene (Chen et al., 2008; Martin et al., 2008), such as the absorbates or charged impurities in the substrates. In 3D TIs, one source of impurities likely relevant to the observed quasi-universal minimum conductivity in our dual-gated BSTS devices could be bulk defects (located near surface) (Skinner et al., 2013), such as those revealed in scanning tunneling microscopy studies (Beidenkopf et al., 2011).

3.5 Dual-gate tunable QHE

Now we start to focus on the transport phenomena in the quantum Hall (QH) regime under a high magnetic field $B$ perpendicular to the top and bottom surfaces in the dual-gated Sample C. Fig. 19 show in color scales
the longitudinal conductivity $\sigma_{xx}$ and Hall conductivity $\sigma_{xy}$ as functions of $V_{tg}$ and $V_{bg}$ at $B=18$ T and $T=0.3$ K in Sample C, with representative cuts at five different $V_{bg}$'s shown in (B) and (D). The $(\nu_t, \nu_b)$ labels (top, bottom) surface filling factors for corresponding quantum Hall states. The $\sigma_{xx}$ curves in (B) are shifted vertically (in consecutive step of $e^2/h$) for clarity (the corresponding zero $\sigma_{xx}$ levels are indicated by the same-colored horizontal dashed lines).

Fig. 19 (A) $\sigma_{xx}$ and (C) $\sigma_{xy}$ shown as 2D color maps, as functions of $V_{tg}$ and $V_{bg}$ at $B=18$ T and $T=0.3$ K in Sample C, with representative cuts at five different $V_{bg}$’s shown in (B) and (D). The $(\nu_t, \nu_b)$ labels (top, bottom) surface filling factors for corresponding quantum Hall states. The $\sigma_{xx}$ curves in (B) are shifted vertically (in consecutive step of $e^2/h$) for clarity (the corresponding zero $\sigma_{xx}$ levels are indicated by the same-colored horizontal dashed lines).
Dirac electrons or holes (see a schematic in Fig. 20). In Fig. 19D, different fixed $V_{bg}$’s (from $-17$ to $40$ V) set $\nu_b$ around consecutive half-integers $-3/2$, $-1/2$, $1/2$, $3/2$ and $5/2$ (such that the bottom surface contributes $\sigma_b = \nu_b^* e^2 / h$ to the total $\sigma_{xy}$), explaining the vertical shift of $e^2 / h$ at QH plateaux of consecutive $V_{tg}$ sweeps.

It is also notable that in Fig. 19, there are a few states with zero quantized Hall conductivity ($\sigma_{xy} = 0$, manifesting as white regions in Fig. 19C, separating the electron-dominated regions in red and the hole-dominated regions in blue) and non-zero $\sigma_{xx}$ minimum, marked by equal and opposite half-integer values of $\nu_t$ and $\nu_b$ thus total $\nu = 0$, for example $(\nu_t, \nu_b) = (-1/2, 1/2)$, $(1/2, -1/2)$ and $(3/2, -3/2)$. These states with total $\nu = 0$, exhibiting zero Hall plateaux (see also Fig. 19D), have non-zero $\sigma_{xx}$ minimum (Fig. 19A) but very large $R_{xx}$ maximum (see Section 3.6, Fig. 21).

### 3.6 Non-local measurements

To further characterize the observed QH and $\nu = 0$ states, we have performed nonlocal transport measurements of $R_{nl} (= V_{nl} / I$, $I$ is the current and $V_{nl}$ is the non-local voltage, see the schematic measurement setup in the inset of Fig. 21B) as functions of $V_{tg}$ and $V_{bg}$ at $B=18$ T and $T=0.3$ K and compared the results with the standard (local) measurements of the longitudinal resistance $R_{xx}$ (Fig. 21A). It’s intriguing that unlike other QH states typified by a zero or minimum in $R_{xx}$, the states with $\nu = \nu_t + \nu_b = 0$ (labeled by $(\nu_t, \nu_b)$ in Fig. 21A with $\nu_t = -\nu_b = 1/2$ or $3/2$) are accompanied by a $R_{xx}$ maximum. The best-developed $\nu = 0$ states are
those at \((\nu_t, \nu_b) = (-1/2, 1/2)\) or \((1/2, -1/2)\), where \(R_{xx}\) reaches 220 k\(\Omega\) \((\rho_{xx} \sim 100\, k\Omega)\), exceeding the resistance quantum \((h/e^2 \approx 25.8\, k\Omega)\) by an order of magnitude. The nonlocal \(R_{nl}\) also becomes very large \((\sim 100\, k\Omega)\) and the similar order of magnitude as \(R_{xx}\) at these two \(\nu = 0\) states while negligibly small at other \((\nu_t, \nu_b)\) states (see Fig. 21B and also the representative cuts in Fig. 21C).

The simultaneously large local and non-local resistance at \(\nu = 0\) states in the QH regime has been reported in other 2D electron-hole systems \((\text{Gusev et al., 2012; Nichele et al., 2014})\) and understood in a picture of dissipative edge channels. We emphasize that the pronounced \(R_{nl}\) signal cannot be explained from \(R_{xx}\) by a classical Ohmic non-local resistance from the stray current connecting the remote leads. Such a contribution \((\sim \rho_{xx} e^{-\pi L/W})\) would decay exponentially with \(L/W(\approx 2.4\) in our case), and be three orders magnitude smaller than the local \(R_{xx}\) (which is the case at \(B = 0\) T, Fig. 22).

As another comparison, the middle panel of Fig. 21C shows the cuts in Fig. 21A and B at \(V_{bg} = 3\, V\), crossing the double-DP (also zeroth LL) of both top and bottom surfaces at \((\nu_t, \nu_b) = (0,0)\), where we observe a relatively large peak in \(R_{xx}\) but significantly smaller \(R_{nl}\). Such a result is consistent with the extended state transport (at the center of zeroth LL) as the current flows through the bulk of the 2D surface.
From the color plots in Figs. 19 and 21, the parallelogram centered around \((\nu_t, \nu_b) = (\nu_t, \nu_b) = (1/2, 1/2)\) state is enclosed by boundaries representing \(N_t = 0\) and \(-1\), \(N_b = 0\) and \(1\) LLs. Similarly, \((\nu_t, \nu_b) = (1/2, -1/2)\) state is bounded by \(N_t = 0\) and \(1\), \(N_b = 0\) and \(-1\) LLs. We conclude that such a \(\nu = 0\) state can exist when the potential difference \(V\) between top and bottom surfaces (equivalently the energy separation between top and bottom surface DPs) is in the range of \(0 < |V| < 2E_{0-1}(\approx 2 \times 50 \text{ meV at } B = 18 \text{T})\), where \(E_{0-1}\) is the 0–1 LL separation of TSS Dirac fermions (Xu et al., 2014). The large energy scale of \(E_{0-1}\) can help make the \(\nu = 0\) and \(\nu = \pm 1\) QH states observable at significantly elevated temperatures as demonstrated below.

### 3.7 Temperature dependence study

We have studied the temperature \((T)\) dependence of the QHE and \(\nu = 0\) states from 0.3 to 50K at \(B = 18\text{T}\) (Fig. 23). At each temperature, the bottom surface density is tuned by \(V_{bg}\) to set \(\nu_b\) near 1/2 (dashed lines) or \(-1/2\) (solid lines), and the peaks in local \(R_{xx}\) and nonlocal \(R_{nl}\) correspond to \((\nu_t, \nu_b) = (-1/2, 1/2)\) or \((1/2, -1/2)\), respectively. The \(R_{xx}\) peaks \((\geq 150 \text{ k}\Omega)\) are seen to be more robust up to the highest temperature \((T = 50\text{K})\) measured while \(R_{nl}\) peaks decrease rapidly (approximately linearly in \(T\), shown in the inset of Fig. 23B) with increasing \(T\) and is nearly suppressed above 50K.

![Fig. 22](image) Local resistance \(R_{xx}\) (A) and non-local resistance \(R_{nl}\) (B) as functions of \(V_{tg}\) at \(V_{bg} = 3 \text{V}\) and \(B = 0\text{T}, T = 0.3\text{K}\). The non-local resistance \(R_{nl}\) shown here is multiplied by a factor of hundred.
Fig. 23 (A) $R_{xx}$ and (B) $R_{nl}$ measured in sample C as functions of $V_{tg}$ for different temperatures at $B = 18 \, T$, where $V_{bg}$ is chosen to set $\nu_b$ at $1/2$ (dashed lines) and $-1/2$ (solid lines), respectively. The inset of (B) shows approximately linear dependence of $R_{nl}$ peaks at $\nu = 0$ on temperature.
We also show the $T$-dependence of $\sigma_{xx}$ and $\sigma_{xy}$ at $(\nu_t, \nu_b) = (1/2, 1/2), (1/2, -1/2), (1/2, 1/2)$ and $(-1/2, -1/2)$ in Fig. 24A and B. The $\sigma_{xy}$ maintains good quantization at $\nu e^2/h$ ($\nu = 0, \pm 1$) up to $T = 50K$ while $\sigma_{xx}$ increases with $T$. The $\sigma_{xx}$ for $\nu = \pm 1$ states is found to show thermally activated behavior at high temperatures (Xu et al., 2014), where the finite $\sigma_{xx}$ is attributed to the thermal excited 2D surface or 3D bulk carriers. Such carriers can shunt the edge-state transport and suppress the nonlocal $R_{nl}$ response at high $T$ (Gusev et al., 2012). We also note that the $\sigma_{xx}$ vs $T$ curves for $\nu = 0$ and $\nu = \pm 1$ states follow the similar trend and have approximately constant separation. We find the averaged separation $\Delta\sigma_{xx} = 1/2 \times \left[ \sigma_{xx}(-1/2, 1/2) + \sigma_{xx}(1/2, -1/2) - \sigma_{xx}(1/2, 1/2) - \sigma_{xx}(-1/2, -1/2) \right]$ to be largely $T$-independent with a value of $(0.27 \pm 0.01)e^2/h$, which we attribute to the conductivity of the quasi-1D dissipative edge channel.

3.8 Dissipationless and dissipative edge states

In this section, I would describe how to understand the QHE in 3D TIs and discuss some remaining open questions.

The surface of a slab-like 3D TI has a closed 2D manifold and the topology of a sphere, thus geometrically different from conventional 2D electron gases with a planar structure (Lee, 2009). The QHE, arising from LL confinement near the edges, is not obvious how it works and whether it exists in a 3D TI (we consider the bulk to be totally insulating/empty here for simplicity) (Lee, 2009; Vafek, 2011; Zhang et al., 2012). In such slab-like 3D TI samples under perpendicular magnetic field, the side surface only
experiences an in-plane field that has no orbital effect while top and bottom surfaces experience magnetic flux along and opposite to the surface normal, respectively. Of course, in the thick limit (sample thickness $t \gg l_B$, where $l_B = \sqrt{\hbar/eB}$ is the magnetic length), the energy spectrum on side surfaces would approach the normal Dirac-like continuum for TSS (Vafek, 2011). In this case, the delocalized states on side surfaces would inevitably short opposite edges (namely side surfaces) and spoil the QHE in a normal Hall-bar structure (Lee, 2009; Vafek, 2011). A “Corbino” structure is proposed to be able to give QHE in this scenario (Lee, 2009).

In the thin sample limit ($t \ll l_B$, without considering any top and bottom surface hybridization effect here), the side surfaces can be viewed as a quai-1D domain boundary that separates the top and bottom surfaces with B pointing outward and inward, respectively, thus can support dissipationless QH edge states (Chu et al., 2011). In this case, when the top and bottom surfaces are doped to the same carrier type (either $n$ or $p$), the corresponding QH edge states (on the side surface) would have the same chirality and give the total $\sigma_{xy} = \nu e^2/h = (\nu_t + \nu_b)e^2/h$, restricted to integer multiples of $e^2/h$. Meanwhile the longitudinal resistivity ($\rho_{xx}$) or conductivity ($\sigma_{xx}$) vanishes because of the ballistic edge states.

When the sample thickness is in the intermediate regime ($t \sim l_B$), it’s not clear whether the well-defined QHE should exist in a conventional Hall-bar (or multiple terminal) structure as there some debates (Brey and Fertig, 2014; Chu et al., 2011; Lee, 2009; Vafek, 2011; Zhang et al., 2012). In this case, the finite-size confinement-induced gap ($\hbar v_F/2t$, which is the energy separation between different sub-bands on side surfaces) is comparable to the $0\rightarrow 1$ LL gap $v_F\sqrt{2eB\hbar}$ on the top and bottom surfaces. A schematic band structure belong this scenario is depicted in the upper panel of Fig. 25A, which shows the Fermi energy $E_F$ residing inside the $0\rightarrow 1$ LL gap and corresponds to $\nu_t = 1/2$ and $\nu_b = 1/2$ (the two surfaces are set to be degenerate with $V=0$ here). The Fermi level passes though some non-chiral edge modes (in pairs) and chiral edge modes. It has been suggested that the QHE should be spoiled by those non-chiral edge modes in a standard Hall-bar measurement (Bruene et al., 2011; Vafek, 2011; Zhang et al., 2012). In our measurement setup (a conventional Hall-bar structure), the contacts connect to the top, bottom and side surfaces, all of which are probed simultaneously. However, for BSTS sample in a wide range of thickness ($\sim 10$ to $\sim 100$ nm, from similar to much larger than the magnetic length $l_B \sim 6$ nm at $B=18$ T), we do observe well-defined QHE. The behavior is similar to what expected for the scenario of $t \ll l_B$. This suggests that the non-chiral edge modes might be canceling each other and become
localized when net chiral modes exist, restoring the QHE (Brey and Fertig, 2014). However, in order to show well-defined QHE, the upper limit of \( t/l_B \) is not known yet. Certainly there has no QHE observed in bulk 3D TI samples (\( t > \text{a few } \mu \text{m} \)).

When the two surfaces have opposite carrier types but one of them dominates, well-defined QH states with \( \nu = \nu_t + \nu_b \) may still be observed, such as the \((-1/2, 3/2)\) state with \( \sigma_{xy} = (-1/2 + 3/2) e^2/h = e^2/h \) and vanishing \( \rho_{xx} \). Previous studies in InAs/(AlSb)/GaSb heterostructure based electron-hole systems (Mendez et al., 1985; Suzuki et al., 2003) also revealed QH effect with \( R_{xy} = h/(\nu e^2) = h/(\nu_e - \nu_h) e^2 \) (\( \nu_e \) and \( \nu_h \) are electron and hole filling factors, both are positive integers) and vanishing \( R_{xx} \) when the AlSb barrier (separating electron and hole gases) is sufficiently thin to enable electron-hole hybridization. Despite the phenomenologically similarities, our QH system is distinctive in the sense that the spatially separated electrons and holes residing on the top and bottom surfaces have half-integer filling factors, and the hybridization only happens at the side surface.

**Fig. 25** (A) and (B) Schematics of surface band structures (upper), showing Landau levels from top and bottom surfaces (blue and red) in the middle of the sample transitioning into side surface sub-bands at sample edge, and edge states in a slab-shaped sample (lower) for \( \nu = 1 \) (A) and \( \nu = 0 \) (B) states. The dashed line indicates a representative Fermi level \( E_f \) and circles in (A) upper label chiral edge modes.
For the \((\nu_t, \nu_b) = (-1/2, 1/2)\) or \((1/2, -1/2)\) state, the carrier density on the top and bottom surfaces are opposite. Since \(E_f\) is within the LL gap on both the surfaces, the finite residual \(\sigma_{xx}\) and large \(R_{nl}\) we observed are indicative of dissipative edge transport. We show a schematic energy spectrum (Morimoto et al., 2015a) of this \(\nu = 0\) state with \(V\) slightly smaller than \(E_{0-1}\) in the upper panel of Fig. 25B, where the Fermi level \(E_f\) resides between the \(N_t = -1\) and 0 LL of top surface (marked in blue), thus \(\nu_t = -1/2\), and also between the \(N_b = 0\) and 1 LL of bottom surface, thus \(\nu_b = 1/2\). Overall, such energy spectrum represents a \((\nu_t, \nu_b) = (-1/2, 1/2)\) and \(\nu = 0\) state. The \(E_f\) crosses an even number (only two shown in this illustrative example in Fig. 25B) of counter-propagating edge modes (arising from sub-bands of the quasi-1D side surface). The disorder can cause scattering and local equilibrium between the counter-propagating modes, giving rise to non-chiral dissipative transport (depicted by a series of conducting loops that can hop between adjacent ones in Fig. 25B lower panel) on the side surface with a large and finite resistance. While the energy spectrum (Fig. 25B) is expected to have a gap (\(\Delta\)) near the edge (due to the hybridization between top (marked with blue) and bottom (red) surface zeroth LLs and approximately the finite-size confinement-induced gap \(\approx h\nu_F/t\approx 10\text{meV}\) opened at DP of the side surface (Morimoto et al., 2015a)), we did not observe a truly insulating state with vanishing \(\sigma_{xx}\) and diverging \(R_{xx}\) (Figs. 21A and 23A). This is likely due to the disorder potential (spatial fluctuation of DP (Beidenkopf et al., 2011)) comparable or larger than \(\Delta\) and thus smearing out this gap (effectively \(E_f\) always crosses the non-chiral edge modes). It would be an interesting question for future studies to clarify whether the weak \(T\)-dependence (at \(T < \sim 50\text{K}\)) of the observed conductance (Fig. 24A), similar to the behavior reported in InAs/GaSb based electron-hole systems (Takashina et al., 2003), may indicate an absence of localization (Anderson, 1958; Dunlap et al., 1989; Kurkijärvi, 1973; Lee, 1984) in such quasi-1D resistive edge channels.

Several recent theories have pointed out that the \(\nu = 1/2–1/2 = 0\) state in the TI QH system may bring unique opportunities to realize various novel physics. It has been suggested that both the \(\nu = 0\) state in TI QHE and an analogous quantum anomalous Hall (QAH) state with zero-Hall-conductance plateau in a magnetic doped TI around the coercive field can be used as platforms to observe the TME effect (Morimoto et al., 2015b; Wang et al., 2015), where an electric (magnetic) field induces a co-linear magnetic (electric) polarization with a quantized magnetoelectric polarizability of \(\pm e^2/2h\).
recently observed in the QAH case in ultrathin (few-nm-thick) films of Cr$_x$(Bi,Sb)$_{2-x}$Te$_3$ at low temperature ($<1$K) (Feng et al., 2015; Kou et al., 2015). In comparison, our samples have much larger thickness (> ~ 50 nm, suggested to be preferable for better developed TME effect (Baasanjav et al., 2014; Wang et al., 2015)), and our $\nu = 0$ state survives at much higher temperatures (~50 K). It has also been proposed that excitonic condensation and superfluidity can occur in thin 3D TIs at the $\nu = 0$ state in QH regime (Tilahun et al., 2011) (in addition to at zero $B$ field (Seradjeh et al., 2009)) induced by spontaneous coherence between strongly-interacting top and bottom surfaces. In future studies, much thinner samples are likely needed to investigate the possibility of such exciton superfluidity.

4. Inter-surface hybridization in thin BSTS

In a relatively thick 3D TI film (thickness $t \gg 10$ nm), the top and bottom surfaces are well separated and their corresponding topological SS Dirac cones are gapless with opposite spin helicities. When the sample is thin enough (typically $\lesssim 10$ nm) to enable hybridization between the two surfaces, a gap $\Delta_0$ is opened at the Dirac point (even though the TRS is still preserved). The SS band structure acquires massive Dirac dispersion $\epsilon = \pm \sqrt{(\hbar v_F k)^2 + (\Delta_0/2)^2}$ (Zhang et al., 2010). Schematics are shown in Fig. 26. Such a crossover of 3D TIs to the two-dimensional (2D) limit is little explored by electronic transport measurements in surface-dominant 3D TI systems (such as BSTS).

Fig. 26 (A and B) Schematics of the two surface Dirac cones with opposite spin helicities for a relatively thick TI sample (A) and the gapped surface band structure due to the hybridization of top and bottom SS for a very thin TI sample (B). The gap size at zero field is denoted as $\Delta_0$. 
In this section, we discuss about the inter-surface hybridization effect in dual-gated BSTS thin samples, upon which we can electrostatically control both surfaces to the charge neutrality point. By reducing the thickness of the BSTS flake, the capacitive coupling between the two surfaces becomes stronger (Fatemi et al., 2014; Kim et al., 2012a). As it can be seen in the color map of 2D conductivity ($\sigma_{xx} = L/(W^*R_{xx})$) vs $V_{tg}$ and $V_{bg}$, the black and white dashed lines tracing the Dirac points of top and bottom surfaces cross and tend to merge together when the thickness $t$ is reduced from 80 to 17 nm (Fig. 27A and B). Further reducing $t$ to $\sim 10$ nm results in the Dirac points from the two surfaces to become indistinguishable (Fig. 27C). When the sample is only a few nm thick (e.g., Device N4 with $t=6$ nm in Fig. 27D), a hard gap opens, as indicated by the highly insulating (two-terminal conductivity $\sigma \ll e^2/h$) blue region at $T=1.6$ K.

![Fig. 27](image-url) (A–D) Measured conductivities (in color scale, with contours) as functions of top and backgate voltages ($V_{tg}$ and $V_{bg}$) in four representative devices with decreasing thickness ($t$). In relatively thick samples N1 (A) and N2 (B), the dashed black and white lines are guides to the eyes indicating the top and bottom surface charge neutrality Dirac points, respectively. Data are measured at temperature $T=0.35$ K in (A–C), and at $T=1.6$ K in (D). All data in this paper are measured with a 4-terminal configuration except for devices thinner than 7 nm (such as N4), which are measured with a 2-terminal configuration.
The minimum conductivity $\sigma_{\text{min}}$ and maximum resistivity $\rho_{\text{max}}$ ($=1/\sigma_{\text{min}}$) are reached when the two surfaces are gated simultaneously to charge neutrality points or Dirac points. In Fig. 27A, we plotted $\rho_{\text{max}}$ as a function of temperature ($T$) for a few representative samples. At $t > 10$ nm, $\rho_{\text{max}}$ shows a metallic behavior ($d\rho_{\text{max}}/dT > 0$), implying a zero or negligible gap. However, at $t < 10$ nm, a strongly insulating behavior ($d\rho_{\text{max}}/dT < 0$) is observed. Around $t = 10$ nm, different samples can behave differently. For example, while device N5 exhibits an insulating behavior, another device N3 exhibits a non-monotonic temperature dependence with its $\rho_{\text{max}}(T)$ close to $h/e^2$ and separating curves with metallic and insulating behaviors. It is consistent with the general observation from previous studies that the critical resistivity for a metal-insulator transition in 2D electron systems is on the order of the resistance quantum $h/e^2$.

In Fig. 28B, we replotted four representative data sets from Fig. 28A as $\sigma_{\text{min}}$ vs $1/T$. The size of non-zero effective transport gap $\Delta_0^*$ in thinner BSTS samples is estimated from thermal activation fit with $\rho_{\text{max}} \propto e^{\Delta_0^*/2k_BT}$ or $\sigma_{\text{min}} \propto e^{-\Delta_0^*/2k_BT}$ in appropriate temperature ranges, beyond which other conduction mechanisms need to be considered to describe the temperature-dependent conduction. For example, in samples with $t = 10$ nm (Device N3 and N5) and 8 nm, the thermal activation fit has been applied to the temperature range between ~3 to ~30 K, where an insulating behavior is observed, to extract $\Delta_0^*$ (note that it could be smaller than the real hybridization gap $\Delta_0$ due to smearing effect of potential fluctuations at different positions). This temperature range is chosen for the following reasons. Usually below 3 K, $\sigma_{\text{min}}$ tends to deviate from a single thermal activation fit. As shown in Fig. 28C of a typical example N5, we can fit the temperature dependent data down to 0.3 K with an empirical formula $\sigma_{\text{min}} = C_1 e^{-T_1/T} + C_2 e^{-(T_2/T)^{1/3}}$, which gives $T_1 = \sim 8$ K and $T_2 = \sim 0.02$ K. The second term has the form of variable range hopping (VRH) for 2D (we caution that other forms may be used that also give reasonable fits), suggesting multiple conduction mechanisms at lower temperatures. Above ~30 K, Dirac points usually shift and the peak become much broader in gate voltage sweeps, indicating additional conducting channels such as impurities or bulk states.

In Fig. 28D, we showed the thermal activation fit for another 6 nm-thick device in the temperature range of 76 to ~300 K, below which $\sigma_{\text{min}}$ becomes too small to be detected accurately due to limitation of our instruments. The good linearity of $\sigma_{\text{min}}$ in logarithmic scale vs $1/T$ indicates a hard gap opening in this device and the fitting gives a gap size $\sim 160$ meV, which is much larger than that determined by ARPES measurements.
on 6 nm Bi$_2$Se$_3$ grown by molecular beam epitaxy (MBE, surface gap barely vanishes at 6 nm) (Zhang et al., 2010) and the theory prediction for Bi$_2$Se$_3$ (a few meV at 6 nm) (Linder et al., 2009; Liu et al., 2010). A much enhanced gap size has also been reported previously on exfoliated Bi$_2$Se$_3$ thin flakes (Cho et al., 2011). Given the material difference and better isolated surface conduction, the gap size in our thin BSTS films grows monotonically with reducing thickness and roughly follows an exponential function $\Delta_0^* \propto e^{-0.7t/[\text{nm}]}$ (Fig. 29 inset). The $\Delta_0^*$ grows by about an order
of magnitude when $t$ is reduced by $\sim 1.4$ nm, comparable to what was found for Bi$_2$Se$_3$ (Kim et al., 2013; Zhang et al., 2010). We don’t have evidence of oscillatory behavior of $\Delta_0^*$ vs $t$ and did not observe any signs of quantum spin Hall effect (Linder et al., 2009; Liu et al., 2010; Lu et al., 2010).

Fig. 29 shows $\sigma_{\text{min}}$ at base temperatures ($T \leq 1.6$ K) for samples with various thicknesses. At large $t$ ($> 20$ nm), saturates around a value close to $4e^2/h$ (Xu et al., 2016). The $\sigma_{\text{min}}$ starts to decrease below $20$ nm and drops abruptly to zero below $\sim 10$ nm. The gap size $\Delta_0^*$ extracted from thermal activation is plotted on the right axis of Fig. 28.

Our data suggest that a measurable transport gap $\Delta_0^*$ (presumably driven by the inter-surface hybridization) opens at the Dirac point below a critical thickness $t_c = 10 \pm 1$ nm in our samples.

5. In-plane magnetic field effect

5.1 Effect of an in-plane magnetic field on relatively thick samples

Resistances of the thicker and thinner samples respond to the in-plane magnetic field differently at low temperatures. For consistency, the samples are mounted with current direction parallel to $B$ (unless otherwise specified).
We have measured multiple samples by either sweeping $V_{tg}$ (with $V_{bg}$ carefully tuned and then fixed at voltages such that these $V_{tg}$ sweeps go through maximum resistivity $\rho_{\text{max}}$) at different in-plane $B$ fields, or measuring $\rho_{\text{max}}$ vs in-plane $B$ at fixed gate voltages. For relatively thick samples such as Device N2 (dual-gating effect at zero field seen in last Chapter) with $t = 17\,\text{nm} > t_c$, the in-plane field up to $\sim 31\,\text{T}$ only induced a relatively small positive MR of $\sim 40\%$ (Fig. 30, noting $\rho_{\text{max}}(B)$ is approximately proportional to $B^2$ at low fields and proportional to $B$ at higher fields). We also noticed that with increasing $B$ field, there is a crossover from the metallic to non-metallic behavior (seen from the temperature dependence of the data at fixed fields, with two examples at 0 and 31.5 T shown in the inset of Fig. 30D) at 6 T.

In relatively thick samples, we shouldn’t expect any MR effect with the in-plane magnetic field as its contribution to the Hamiltonian can be

![Fig. 30](image)

Fig. 30  (A and B) Resistivity of Device N2 measured as a function of the top-gate voltage ($V_{tg}$) at various in-plane $B$ fields at $T = 0.3\,\text{K}$, where (A) is the raw data set while (B) is the generated 3D contour plot. The corresponding $V_{bg}$’s are carefully tuned such that the $V_{tg}$ sweeps go through the global resistivity maximum $\rho_{\text{max}}$ (C and D) The $\rho_{\text{max}}$ (and the corresponding MR, right axis) measured (at fixed $V_{bg}$ and $V_{tg}$ labeled in the figure) as a function of in-plane $B$ field measured at $T = 0.3\,\text{K}$ (C) and multiple temperatures (D).
eliminated by a gauge transformation. The small positive MR observed here are of interests for future studies and could be related to magnetic field induced spin-flip scattering process (Taskin et al., 2017).

5.2 Large negative MR in thinner samples

In this section, we’ll mainly focus on the in-plane measurements on thinner devices N3 and N5 (see an optical image in Fig. 31A and B) both with thickness around 10 nm. At zero field, Device N5 shows maximum resistivity $\rho_{\text{max}}$ near $V_{\text{bg}} \approx 30 \text{ V}$ and $V_{\text{tg}} \approx 1 \text{ V}$ (note that depending on gate sweep histories, the exact peak position could vary slightly) at zero field and base temperature $T=0.3 \text{ K}$.

At small in-plane fields ($<5 \text{ T}$), devices N3 and N5 also show some positive MR (for N5, we also observed an additional tiny cusp with negative MR near 0 T, see Fig. 32). Such low-field features in thinner devices disappear when we increase the temperature to just a few Kelvin (see Fig. 33), thus are attributed to phase coherent transport (Liao et al., 2015; Lin et al., 2013). In the following, we mainly focus on the higher magnetic field data showing a giant negative MR that has only been observed in thin samples with insulating behavior. For example, in Device N3 (Fig. 32A and B), $\rho_{\text{max}}$ drops dramatically above $\sim 5 \text{ T}$ and saturates at high field ($\sim 30 \text{ T}$).

![Fig. 31](image_url) (A and B) Optical images for the BSTS device N5 before and after top-gate metal deposition. The BSTS flake is etched into a Hall-bar structure and covered with a hexagonal boron nitride (h-BN) flake ($\sim 24 \text{ nm}$ thick) as top-gate dielectric. (C) The $\rho_{xx}$ as a function of $V_{\text{bg}}$ and $V_{\text{tg}}$ for Device N5. The behavior is similar to that of Device N3 with resistivity peak along the total charge neutral line (dashed white line). However, Device N5 is initially more n doped and has a larger $\rho_{\text{max}}$ ($\sim 10 \text{ h/e}^2$) as identified from the maximum resistivity peak in the line cut (shown in the inset) along the dashed line near $V_{\text{bg}} = 34 \text{ V}$. 
Fig. 32  (A and C) Resistivity of three representative devices (N3 and N5) measured as a function of the top-gate voltage ($V_{tg}$) at various in-plane $B$ fields at $T=0.3$ K. The corresponding $V_{bg}$’s are carefully tuned such that the $V_{tg}$ sweeps go through the global resistivity maximum max. (B,D) The $\rho_{max}$ (and the corresponding MR, right axis) measured (at fixed $V_{bg}$ and $V_{tg}$ labeled in the figure) as a function of in-plane $B$ field for the two devices (N3 and N5, respectively).

Fig. 33  (A) The $\rho_{max}$ of device N5 ($t\sim 10$ nm) as a function of temperature at different in-plane $B$ from 11.4 up to 45T. Inset shows the corresponding thermal activation fittings, while the extracted $\Delta_B^*$ is plotted (filled circles) as a function of $B$ in (B) along with data measured from another (second) cool down. (B) The linear fits for both cool downs (solid and dash lines, respectively) indicate a gap closing at $B_c$ between 35 and 40T, consistent with the inset showing the converging (at $B_c = 36$ T) of all the linear fittings from Ln[(2e^2/h)$\rho_{max}$] vs $B$ at different temperatures.
to \( \sim 0.45h/e^2 \). Notably for the more insulating sample N5 (Fig. 32C and D), \( \rho_{\text{max}} \) drops by a factor of \( \sim 20 \) (giving an MR \( \sim 95\% \)) from a very resistive value of \( \sim 12h/e^2 \) at \( B = 0 \) T to a value \( \sim 0.55h/e^2 \) again close to \( h/2e^2 \) at \( B = 45 \) T.

5.3 In-plane magnetic field induced insulator-semimetal transition

We performed systematic \( V_{\text{tg}} \) sweeps (fixed \( V_{\text{bg}} = 30 \) V) to extract \( \rho_{\text{max}} \) with in-plane \( B \) fields from 11.4 to 45 T in Device N5 at various temperatures. As shown in Fig. 33A, the insulating behavior of \( \rho_{\text{max}}(T) \) is strongly suppressed at higher fields. At the highest field (45 T), \( \rho_{\text{max}} \) saturates to a value close to \( \sim h/2e^2 \) and becomes relatively insensitive to temperature. We estimated the thermal activation gap \( \Delta_B^\ast \) at each \( B \) field from the slope of \( \text{Ln}[(2e^2/h)\rho_{\text{max}}] \) vs \( 1/T \) in the temperature range of 3 to 22 K (Fig. 33A inset) and plotted it vs the corresponding \( B \) in Fig. 33B, which also displays the gap size measured in another (second) cool down for \( B \) up to 18 T. The gap size \( \Delta_B^\ast \) is found to differ slightly over different cool downs but exhibits a similar linear dependence on \( B \) in the intermediate field range (5 T \( \sim 30 \) T).

Extrapolating the linear fits in Fig. 33B to zero suggests that the gap would close at a critical field \( (B_c) \) between 36 to 40 T, around which we observe the sample (N5) to become metallic \( (d\rho_{\text{max}}/dT > 0 \) see Fig. 33A) below \( T \sim 2 \) K. However, some non–metallic behavior \( (d\rho_{\text{max}}/dT < 0) \) can still be observed between 2 to 22 K even at the highest fields (Fig. 33A) and fitted to a thermal activation, giving \( \Delta_B^\ast \) that deviates from the red solid line (Fig. 33B). A non–metallic behavior under large in–plane magnetic fields was also observed in gapless samples such as N2 with \( t = 17 \) nm (Fig. 30D) The reason for this behavior remains to be understood (note a hybridization gap opening is not expected in such thick samples). We have also verified that \( \text{Ln}[(2e^2/h)\rho_{\text{max}}] \) is linear with \( B \) (\( \sim 25 \) T) at different temperatures and all the fitted lines converge to a critical field of \( B_c \sim 36 \) T (inset of Fig. 33B). This also suggests \( \Delta_B^\ast \propto (B_c - B) \), with a saturation resistivity \( \sim h/2e^2 \) (when \( \Delta_B^\ast \sim 0 \)) and gap closing at \( B_c \sim 36 \) T at fixed temperatures.

5.4 Discussion

Such large negative MR and the corresponding insulator–semimetal transition have not been reported before in thin nonmagnetic 3D TIs with surface dominant conduction. The relatively small negative MR previously shown
in Bi$_2$Se$_3$ with in-plane magnetic fields are attributed to bulk carriers and can be understood by semiclassical Boltzmann formalism considering the Berry curvature and orbital moment of the bulk states (Dai et al., 2017). Our observations of distinct behavior between thick and thin BSTS samples may be interpreted in terms of a theoretical prediction by Zyuzin et al. (2011). Generally, in thick TIs the in-plane magnetic field $B$ (set to be along the $x$ direction) can introduce opposite shifts (along $k_y$) of top and bottom surface Dirac cones in the momentum space. This does not produce any MR effect in thick 3D TIs but will prevent the two Dirac points from annihilation and eliminate the hybridization gap in thin TIs (schematics shown in Fig. 34A). Semiclassically, a spin-helical SS electron with spin orientated along the $B$ field (thus spin magnetic moment $-g\mu_B/2$, with $g$ being the in-plane spin $g$-factor and $\mu_B \approx 0.058\text{meV/T}$ the Bohr magneton) moves

![Fig. 34](A) Predicted evolution of the surface band structure from a trivial insulator toward a 2D topological semimetal (TSM), upon increasing in-plane $B$ field in a thin TI sample with hybridized SS. The gap $\Delta_B$ decreases linearly with $B$ and is closed when $\Delta_0 = |E_B|$, where $E_B = (g\mu_B - e\nu_F)B$. Red/blue arrows depict spin textures of the gapless Dirac cones, separated in the momentum space for 2D TSM, respectively. (B) Schematic of the orbital motion (normal to $B$ and $x$ direction) of spin-helical SS electrons for a 3D TI film.
clockwise around the circumference (Fig. 34B) with orbital magnetic moment (parallel to $B$ field) $\mu_{orb} = \frac{e}{2l(w+t)/v_F} \cdot wt \approx etv_F/2$ (noting width $w > t$ in our samples) (Minot et al., 2004). Both the spin and orbital magnetic moments couple to the $B$ field, giving rise to an effective Zeeman energy $E_B = g_{eff} \mu_B B = (g\mu_B - etv_F)B$ (total effective $g_{eff} = g - etv_F/\mu_B$, the second term can also be considered as due to the Aharonov-Bohm phase gradient between the two opposite surfaces). In thin TIs with hybridization gap $\Delta_0$ (at zero $B$ field), one can show (Zyuzin et al., 2011) that the massive Dirac band is spin-split by the above “effective Zeeman energy,” shrinking the hybridization gap linearly with $B$ as $\Delta_B = \Delta_0 - |E_B|$. The gap vanishes at a critical field $B_c$ ($|E_B| = \Delta_0$), where the dispersion near $k = 0$ becomes quadratic along $k_y$ and linear along $k_x$ (see Fig. 34A). With further increasing $B (> B_c)$, two Dirac points are restored and separated by $\sqrt{E_B^2 - \Delta_0^2/\hbar v_F}$ along $k_y$. This gives rise to a distinct 2D TSM that is topologically stable as long as translational symmetry is preserved (Zyuzin et al., 2011).

The above mechanism can qualitatively explain the trend we observed in experiments. However, the slope of the linear fitting yields the gap closing rate $|E_B| \approx 0.02 \text{meV/T}$ (corresponding to $|g_{eff}| \approx 0.33$). This is nearly two orders of magnitudes smaller compared with $|E_B| \approx 1 \text{meV/T}$ estimated for a 10-nm sample by Zyuzin et al., 2011, which assumed $g = 2$, leading to a negligibly small contribution from the Zeeman effect. Our results imply a large $g$, giving rise to a Zeeman term $(g\mu_B B)$ that is comparable with the orbital term $(etv_F B)$. Thus, the two nearly cancel to give a small $|E_B|$, which can be used to extract the in-plane SS $g$-factor if the Fermi velocity $v_F$ is known. Assuming a typical $v_F = 3.5 \times 10^5 \text{m/s}$ for topological SS with purely linear dispersion, we get a $g$-factor of $\sim 60$. In actual 3D TI materials such as BSTS, the surface Dirac cone contains substantial nonlinearity that can be described by a quadratic mass term added to the SS Hamiltonian. Subsequently, a reduced $v_F = 1.3 \times 10^5 \text{m/s}$, which describes the linear part in the Hamiltonian, yields a $g$-factor of $\sim 20$. It has been pointed out that the Zeeman coupling of the SS carriers can be highly anisotropic (Chu et al., 2011). In previous experiments, only an out-of-plane SS $g$-factor is determined and found to vary significantly in different TI materials (Fu et al., 2016; Taskin and Ando, 2011). Our study provides a method to extract the in-plane $g$-factor of SS carriers. We have to note that in our experiments, the gap extracted from thermal activation is an effective transport gap ($\Delta_B^*$) and can be smaller than the real band gap ($\Delta_B$) due to
disorder-induced smearing, namely $\Delta B^* = \Delta B - \delta$, where $\delta$ is a correction due to the potential fluctuations (likely to be on the order of several meV or higher (Nam et al., 2019)) in the system. Therefore, the observed apparent metallic behavior ($\Delta B^*$ reaching 0) in Device N5 above $B_c \sim 36$ T does not necessarily indicate the realization of the 2D TSM phase, which requires closing the real gap $\Delta B$ and possibly much larger magnetic field than $B_c$ (noting the relatively small gap closing rate of 0.02 meV/T in light of the estimated $\delta \sim$ meV in our BSTS samples). It might be easier to realize the 2D TSM phase (at lower B field) in other TI systems with a smaller or even negative g-factor (so the gap closing rate can be much larger than that in our samples). It would also be interesting for future studies to clarify whether the saturation resistivity $\sim h/2e^2$ is related to the modification of band structure and magnetic field induced spin-flip scatterings (Lin et al., 2013).

References


