Disentangling surface and bulk transport in topological-insulator p-n junctions

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: BACKES, D. ... et al, 2017. Disentangling surface and bulk transport in topological-insulator p-n junctions. Physical Review B, 96 (12), 125125.

Additional Information:

- This paper was published in the journal Physical Review B and the definitive published version is available at https://doi.org/10.1103/PhysRevB.96.125125.

Metadata Record: https://dspace.lboro.ac.uk/2134/34127

Version: Accepted for publication

Publisher: © American Physical Society

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Disentangling surface and bulk transport in topological insulator p-n junctions

Dirk Backes,1,* Danhong Huang,2 Rhodih Mansell,1 Martin Lanius,3 Jörn Kampmeier,3 David Ritchie,1 Gregor Mussler,3 Godfrey Gumbs,4 Detlev Grützmacher,3 and Vijay Narayan1,†

1Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom
2Air Force Research Laboratory, Space Vehicles Directorate, Kirtland Air Force Base, New Mexico 87117, USA
3Peter Grünberg Institute (PGI-9), Forschungszentrum Jülich, 52425 Jülich, Germany
4Department of Physics and Astronomy, Hunter College of the City University of New York, 695 Park Avenue, New York, New York 10065, USA
(Dated: August 19, 2017)

By combining n-type Bi$_2$Te$_3$ and p-type Sb$_2$Te$_3$ topological insulators, vertically stacked p-n junctions can be formed, allowing to position the Fermi level into the bulk band gap and also tune between n- and p-type surface carriers. Here we use low-temperature magnetotransport measurements to probe the surface and bulk transport modes in a range of vertical Bi$_2$Te$_3$/Sb$_2$Te$_3$ heterostructures with varying relative thicknesses of the top and bottom layers. With increasing thickness of the Sb$_2$Te$_3$ layer we observe a change from n- to p-type behavior via a specific thickness where the Hall signal is immeasurable. Assuming that the the bulk and surface states contribute in parallel, we can calculate and reproduce the dependence of the Hall and longitudinal components of resistivity on the film thickness. This highlights the role played by the bulk conduction channels which, importantly, cannot be probed using surface sensitive spectroscopic techniques. Our calculations are then buttressed by a semi-classical Boltzmann transport theory which rigorously shows the vanishing of the Hall signal. Our results provide crucial experimental and theoretical insights into the relative roles of the surface and bulk in the vertical topological p-n junctions.

PACS numbers: 73.20.-r, 73.25.+i, 73.50.-h

I. INTRODUCTION

Topological insulators (TIs) are bulk insulators with exotic ‘topological surface states’ (TSS) which are robust to backscattering from non-magnetic impurities, exhibit spin-momentum locking and have a Dirac-like dispersion. These unique characteristics present several opportunities for applications in spintronics, thermoelectricity, and quantum computation. However, a major drawback of ‘early generation’ TIs such as Bi$_{1-x}$Sb$_x$ and Bi$_2$Se$_2$Te that is the Fermi level $E_F$ intersects the conduction/valence bands, thus giving rise to finite conductivity in the bulk. This non-topological conduction channel conducts in parallel to the TSS and in turn subverts the overall topological nature. Thus, in order to create bona fide TIs, the Fermi level $E_F$ needs to be tuned within the bulk bandgap, and this has been previously achieved by means of electrical gating, doping, or, as recently reported, by creating p-n junctions from two different TI films.

In Ref. 14 a ‘vertical topological p-n junction’ was realized by growing an n-type Bi$_2$Te$_3$ layer capped by a layer of p-type Sb$_2$Te$_3$, and it was shown that varying the relative layer thicknesses serves to tune $E_F$ without the use of an external field. Importantly, such bilayer systems are expected to be significantly less disordered than doped materials such as (Bi$_{1-x}$Sb$_x$)$_2$Te$_3$ in which inhomogeneity of the dopants is a constant problem. Furthermore, and in sharp contrast to doped TIs, the intrinsic p and n character of the individual layers presents remarkable opportunities towards the observation of novel physics including Klein tunneling, spin interference effects at the p-n interface, and topological exciton condensates. However, currently there exists little understanding of the bulk conduction in such topological p-n junctions, primarily because ARPES used in Ref. 14 is a surface-sensitive method. This is especially noteworthy in light of the fact that the band structure varies along the depth of the TI p-n junction slab, in sharp contrast to the essentially constant band gap within the bulk of (Bi$_{1-x}$Sb$_x$)$_2$Te$_3$-type compounds. Understanding and minimizing the bulk conduction channels in TI p-n junctions is crucial in order to realize their technological potential as well as to gain access to the exotic physics they can host.

II. EXPERIMENT

Bi$_2$Te$_3$/Sb$_2$Te$_3$-bilayers (BST) were grown on phosphorous doped Si substrates using molecular beam epitaxy (MBE). Details of the MBE sample preparation can be found in Ref. 14. In all the samples, the bottom Bi$_2$Te$_3$-layer had thickness $t_{BiTe} = 6$ nm while the top Sb$_2$Te$_3$ layers had thicknesses $t_{SbTe} = 6.6$ nm (BST6), 7.5 nm (BST7), 15 nm (BST15), and 25 nm (BST25), respectively. The layers were patterned into Hall bars of width $W = 200$ µm and length $L = 1000$ µm using photolithography as a mask for ion milling, and Ti/Au contact pads were deposited for electrical contact. Low-T electrical
measurements were carried out using lock-in techniques in a He-3 cryostat with a base temperature of 280 mK and a 10 T superconducting magnet. Both longitudinal ($R_{xx}$) and transverse ($R_{xy}$) components of resistance were measured.

III. RESULTS

Figure 1(a) shows the longitudinal magnetoresistance (MR) $\equiv (R_{xx}(B) - R_{xx}(0))/R_{xx}(0)$ of the various samples considered. We find that above $\sim 2$ T the MR in BST6 and BST7 is manifestly linear whereas the MR in BST15 and BST25 appears to be neither purely linear nor quadratic. While there is experimental evidence suggesting an association between linear MR and linearly dispersive media$^{20–22}$, as well as a theoretical basis for this association$^{23}$, we note that disorder can also render giant linear MR$^{24,25}$ by admixing longitudinal and Hall voltages. In Fig. 1(b) we see that $R_{xy}$ is linear in $B$ and its slope changes sign from positive (BST6) to negative (BST15 and BST25). This is simply a reflection of different charge carrier types of Bi$_2$Te$_3$ (n-type) and Sb$_2$Te$_3$ (p-type), where electrons (holes) dominate transport when Sb$_2$Te$_3$ is thin (thick). Intriguingly, Fig. 1(c) shows $R_{xy}$ vs $B$ measured in two different Hall bar devices of BST7 to be strongly non-linear and non-monotonic. Qualitatively, it appears as if $R_{xy}$ is picking up a large component of $R_{xx}$ despite the Hall probes being aligned to each other with lithographic ($\mu$m-scale) precision.

![Figure 1](image1.png)

**Figure 1.** (a) MR and (b+c) $R_{xy}$ as a function of $B$ for different $t_{SBTe}$. All curves are measured at 280 mK. The high field MR is linear for thin samples and changes to parabolic for thicker samples. Cusp-like deviations at low fields are due to WAL corrections. The sign change of the slope in (b) indicates a net excess of $p$-type carriers. The investigation of this discrepancy is the major focus of this manuscript.

![Figure 2](image2.png)

**Figure 2.** (a+b) Weak antilocalization peaks for 2 different Sb$_2$Te$_3$-thicknesses and at 3 different temperatures. Fits to the measurements, based on the HLN model, are shown in straight red lines, while curves with a $\alpha = 0.5$ (green dashed line) and 1 (blue dashed-dotted line) allow to estimate the error. (c) $l_\phi$ as a function of $T$ for various $t_{SBTe}$ in a log-log plot. All curves are proportional to $\propto T^{-0.5}$ (dashed line) but shifted with respect to each other. (d) $\alpha$ as a function of $T$ for various $t_{SBTe}$.
electron scattering processes. The second fitting parameter $\alpha$ is depicted in Fig. 2(d) and we find values consistent with $\alpha = 0.5$ (error estimates on $\alpha$ can be found in Fig. 2(a) and a discussion in Appendix A). This is consistent with several previous reports on TI thin films.$^{9,29–31}$

IV. DISCUSSION

A. 3-channel model

Having ascertained that the transport characteristics of the Bi$_2$Te$_3$/Sb$_2$Te$_3$ heterostructures are consistent with conventional TI behaviour, we now proceed to understand the Hall characteristics. It is well-known that the TIs Bi$_2$Te$_3$ and Sb$_2$Te$_3$ show bulk conduction in addition to the TSS. Thus, we start with a simple picture of three independent conduction channels: bulk $n$- and $p$-type layers corresponding to the Bi$_2$Te$_3$ and Sb$_2$Te$_3$ layers, respectively, and a TSS on the top surface. While in principle a TSS exists also at the interface with the substrate, it is expected that its contribution to the conductivity is largely diminished due to the strongly disordered TI-substrate interface$^{31,32}$. Thus as a first approximation, we do not consider the bottom TSS.

Our starting point is the expressions for $\sigma_{xx}$ and $R_H$ in a multi-channel system$^{33–35}$

$$\sigma_{xx} = e \left( n_p \mu_p - n_n \mu_n \pm e n_t \mu_t \right)$$

$$R_H(t_{SbTe}) \equiv \frac{1}{\epsilon \cdot n_{eff}} = \frac{n_p \mu_p^2 - n_n \mu_n^2 \pm n_t (t_{SbTe}) \mu_t^2}{\epsilon (n_p \mu_p + n_n \mu_n + n_t (t_{SbTe}) \mu_t)^2} \quad (3)$$

Here $n_{eff}$ is the effective carrier concentration, $e$ is the charge of an electron and $-e$ is the charge of a hole, the subscript $p$, $n$, and $t$ signify bulk electrons, bulk holes, and surface carriers, respectively, $n_t$ are carrier concentrations, and $\mu_t$ represent the mobility of the charge carriers. The $\pm$ indicates, respectively, negative ($t_{SbTe} < 20 \text{ nm}$) and positive charge carriers ($t_{SbTe} > 20 \text{ nm}$) in the TSS.

The following literature values for the bulk layers are assumed: $n_{BiTe} = 8 \times 10^{19} \text{ cm}^{-3}$ and $\mu_n = 50 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$ for Bi$_2$Te$_3$, and $n_{SbTe} = 4.5 \times 10^{19} \text{ cm}^{-3}$ and $\mu_p = 300 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$ for Sb$_2$Te$_3$. In order to compare $n_{BiTe}$ and $n_{SbTe}$ to the TSS carrier concentration, we convert them to effective areal densities as $n_t \equiv n_{BiTe} \cdot t_{BiTe}$ and $n_p \equiv n_{SbTe} \cdot t_{SbTe}$. It can be shown that $n_t \propto E_B^2$ where $E_B$ is the difference between $E_F$ and Dirac point (see Eq. B3, Appendix B) and $E_B$, in turn, can be retrieved from ARPES measurements in Ref.$^{14}$. $\mu_t$ is used as a fitting parameter.

Figure 3(a) shows $R_H$ as predicted by the model using the above parameters to be in good agreement with the measured values. However, for the same parameters we find that $R_{xx} \equiv (L/W)\sigma_{xx}$ is significantly underestimated especially for low $t_{SbTe}$ (Fig. 3(b)). A likely source of this discrepancy is that the bulk $\mu_t$ values are not applicable for the ultra-thin films. This is especially so considering the fact that a depletion zone will form at the $p$-$n$ interface. Determining the exact profile of the charge carrier density at the interface is beyond the scope of this paper and instead, we demonstrate that an ad-hoc thickness-dependent reduction of $\mu_t$ of the bulk layers with all other parameters unchanged, can significantly improve the quality of the predictions. Figure 3(d) shows the result of a fit in which $\mu_p$ and $\mu_n$ are reduced to 20% of their bulk value in BST6 and BST7, and to 95% of their bulk value in BST15 and BST25. Not only do we obtain excellent agreement with the $R_H$ data, the model is also able to accurately predict $R_{xx}$ (Fig. 3(e)). The obtained value of $\mu_t = 281 \pm 17 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$ is well within the range of previous studies in ultra-thin TIs where the TSS dominate transport$^{11}$.

Figure 3(f) shows the important physical insight we arrive at on the basis of this simple model: the bulk contribution is drastically reduced in thin films (see Fig. 3(c)), with the TSS eventually dominating the overall conductivity $\sigma_{tot}$ (see Fig. 3(f)).

To test this conclusion we measure samples with top-gate electrodes which enable the tuning of the Fermi level $E_F$ via a gate voltage $V_G$. A variation of $E_F$ should lead to perceptible changes of the transport properties.
of the TSS (see Fig. 4(b)) while transport through the bulk should be less affected due to screening. As can be seen in Fig. 4(a) this is indeed the case, with the resistance of the thin, TSS dominated sample much more dependent on $V_G$ than the thick, bulk dominated sample. The resistance of the thin sample is maximized when $V_G = -12 \, V$, likely corresponding to the alignment of $E_F$ with the Dirac point. Thus, broadly speaking, despite where the Fermi velocity of a Dirac cone, $v_F$ is the Fermi energy of a Dirac cone, and $\rho_s(u_s) = u_s/(2\pi e^2 v_F^2)$ is the surface density-of-states of a Dirac cone.

The bulk mobility tensors $\tilde{\mu}_{c,v}(B,z)$ are given by

$$
\tilde{\mu}_{c,v}(B,z) = \frac{\mu_0(z)}{1 + \frac{\mu_0^2(z)}{B^2}} \begin{bmatrix}
1 & -\mu_0(z)B \\
-\mu_0(z)B & 1
\end{bmatrix} ,
$$

(6)

where $\mu_0(z) = e_\gamma c_h \tau_{p(e,h)}(z) / m_{c,h}^*$. A derivation of the bulk mobility tensor can be found in Appendix D. The bulk conductivity tensor is then calculated as

$$
\tilde{\sigma}_{c,v}(B) =
$$

(7)

$$
\tilde{\sigma}_{c,v}(B) =
$$

(7)

Likewise, the surface mobility tensor is

$$
\tilde{\mu}_{s}^\pm(B) = \pm \frac{\mu_1}{1 + \frac{\mu_1^2}{B^2}} \begin{bmatrix}
1 & \mp \mu_1 B \\
\mp \mu_1 B & 1
\end{bmatrix} ,
$$

(8)

where $\mu_1 = 4e_\epsilon \tau_{p(e,h)}^2 v_F^2 / \sigma_i e^3$, $\epsilon_s$ is the host dielectric constant, and $\sigma_i$ is the surface density of impurities. This corresponds to a surface conductivity tensor given by

$$
\tilde{\sigma}_s^\pm(B) =
$$

(9)

Similarly, one obtains the surface current per length as

$$
\mathbf{j}_s^\pm = \frac{e\tau_{p(e,h)} k_F^2}{\tau_{sp} v_F} \mathbf{v}_s^\pm(u_s) \left\{ \left[ \tilde{\mu}_{s}^\pm(B) \cdot \mathbf{E} \right] \cdot \mathbf{v}_s^\pm(u_s) \rho_b(u_s) \right\} ,
$$

(5)

where the $\pm$ denote when the Fermi level lies above and below the Dirac point, respectively, $\tau_s$ and $\tau_{sp}$ are surface energy- and momentum relaxation times, $k_F^2 = \sqrt{4\pi n_s}$, where $n_s$ is the areal density of surface electrons, $v_F$ the Fermi velocity of a Dirac cone, $\mathbf{v}_s^\pm(k_s) = \pm (k_s / k_F^2) \mathbf{v}_F$, $u_s = \hbar v_F k_F^2$ is the Fermi energy of a Dirac cone, and $\rho_b(u_s) = u_s/(2\pi e^2 v_F^2)$ is the surface density-of-states of a

Clearly consistent with the observation of ‘no’ Hall slope in BST7.

### B. Semi-classical theory

Although our simplistic model offers useful physical insights, for a more microscopic understanding it is desirable that one is not dependent on ad-hoc assumptions and/or a large number of experimental parameters. In the following we present a semi-classical theory for calculating magneto-conductivity tensors of surface and bulk charge carriers in a topological $p$-$n$ junction using zeroth and first-order Boltzmann moment equations\textsuperscript{37}. Assuming the $p$-$n$ interface to be in the $x$-$y$ plane, then under a parallel external electric field $\mathbf{E} = (E_x, E_y, 0)$ and a perpendicular magnetic field $\mathbf{B} = (0, 0, B)$, the total current per length in a $p$-$n$ junction structure is given by

$$
\int_{-L_D}^{L_D} dz \left[ \mathbf{j}_s^0(\mathbf{B}) + \mathbf{j}_s^1(\mathbf{B}) \right] + \mathbf{j}_b^0,\text{ where } L_D \text{ and } L_A \text{ are the thickness of the } p \text{ region (donors) and } n \text{ region (acceptors), respectively. Here } \mathbf{j}_s^0 \text{ indicate the current densities with } i = c, v \text{ or } s \text{ for conduction valance band, wave surface and, respectively. The superscript } || \text{ is included to emphasize that the current considered is parallel to the } p$-$n \text{ interface as is experimentally the case. The bulk current densities are given by}

$$
\mathbf{j}_b^0 = \frac{2e\gamma c_h m_s^* \tau_{e,h}(z)}{\tau_{p(e,h)}(z)} \mathbf{v}_s^0(u_s(z)) \left\{ \left[ \tilde{\mu}_{0}^0(B) \cdot \mathbf{E} \right] \cdot \mathbf{v}_s^0(u_s(z)) \right\} D_{c,v}(u_s(z)) ,
$$

(4)
This provides a very useful microscopic ground-
to that in Eq. 3, but arrived at in a more rigorous
the obtained result is qualitatively unchanged. Impor-
tant, the physical content of Eq. 12 is essentially iden-
tical to that in Eq. 3, but arrived at in a more rigorous
fashion. This provides a very useful microscopic ground-

\[ \mathbf{\sigma} = e \sigma_s \left( \frac{\tau_s}{\tau_s \tau_{sp}} \right) \mathbf{\mu}_s (B) . \] (9)

\[ \mathbf{\sigma}_{\text{tot}} (B) = e \mathbf{\mu}_s (B) A_{\text{h}} \left[ (L_A - W_p) + \int_0^{W_p} dz \exp \left( -\frac{\beta e \mu h, N_A}{2 e_0 \epsilon_l D_h} z^2 \right) \right] - e \mathbf{\mu}_s (B) N_A A_{\text{e}} \]

\[ \times \left[ (L_D - W_n) + \int_0^{W_n} dz \exp \left( -\frac{\beta e \mu e, N_D}{2 e_0 \epsilon_l D_e} z^2 \right) \right] + e \mathbf{\mu}_s (B) \left( -\frac{\alpha_0^2}{4\pi h^2 v_F^2} \right) (L_A - L)^2 A_s , \] (10)

where \( \alpha_0 \) and \( L_0 \) are constants to be determined experi-
mentally, \( N_{\text{D,A}} \) are doping concentrations, \( W_n \) and \( W_p \)
are the thicknesses of the depletion zones for donors and
acceptors in a \( p-n \) junction, \( \mu_{e,h} \) are \( \mu_0 (z) \) evaluated at
\( n_{e,h} (z) = N_{\text{D,A}} \), \( D_{e,h} \) are diffusion coefficients, \( \beta = 4/3 \)
(\( \beta = 7/3 \)) for longitudinal (Hall) conductivity. In addi-
tion, the averaged mobilities \( \bar{\mu}_e, v_B (B) \) are defined by their
values of \( \tau_{p(e,h)} (z) \) at \( n_{e,h} (z) = N_{\text{D,A}} \), and three coeffi-
cients are \( A_s = \tau_s / \tau_{sp} \approx 3/4 , \)

\[ A_{e,h} = \frac{\tau_{e,h} (z)}{\tau_{p(e,h)} (z)} \bigg|_{n_{e,h} (z) = N_{\text{D,A}}} \]

\[ = \frac{1}{6} \left( \frac{Q_c}{k_{F,e}^h} \right)^2 \left[ \frac{2 \ln \left( 2k_{F,e}^h \right) - 1}{Q_c} \right] \]

\[ = \frac{Q_c^2}{6 (3 \pi^2 N_{\text{D,A}})^{2/3} } \left\{ 2 \ln \left[ \frac{2 (3 \pi^2 N_{\text{D,A}})^{1/3} }{Q_c} \right] - 1 \right\} , \] (11)

where \( 1/Q_c \) is the Thomas-Fermi screening length. More
details on the derivation of the conductivity tensors can
be found in Appendix E.

From Eq. 10 one can see that there exists a critical
value of \( L_A = L^* \) at which the total Hall conductivity
becomes zero, which is determined from the following
quadratic equation

\[ \frac{\mu_{h}^2 N_A A_{h}}{1 + \mu_{h}^2 B^2} \left\{ (L^* - W_p) + \int_0^{W_p} dz \exp \left( -\frac{7 e \mu h, N_A}{6 e_0 \epsilon_l D_h} z^2 \right) \right\} - \frac{\mu_{e}^2 N_D A_e}{1 + \mu_{e}^2 B^2} \left\{ (L_D - W_n) + \int_0^{W_n} dz \exp \left( -\frac{7 e \mu e, N_D}{6 e_0 \epsilon_l D_e} z^2 \right) \right\} + \frac{\mu_{e}^2}{1 + \mu_{e}^2 B^2} \left( \frac{\alpha_0^2}{4\pi h^2 v_F^2} \right) (L^* - L_0)^2 A_s = 0 , \] (12)

where the sign \(+/(-)\) corresponds to \( L_A > L_0 \) \( (L_A < L_0) \)
for the contribution of the lower (upper) Dirac cone.

We note that in arriving at the above equations we
have not considered scattering between the TSS and bulk
layers. Including these will modify energy-relaxation
times for both bulk and surface states, although no ana-
lytical expression for these can be obtained even at low
\( T \). We leave a numerical evaluation of the problem for
a later manuscript. For the purposes of this manuscript,
we stress that the inclusion of this coupling only serves
to modify the three coefficients \( A_e, A_h \), and \( A_s \), and thus
the obtained result is qualitatively unchanged. Impor-
tantly, the physical content of Eq. 12 is essentially iden-
tical to that in Eq. 3, but arrived at in a more rigorous
fashion. This provides a very useful microscopic ground-

\[ \mathbf{\sigma}_{\text{tot}} (B) = e \mathbf{\mu}_s (B) + e \mathbf{\mu}_c (B) + e \mathbf{\mu}_s (B) \] is obtained as

\[ \mathbf{\sigma}_{\text{tot}} (B) = e \mathbf{\mu}_s (B) A_{\text{h}} \left[ (L_A - W_p) + \int_0^{W_p} dz \exp \left( -\frac{\beta e \mu h, N_A}{2 e_0 \epsilon_l D_h} z^2 \right) \right] - e \mathbf{\mu}_s (B) N_A A_{\text{e}} \]

\[ \times \left[ (L_D - W_n) + \int_0^{W_n} dz \exp \left( -\frac{\beta e \mu e, N_D}{2 e_0 \epsilon_l D_e} z^2 \right) \right] + e \mathbf{\mu}_s (B) \left( -\frac{\alpha_0^2}{4\pi h^2 v_F^2} \right) (L_A - L)^2 A_s , \] (10)

V. CONCLUSION

In conclusion, we have reported low-\( T \) magnetotrans-
port measurements on vertical topological \( p-n \) junctions
and understood the data within a three-channel model
for the Hall resistance. It provides useful insights into
the complex interplay of the bulk and TSS in the multi-
layered TI, explains the sign change of \( R_H \) with varying
\( I_{\text{SbTe}} \) and delivers values for the mobility of the TSS of
281 cm\(^2\)V\(^{-1}\)s\(^{-1}\). We then develop a Boltzmann trans-
port theory which provides a clear microscopic foundation for our model. Our work paves the way for the study of other complex TI heterostructures, where bulk states and TSS of different carrier types coexist. In future, our method can be applied to improved topological p-n junctions in which a top and bottom TSS can form novel Dirac fermion excitonic states.

**ACKNOWLEDGMENTS**

D.B., D.R. and V.N. acknowledge funding from the Leverhulme Trust, UK, D.B., R.M., D.R., and V.N. acknowledge funding from EPSRC (UK). DH would like to thank the support from the Air Force Office of Scientific Research (AFOSR). G.M., M.L., J.K. and D.G. acknowledge financial support from the DFG-funded priority programme SPP1666.

**Appendix A: Error estimates for α**

Figure 2(a) compares the results when 1) α and lφ were both fitting variables (red line) or 2) when lφ alone was used as a fitting variable and α was kept constant. We find that the fit for α = 1 (blue dashed-dotted line) is of a significantly poorer quality, indicating clearly that the data is consistent with the existence of one WAL mode. This errors become significantly larger as T is increased (here not shown) and thus one must not over interpret the apparent increase in α with T in Fig. 2(d).

**Appendix B: TSS electron density**

The density of states in the dirac cone is given by

\[ g(k)dk/2\pi^2 L = 2\pi k dk/2\pi^2 = kdk/(2\pi/L)^2 \]  
(B1)

The relation between the binding energy \( E_B \), i.e. the difference between the Fermi energy and the Dirac point, and the Fermi wave vector \( k_F \) is

\[ E_B = \beta k_F = hv_F k_F \]  
(B2)

and can be retrieved from ARPES measurements in Ref. 14, carried out using samples from the same growth process and identical material parameters. For \( E_B = 215\text{meV} \), \( k_F \approx 0.1\text{Å} \) (see Fig. 4(h) in Ref. 14), thus \( \beta = E_B/k_F = 3.44 \cdot 10^{-29} \text{J m} \). From \( \beta \), a Fermi velocity of \( 3.26 \cdot 10^5 \text{m/s} \) can be derived.

The electron density of the TSS is

\[ n_t = k_F^2/4\pi = E_B^2/4\pi \beta^2 \]  
(B3)

Furthermore, the relation between \( E_B \) and the Sb2Te3-thickness is linear (\( dE_B/dt_{SbTe} = 1.62 \cdot 10^{-12} \text{J/m} \), see Fig. 5) and

\[ n_t = (dE_B/dt_{SbTe} \cdot t_{SbTe})^2/4\pi \beta^2 \]  
(B4)

**Appendix C: Derivation of \( R_H \) and \( n_{\text{eff}} \)**

The force acting on charges in the TSS (index t), bulk-Sb2Te3 (p) and bulk-Bi2Te3 (n) originate from an electric field \( \vec{E} \) in y-direction and a magnetic field \( \vec{B} \) in z-direction:

\[ -F_{ny} = eE_y + ev_nx_Bz \]
\[ -F_{ty} = eE_y + ev_k B_z \]
\[ F_{py} = eE_y - ev_{px} B_z \]  
(C1)

Using \( v = \frac{\mu}{e} F \) with \( \mu \) the mobility, we obtain

\[ \frac{v_{ny}}{\mu_n} = E_y + v_n x B_z \]
\[ \frac{v_{ty}}{\mu_t} = E_y + v_t x B_z \]  
(C2)

\[ \frac{v_{py}}{\mu_p} = E_y - v_p x B_z \]

Furthermoe, no charge current is flowing in y-direction

\[ J_y = J_n + J_k + J_p \]
\[ = e(n_n v_{ny} + e n_k v_{ty} + e n_p v_{py}) = 0 \]  
(C3)

Inserting the velocities in the previous equation gives

\[ n_n \mu_n (E_y + \mu_n x B_z) \]
\[ = -(n_k \mu_k (E_y + \mu_k x B_z) + n_p \mu_p (E_y - \mu_p x B_z)) \]
\[ \Rightarrow E_y (n_n \mu_n + n_k \mu_k + n_p \mu_p) \]
\[ = B_z E_x (-n_n \mu_n^2 + n_k \mu_k^2 + n_p \mu_p^2) \]  
(C4)

The charge current in x-direction is

\[ J_x = e n_n v_{nx} + e n_k v_{kx} + e n_p v_{px} \]
\[ = (n_n \mu_n + n_k \mu_k + n_p \mu_p) e E_x \]  
(C5)

\( E_x \) can now be replaced, resulting in

\[ e E_x (n_n \mu_n + n_k \mu_k + n_p \mu_p)^2 \]
\[ = B_z J_x (-n_n \mu_n^2 + n_k \mu_k^2 + n_p \mu_p^2) \]
\[ \Rightarrow R_H = \frac{B_z J_x}{E_x} = \frac{-n_n \mu_n^2 + n_k \mu_k^2 + n_p \mu_p^2}{e (n_n \mu_n + n_k \mu_k + n_p \mu_p)^2} \]  
(C6)
Both $n_p$ and $n_t$ are depending on the thickness of the Sb$_2$Te$_3$-thickness, $t_{SbTe}$, with

$$n_p = n_{SbTe} \cdot t_{SbTe}$$

$$n_t(t_{SbTe}) = \frac{(dE_B/dSbTe \cdot (t_{SbTe} - t_0))^2}{4\pi\beta^2}$$

Thus $R_H(t_{SbTe})$ is a function of the Sb$_2$Te$_3$-thickness of the form

$$R_H(t_{SbTe}) = \frac{-n_n(t_{SbTe})\mu_n^2 + n_t(t_{SbTe})\mu_t^2 + n_p\mu_p^2}{e(n_n(t_{SbTe})\mu_n + n_t(t_{SbTe})\mu_t + n_p\mu_p)^2}$$

where $dE_B/dSbTe$ can be gained from Fig. 5.

Because of the entity $R_H = -1/(e \cdot n_{eff})$, the 'effective' 2-dimensional charge density is given by

$$n_{eff} = \frac{(n_n(t_{SbTe})\mu_n + n_t(t_{SbTe})\mu_t + n_p\mu_p)^2}{-n_n(t_{SbTe})\mu_n^2 + n_t(t_{SbTe})\mu_t^2 + n_p\mu_p^2}$$

as well as the source vector $s$, given by

$$s = \left[ \begin{array}{c} q_1(r_{11}E_1 + r_{12}E_2 + r_{13}E_3) \\ q_2(r_{21}E_1 + r_{22}E_2 + r_{23}E_3) \\ q_3(r_{31}E_1 + r_{32}E_2 + r_{33}E_3) \end{array} \right] ,$$

we can reduce the linear equations to a matrix equation

$$\hat{\mathbf{C}} \cdot \mathbf{v}_d = s$$

with a formal solution $\mathbf{v}_d = \hat{\mathbf{C}}^{-1} \cdot s$. Explicitly, we find the solution $\mathbf{v}_d = \{v_1, v_2, v_3\}$ for $j = 1, 2, 3$ as

$$\frac{\partial \mathbf{v}_d(t|z)}{\partial t} = -e\hat{\mathbf{M}}^{-1}_e(z) \cdot \mathbf{v}_d(t|z)$$

$$- \mathbf{E}(t) \times \mathbf{v}_d(t|z) = 0 ,$$

as well as the source vector $s$, given by

$$s = \left[ \begin{array}{c} q_1(r_{12}B_3 - r_{13}B_2) \\ q_1(r_{11}B_1 - r_{12}B_3) \\ q_1(r_{11}B_1 - r_{11}B_3) \end{array} \right] ,$$

and $q = -e$. By defining the coefficient matrix $\hat{\mathbf{C}}$ for the above linear equations, i.e.,

$$\hat{\mathbf{C}} = \left[ \begin{array}{ccc} 1 + q_1(r_{12}B_3 - r_{13}B_2) & q_1(r_{11}B_1 - r_{11}B_3) & q_1(r_{11}B_1 - r_{11}B_3) \\ q_2(r_{22}B_3 - r_{23}B_2) & 1 + q_2(r_{23}B_1 - r_{23}B_3) & q_2(r_{22}B_3 - r_{23}B_2) \\ q_3(r_{32}B_3 - r_{33}B_2) & q_3(r_{33}B_1 - r_{33}B_3) & 1 + q_3(r_{33}B_1 - r_{33}B_3) \end{array} \right] ,$$

we get the following group of linear inhomogeneous equations for $\mathbf{v}_d = \{v_1, v_2, v_3\}$
\(
\hat{\Delta}_1 = \begin{bmatrix}
q_1(r_{11}E_1 + r_{12}E_2 + r_{13}E_3) & q_1(r_{13}B_1 - r_{11}B_3) & q_1(r_{11}B_2 - r_{12}B_1) \\
q_2(r_{21}E_1 + r_{22}E_2 + r_{23}E_3) & q_2(r_{23}B_1 - r_{21}B_3) & q_2(r_{21}B_2 - r_{22}B_1) \\
q_3(r_{31}E_1 + r_{32}E_2 + r_{33}E_3) & q_3(r_{33}B_1 - r_{31}B_3) & q_3(r_{31}B_2 - r_{32}B_1)
\end{bmatrix},
\)

\(
\hat{\Delta}_2 = \begin{bmatrix}
1 + q_1(r_{13}B_1 - r_{11}B_3) & q_1(r_{11}E_1 + r_{12}E_2 + r_{13}E_3) & q_1(r_{11}B_2 - r_{12}B_1) \\
q_2(r_{23}B_1 - r_{21}B_3) & q_2(r_{21}E_1 + r_{22}E_2 + r_{23}E_3) & q_2(r_{21}B_2 - r_{22}B_1) \\
q_3(r_{33}B_1 - r_{31}B_3) & q_3(r_{31}E_1 + r_{32}E_2 + r_{33}E_3) & q_3(r_{31}B_2 - r_{32}B_1)
\end{bmatrix},
\)

\(
\hat{\Delta}_3 = \begin{bmatrix}
1 + q_1(r_{13}B_1 - r_{11}B_3) & q_1(r_{11}E_1 + r_{12}E_2 + r_{13}E_3) & q_1(r_{11}B_2 - r_{12}B_1) \\
q_2(r_{23}B_1 - r_{21}B_3) & q_2(r_{21}E_1 + r_{22}E_2 + r_{23}E_3) & q_2(r_{21}B_2 - r_{22}B_1) \\
q_3(r_{33}B_1 - r_{31}B_3) & q_3(r_{31}E_1 + r_{32}E_2 + r_{33}E_3) & q_3(r_{31}B_2 - r_{32}B_1)
\end{bmatrix}.
\)

By assuming \( r_{ij} = 0 \) for \( i \neq j \), \( r_{jj} = 1/m_j^* \) and introducing the notation \( \mu_j = q_3/m_j^* \), we find

\[
\hat{\mathcal{C}} = \begin{bmatrix}
1 & -\mu_1 B_3 & \mu_1 B_2 \\
\mu_2 B_3 & 1 & -\mu_2 B_1 \\
-\mu_3 B_2 & \mu_3 B_1 & 1
\end{bmatrix},
\]

\[
\hat{\Delta}_1 = \begin{bmatrix}
1 & \mu_1 E_1 & \mu_1 \mu_2 B_1 \\
\mu_2 B_3 & 1 & -\mu_2 B_1 \\
-\mu_3 B_2 & \mu_3 B_1 & 1
\end{bmatrix},
\]

\[
\hat{\Delta}_2 = \begin{bmatrix}
1 & \mu_1 E_1 & \mu_1 B_2 \\
\mu_2 B_3 & 1 & -\mu_2 B_1 \\
-\mu_3 B_2 & \mu_3 B_1 & 1
\end{bmatrix},
\]

\[
\hat{\Delta}_3 = \begin{bmatrix}
1 & -\mu_1 B_3 & \mu_1 E_1 \\
\mu_2 B_3 & 1 & \mu_2 E_2 \\
-\mu_3 B_2 & \mu_3 B_1 & \mu_3 E_3
\end{bmatrix},
\]

and

\[
\hat{\mu}_c(B) = \frac{\mu_0}{1 + \mu_0 B^2} \begin{bmatrix}
1 + \mu_0 B^2 & -\mu_0 B_3 & \mu_0 B_2 + \mu_0^2 B_1 B_3 \\
\mu_0 B_3 & 1 + \mu_0^2 B_2 B_1 & -\mu_0 B_2 + \mu_0^2 B_1 B_3 \\
-\mu_0 B_2 & \mu_0^2 B_3 B_1 & 1 + \mu_0^2 B_2^2
\end{bmatrix},
\]

where \( B^2 = B_1^2 + B_2^2 + B_3^2 \). By taking \( B = \{0, 0, B\} \), we find from Eq. (D10) that

\[
\hat{\mu}_s(B) = \frac{\mu_1}{1 + \mu_1^2 B^2} \begin{bmatrix}
1 & \mu_1 B \\
-\mu_1 B & 1
\end{bmatrix},
\]

where \( \mu_1 = e \tau_{se} v_F / (\hbar k_T^2) \), \( k_T^2 = \sqrt{4 \pi \sigma_s} \), and \( \sigma_s \) is the areal density of surface electrons.

**Appendix E: Bulk and surface conductivity tensors**

Under a parallel external electric field \( E = (E_x, E_y, 0) \) and a perpendicular magnetic field \( B = (0, 0, B) \), the total parallel current per length in a p-n junction structure is given by \( \int_{-L_d}^{L_d} dz [j_{\parallel}^b(z) + j_{\parallel}^s(z)] + j_{\parallel}^s \), where \( L_d \) and \( L_s \) are the distribution ranges for donors and acceptors, respectively. Here, by using the second-order Boltzmann moment equation\(^{42}\), the bulk current densities are found to be
\[
\mathbf{j}^\parallel_{c,v}(z) = \frac{2e\gamma_{c,h} m_{c,h}^* \tau_{c,h}(z)}{\tau_{p(c,h)}(z)} \mathbf{v}^\parallel_{c,v}(u_{c,v}(z)) \left\{ \left[ \frac{\mu^\parallel_{c,v}(\mathbf{B},z)}{\mu^\parallel_{c,v}(u_{c,v}(z))} \right] \cdot \mathbf{v}^\parallel_{c,v}(u_{c,v}(z)) \right\} D_{c,v}[u_{c,v}(z)] , \tag{E1}
\]

where \( D_{c,v}[u_{c,v}(z)] = \left( \sqrt{u_{c,v}(z)/4\pi^2} \right) (2m_{c,h}^*/h^2)^{3/2} \) is the electron and hole density-of-states per spin, \( u_{c,v}(z) = (h\kappa_{c,h}^*/2m_{c,h}^*) \) are Fermi energies and wave vectors in a bulk, \( m_{c,h}^* \) are effective masses of electrons and holes, \( \tau_{c,h}(z) \) and \( \tau_{p(c,h)}(z) \) are energy- and momentum relaxation times, respectively. \( \kappa_{c,h} \) is the Thomas-Fermi screening length.

In addition, the bulk energy-relaxation times \( \tau_{c,h}(z) = e\gamma_{c,h} \int_{-L_A}^{L_D} dz n_{c,h}(z) \left[ \frac{\tau_{c,h}(z)}{\tau_{p(c,h)}(z)} \right] \mu^\parallel_{c,v}(\mathbf{B},z) \right\} D_{c,v}[u_{c,v}(z)] , \tag{E3}
\]
and \( \mathbf{v}^\pm_{s}(k) = \pm (k_\parallel/k_\perp) v_F \).

From Eq. (E1), we find the bulk conductivity tensor as

\[
\tilde{\sigma}^\parallel_{c,v}(\mathbf{B}) = e\gamma_{c,h} \int_{-L_A}^{L_D} dz n_{c,h}(z) \left[ \frac{\tau_{c,h}(z)}{\tau_{p(c,h)}(z)} \right] \mu^\parallel_{c,v}(\mathbf{B},z) . \tag{E4}
\]

On the other hand, from Eq. (E2) we get the surface conductivity tensor, given by

\[
\tilde{\sigma}^\pm_{s}(\mathbf{B}) = e\gamma_s \left( \frac{\tau_s}{\tau_{sp}} \right) \mu^\pm_{s}(\mathbf{B}) . \tag{E5}
\]

Therefore, the total conductivity tensor \( \tilde{\sigma}_{tot}(\mathbf{B}) = \tilde{\sigma}^\parallel_{c,v}(\mathbf{B}) + \tilde{\sigma}^\parallel_{e,h}(\mathbf{B}) + \tilde{\sigma}^\pm_{s}(\mathbf{B}) \) can be obtained from

\[
\tilde{\sigma}_{tot}(\mathbf{B}) = e\tilde{\mu}^\parallel_{c,v}(\mathbf{B}) N_A A_h \left[ (L_A - W_p) + \int_0^{W_p} dz \exp \left( -\frac{\beta e\tilde{\mu}_h N_A}{2e_0e_r D_h} z^2 \right) \right] \\
- e\tilde{\mu}^\parallel_{e,h}(\mathbf{B}) N_D A_e \left[ (L_D - W_n) + \int_0^{W_n} dz \exp \left( -\frac{\beta e\tilde{\mu}_h N_D}{2e_0e_r D_e} z^2 \right) \right] + e\tilde{\mu}^\pm_{s}(\mathbf{B}) \left( \frac{\alpha_0^2}{4\pi h^2 v_F^2} \right) (L_A - L_0)^2 A_s , \tag{E6}
\]

where \( \alpha_0 \) and \( L_0 \) are constants to be determined experimentally, \( N_{D,A} \) are doping concentrations, \( W_n \) and \( W_p \) are depletion ranges for donors and acceptors in a \( p-n \) junction, \( \mu_{c,h} \) are mobility coefficients at \( n_{c,h}(z) = N_{D,A} \), \( D_{c,h} \) are diffusion coefficients, and \( \beta = 4/3 \) (\( \tilde{\gamma} = 7/3 \)) for longitudinal (Hall) conductivity. In addition, the averaged mobilities \( \tilde{\mu}^\parallel_{c,v}(\mathbf{B}) \) are defined by their values of \( \tau_{p(c,h)}(z) \) at \( n_{c,h}(z) = N_{D,A} \), and three introduced coefficients are \( A_s = \tau_s/\tau_{sp} \approx 3/4 \),

\[
\frac{1}{\tau_{c,h}(z)} = \left[ \frac{2n_i}{n_{c,h}(z)\pi h Q_c^2} \right] \left( \frac{e^2}{\epsilon_0 \epsilon_r} \right) ^2 \times \int_0^{k_{F,c,h}(z)} dk \mathcal{D}_{c,v}(k) \left( \frac{4k^2}{4k^2 + Q_c^2} \right) = \left[ \frac{n_i m_{c,h}^*}{8n_{c,h}(z)\pi^3h^3Q_c^2} \right] \left( \frac{e^2}{\epsilon_0 \epsilon_r} \right) ^2 \times \left\{ \left( 2k_{F,c,h}(z) \right)^2 - Q_c^2 \ln \left( \frac{2k_{F,c,h}(z)^2}{Q_c^2} \right) \right\} , \tag{E7}
\]

and the surface energy-relaxation time \( \tau_s \) is found to be37,40,41

\[
\frac{1}{\tau_s} = \frac{2r_i}{\pi^2 \sigma_s h^2 v_F} \left( \frac{e^2}{\epsilon_0 \epsilon_r} \right) ^2 \times \int_0^\pi d\phi \int_0^{k_F} dk_{\parallel} \left( \frac{k_{\parallel}^2}{q_c + 2k_{\parallel} |\cos \phi|} \right) ^2 , \tag{E8}
\]

where \( n_i \) and \( \sigma_i \) are the impurity concentration and sur-
Finally, the bulk chemical potentials for electrons \([u_e(z)]\) and holes \([u_v(z)]\) are calculated as

\[
[u_{e,h}(z)]^{3/2} = 3\pi^2 \left( \frac{\hbar^2}{2m_{e,h}} \right)^{3/2} n_{e,h}(z), \tag{E9}
\]

and the carrier density functions are

\[
n_{e,h}(z) = N_{D,A} \times \exp \left\{ -\gamma_{e,h} \left( \frac{\mu_{e,h}}{D_{e,h}} \right) \left[ \Phi(z) + \gamma_{e,h}(E^b_{p,e}/e) \right] \right\}. \tag{E10}
\]

Here, the expression for the introduced potential function \(\Phi(z)\) is given by

\[
\Phi(z) = \begin{cases} 
-\frac{E_{p}^b}{e} & , \\
-\frac{E_{p}^b}{e} + (eN_A/2\epsilon_{f,e})(z + W_p)^2 & , \\
\frac{E_{p}^b}{e} - (eN_D/2\epsilon_{f,h})(W_n - z)^2 & , \\
\frac{E_{p}^b}{e} & ,
\end{cases} \quad z < -W_p \\
-\frac{E_{p}^b}{e} < z < 0 \\
0 < z < W_n \\
\Phi(0) = 0, \quad z > W_n 
\] \tag{E11}

and \(E_{p}^b\) is the Fermi energy of electrons (holes) at zero temperature and defined far away from the depletion region.

\[ * \text{db}639@cam.ac.uk \]
\[ † \text{vn}237@cam.ac.uk \]


