Anomalous Finite Size Effects on Surface States in the Topological Insulator Bi$_2$Se$_3$

Jacob Linder,$^{1}$ Takehito Yokoyama,$^{2}$ and Asle Sudbø$^{1}$

$^1$Department of Physics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway
$^2$Department of Applied Physics, University of Tokyo, Tokyo 113-8656, Japan

(Dated: Received October 26, 2009)

Abstract: We study how the surface states in the strong topological insulator Bi$_2$Se$_3$ are influenced by finite size effects, and compare our results with those recently obtained for 2D topological insulator HgTe. We demonstrate two important distinctions: (i) the gapping due to hybridization of the surface states features an oscillating exponential decay as a function of $L$ in Bi$_2$Se$_3$, while this decay is absent for HgTe. Our findings suggest that Bi$_2$Se$_3$ is suitable for nanoscale applications. In particular, topological insulators could find applications in quantum computing or spintronics. Also, we propose a way to experimentally detect both of the predicted effects.

PACS numbers: 73.43.-f, 72.25.Dc, 85.75.-d

I. INTRODUCTION

A new state of matter, known as the topologically insulating state, has recently been experimentally observed after the successful prediction of its existence in HgTe. This state is characterized by the topological protection of the conducting states that form at the edges (in 2D) or surfaces (in 3D) of such materials, whereas the bulk states remain insulating due to a charge excitation gap. This distinguishes topological insulators from conventional insulators that do not feature such edge/surface states, and introduces several interesting contrasts from conventional insulators that do not feature such edge/surface states, and introduces several interesting contrasts

It has recently been realized that Bi$_2$Se$_3$ is a three-dimensional topological insulator with a large charge excitation gap in the bulk. The surface states have an energy dispersion that is linear in momentum and thus form a Dirac cone at low energy, similarly to graphene. In stark contrast, however, Kramers theorem does not guarantee the survival of the edge states in graphene in the presence of perturbations since it holds an even number of Dirac cones inside the Fermi contour. In Bi$_2$Se$_3$, the number of Dirac points inside the Fermi arc is odd which activates the protection from Kramers degeneracy theorem. This has very recently been experimentally observed.

The prospect of utilizing the protected surface states in topological insulators such as Bi$_2$Se$_3$ in actual devices related to quantum computing or spintronics demands that finite size effects are taken seriously. This fact is underlined by the finding of Ref. which showed that the edge states of HgTe quantum wells become gapped due to a finite-size effect as the width $L$ is decreased. The gapping was shown to become experimentally measurable around $L \approx 200$ nm, suggesting that the material loses its exotic edge state properties in this region. Clearly, this places severe restrictions on potential use of HgTe quantum wells in applications on the nm-scale.

In this work, we will demonstrate how the situation changes dramatically when considering Bi$_2$Se$_3$. Within a combined analytical and numerical approach, we show how the surface states in Bi$_2$Se$_3$ display a remarkable robustness towards finite-size effects, becoming measurably gapped only when the width drops to a few nm. This means that samples of Bi$_2$Se$_3$ can be made several tens of times smaller than HgTe while still retaining their characteristic surface states giving rise to the quantum spin Hall effect. We explain this observation in terms of the large charge excitation gap $|M|$ in Bi$_2$Se$_3$ which gives rise to a very short localization length of the surface states. Moreover, we show how the finite-size induced gapping in the surface states displays a qualitatively different dependence on $L$ in Bi$_2$Se$_3$ compared to HgTe, namely an oscillating exponential decay. Both of these effects can be measured in a two-terminal geometry as sketched in Fig. 1.

II. THEORY

The effective low-energy Hamiltonian for Bi$_2$Se$_3$ centered around the $\Gamma$ point in the Brillouin Zone may be written as:

$$\hat{H} = \left( \begin{array}{cc} \varepsilon_{k\perp} + M_{k\perp} & A_1 k_x \tau_z \\ A_2 k_x \tau_z & \varepsilon_{k\perp} + M_{k\perp} \end{array} \right) \approx \left( \begin{array}{cc} A_2 k_x \tau_z & A_1 k_x \tau_z \\ A_1 k_x \tau_z & \varepsilon_{k\perp} + M_{k\perp} \end{array} \right),$$

(1)
the bulk bands and the surface states with a gap of \( \Gamma \). Here, we have defined the following quantities:

\[
\varepsilon_k = C + D_1 k_\perp^2 + D_2 k_\perp^4, \\
\mathcal{M}_k = M - B_3 k_\perp^2 - B_2 k_\perp^4,
\]

and \( k_\perp = k_\perp \pm \ell k_y \). Here, \( \tau_x \) and \( \tau_z \) are the Pauli matrices in standard notation, while the parameters \( \{ A, B, C, D, M \} \) describe the band-structure in Bi\(_2\)Se\(_3\), obtainable by first-principles calculations. Such a fitting procedure was undertaken in Ref\(^{[18]}\), with the result \( C = -6.8 \times 10^{-3} \text{ eV} \), \( M = 0.28 \text{ eV} \), and

\[
A_1 = 2.2 \text{ eV} \cdot \text{Å}, \quad A_2 = 4.1 \text{ eV} \cdot \text{Å}, \quad B_1 = 10 \text{ eV} \cdot \text{Å}^2, \\
B_2 = 56.6 \text{ eV} \cdot \text{Å}^2, \quad D_1 = 1.3 \text{ eV} \cdot \text{Å}^2, \quad D_2 = 19.6 \text{ eV} \cdot \text{Å}^2.
\]

It should be noted that there is an anisotropy along the \( \hat{z} \)-axis and that the full 3D structure of Bi\(_2\)Se\(_3\) has been taken into account. The basis \( \Psi \) we have used for the \( 4 \times 4 \) Hamiltonian is

\[
\Psi = (|P1^+, \uparrow\rangle, |P2^-, \uparrow\rangle, |P1^+, \downarrow\rangle, |P2^-, \downarrow\rangle)^T,
\]

where \( \text{tr} \) denotes the transpose operation. The above \( p \)-orbital states are the relevant ones near the Fermi level of Bi\(_2\)Se\(_3\) and may be classified according to their parity \( \pm \) since inversion symmetry is preserved on the lattice. The states \( |P1\rangle \) stem from the Bi atoms, whereas \( |P2\rangle \) stems from the Se atoms. The \( |P1^+, \sigma\rangle \) and \( |P2^-, \sigma\rangle \) states (\( \sigma =\uparrow,\downarrow \)) have opposite parity and the order of them near the Fermi level is interchanged when spin-orbit coupling is taken into account\(^{[19]}\). In this way, the spin-orbit coupling effect is responsible for driving the system into a topologically insulating phase.

It is interesting to observe that the Hamiltonian Eq. \((1)\) contains spin-mixing terms due to the off-diagonal entries \( A_2 k_\perp \tau_x \). This in contrast to the 2D topological insulator HgTe, which is diagonal in spin-space. As a result, one might expect new features in the spin-current of Bi\(_2\)Se\(_3\) carried by the topological surface states.

### III. RESULTS AND DISCUSSION

We will consider a finite width \( L \) in the \( z \)-direction with open boundary conditions at the edges, i.e.

\[
\Psi(x, y, z = \pm L/2) = 0.
\]

Since the translational symmetry is broken along the \( z \)-direction, we perform a Peierls substitution \( k_z \to -i \tilde{\partial}_z \) in the Schrödinger equation \( \hat{H}\Psi = \varepsilon \Psi \). Assuming a plane-wave solution

\[
\Psi \sim e^{i\Lambda z},
\]

where \( \Lambda \) determines whether the mode is evanescent or propagating, we solve the secular equation to find the allowed eigenvalues for \( \Lambda \). These read \( \Lambda = \pm \Lambda_\alpha \), where

\[
\Lambda_\alpha = [A_1^2 - 2D_1(C - \varepsilon + D_2 k_\perp^2) + 2B_1(B_2 k_\perp^2 - M)]^{1/2} + \frac{\alpha \sqrt{R}}{\sqrt{2(B_1^2 - D_1^2)}}, \quad \alpha = \pm 1,
\]

with the definition

\[
R = A_2^2 [A_1^2 - 4D_1(C - \varepsilon + D_2 k_\perp^2) - 4B_1(M - B_2 k_\perp^2)] - 4A_2^2 k_\perp^2(B_1^2 - D_1^2) + 4[B_1(C - \varepsilon + D_2 k_\perp^2)]^2 + D_1(M - B_2 k_\perp^2)^2.
\]
As demanded by consistency, Eq. (7) reduces to the result of in the limiting case of zero anisotropy and $C = 0$. The total wavefunction $\Psi$ is a superposition of the terms $e^{\pm \Lambda_\alpha z}$ with belonging normalization coefficients. The open boundary conditions at $z = \pm L/2$ allow us to write down an implicit equation for the energy eigenvalues of $\Psi$. The values of $\varepsilon$ solving this equation then correspond to the bulk states and, where possible, surface-bound states. We arrive at the following energy eigenvalue equation:

$$
\sum_\alpha t_\alpha = \Lambda_+^2 + \Lambda_-^2 - (\Lambda_+^2 - \Lambda_-^2)^2(B_1^2 - D_1^2)/A_1^2 \Lambda_+ \Lambda_-,
$$

where $t_\alpha = \tanh(\lambda_\alpha L/2)$. For arbitrary values of $k_\perp$, Eq. (9) cannot be solved analytically. Instead, we employ a numerical solution of Eq. (9) to find the allowed values of $\varepsilon$ for a given value of $k_\perp$. All other material parameters are specified in Eq. (3).

To highlight the dramatic difference between the surface states in $\text{Bi}_2\text{Se}_3$ and the edge states in HgTe, we show in Fig. 2 the energy dispersion as a function of the transverse momentum $k_\perp$. The plots for HgTe were obtained by utilizing the results in Ref. [20]. In the lower panel, it is seen how a gap $\Delta$ opens between the edge states at the $\Gamma$ point $k_\perp = 0$ as $L$ decreases. Let us emphasize here that the gap $\Delta$ between the edge/surface states is to be distinguished from the gap $|M|$ between the bulk energy bands. For large $L \simeq 1000$ nm, the edge states remain ungapped for HgTe. In sharp contrast, the edge states in $\text{Bi}_2\text{Se}_3$ remain completely ungapped in the entire regime $L \in [150, 1000]$ nm considered in Fig. 2. In fact, we find that a measurable gap $\Delta$ does not begin to open until widths of $L \leq 10$ nm are reached. This fact suggests that much smaller samples that retain their conducting surface states can be fabricated in the case of $\text{Bi}_2\text{Se}_3$ than in the case of HgTe, which is our first main result. Such an observation is crucial for the prospect of utilizing topological insulators such as $\text{Bi}_2\text{Se}_3$ and HgTe in applications linked to quantum computing or spintronics.

In order to understand the large quantitative difference between the necessary width $L$ that induces gapping between the surface/edge states in $\text{Bi}_2\text{Se}_3$ and HgTe, we study more carefully the eigenvalues $\Lambda_\alpha$. The physical interpretation of these quantities is that $\text{Re}\{\Lambda_\alpha\}$ corresponds to an inverse localization length (or, alternatively, penetration depth into the bulk) for the surface states. Therefore, the largest of the length scales $(\text{Re}\{\Lambda_+\})^{-1}$ and $(\text{Re}\{\Lambda_-\})^{-1}$ mainly determines the density profile for the surface states and their penetration into the bulk. We have verified numerically that

$$
(\text{Re}\{\Lambda_-\})^{-1} \geq (\text{Re}\{\Lambda_+\})^{-1}
$$

for all energies inside the bulk gap $\pm |M|$ for both $\text{Bi}_2\text{Se}_3$ and HgTe, so that $\alpha = -1$ will determine the penetration depth of the surface states into the bulk. The penetration depth of the surface states can be estimated by

$$
\xi = \hbar v_F / |M|,
$$

where the Fermi velocity is provided by $v_F = A/h$ for HgTe and $v_F = A_2/h$ for $\text{Bi}_2\text{Se}_3$. Since $A \simeq A_2$ while $|M|_{\text{Bi}_2\text{Se}_3} \gg |M|_{\text{HgTe}}$, the large difference in the distribution length of the surface/edge states stems from the sizable charge excitation gap $|M|$ in $\text{Bi}_2\text{Se}_3$. It is tempting to draw an analogy to the midgap Andreev-bound states in $d$-wave superconductors induced at the interface which extend a distance into the bulk proportional to the inverse of the superconducting gap.

Our second main result is related to the manner in which the gap between the surface states in $\text{Bi}_2\text{Se}_3$ depends on the width $L$, in effect $\Delta = \Delta(L)$. It is instructive to first recall that $\Delta(L)$ in HgTe was shown to exhibit a purely exponential decay in Ref. [20]. To investigate the situation in $\text{Bi}_2\text{Se}_3$, we have solved numerically for $\Delta(L)$ and plotted the result in Fig. 3. It is seen that a qualitatively different scenario from HgTe transpires: the decay with $L$ is highly non-monotonous, and in fact features a superimposed oscillatory pattern on the exponential decay. The experimental signature of such an oscillatory decay would be to measure the conductance at a fixed chemical potential for several samples with different widths $L$ and see how the conductance appears and reappears, as we shall describe below.

We now proceed to explain the origin of the oscillatory decay of the gap $\Delta(L)$ found in $\text{Bi}_2\text{Se}_3$, considering the $\Gamma$ point $k_\perp = 0$. The crucial observation in this context is that the eigenvalues $\Lambda_\alpha$ are not purely real in the bulk insulating regime $\varepsilon \in [C - |M|, C + |M|]$. This is in contrast to HgTe, where $\Lambda_\alpha$ are purely real in this regime. For $\text{Bi}_2\text{Se}_3$, we find...
FIG. 4: (Color online) Plot of the two-terminal conductance at $T = 20$ K and $T = 200$ K as a function of $\mu$ and $L$.

that

$$\Lambda_+ = \Lambda^- \text{ for } \varepsilon \in \{C - |M|, C + |M|\},$$  \hspace{1cm} (12)

as can be verified directly from Eqs. (3) and (7) since $R < 0$. As a consequence, whereas a gap dependence of the type $e^{-\Delta L}$ found for HgTe\textsuperscript{20} dictates an exponential decay, it will give a superimposed oscillatory pattern on top of the exponential decay in the case of Bi\textsubscript{2}Se\textsubscript{3}. The natural question is then: is it possible to identify the reason for why $R < 0$ in Bi\textsubscript{2}Se\textsubscript{3} whereas $R > 0$ in HgTe, leading to qualitatively different behavior of the surface states gap? Analyzing the expression for $R$ in Eq. (7) at $k_\perp = 0$, it seen that neither the 3D nature or the anisotropy of the former material can be the reason, since there is no mixing between indices with subscript ‘1’ and ‘2’. Therefore, $R < 0$ seems to occur as a direct result of the material parameters given in Eq. (3). In principle, since all the parameters in the effective Hamiltonian of HgTe depend on the thickness $d$ of the sample (which also determines whether the material is in the trivial or topologically insulating state), it might be possible to obtain a conversion from exponential decay to oscillating decay also in HgTe, although this clearly warrants a separate investigation based on first-principles calculations.

The simplest way to experimentally detect both the gapping of the surface states and their unusual dependence on $L$ is arguably a two-terminal geometry, as shown in Fig. 1. In that case, the conductance can be evaluated in the Landauer-Büttiker framework at a finite temperature $T$. Considering a zero-bias situation $eV \rightarrow 0$, the transmission coefficient $T$ may be written as

$$T(\varepsilon) = N_c[\Theta(\varepsilon - \Delta/2) + \Theta(-\varepsilon - \Delta/2)],$$  \hspace{1cm} (13)

where $\Theta$ is the Heaviside-step function, $N_c$ is the number of conducting channels on the surfaces $z = \pm L/2$, and $\varepsilon$ is the quasiparticle energy. Similarly to Ref.\textsuperscript{20}, one arrives at the following expression for the conductance normalized against $N_c$:

$$G = \frac{e^2}{h}\left\{1 + [1 + e^{\beta(\Delta/2 - \mu)}]^{-1} - [1 + e^{\beta(-\Delta/2 - \mu)}]^{-1}\right\},$$  \hspace{1cm} (14)

with $\beta^{-1} = k_B T$. As a direct consequence of the gap $\Delta$, the conductance is suppressed as $T \rightarrow 0$. In Fig. 4 we plot the conductance as a function of $L$ and the chemical potential $\mu$, comparing the temperatures $T = 20$ K and $T = 200$ K. As seen, the conductance displays oscillations for a fixed $\mu$, and the oscillation length depends on the value of the chemical potential. This is qualitatively completely different from HgTe, where the conductance displays a monotonic dependence on $L$. The oscillatory features become more smeared at elevated temperatures, as expected.

Finally, we note that after this paper was submitted for publication, we learned about the very recent work of H.-Z. Lu et al.\textsuperscript{22} and C.-X. Liu et al.\textsuperscript{23}, in which similar finite size effects have been predicted.

IV. SUMMARY

In summary, we have investigated finite size effects on the surface states in the strong topological insulator Bi\textsubscript{2}Se\textsubscript{3}, comparing the results also with those recently reported for HgTe\textsuperscript{20}. We demonstrate that the surface states respond differently to finite size effects in these materials, both quantitatively and
qualitatively. First of all, while the edge states become measurably gapped around $L \approx 200 \text{ nm}$ in HgTe\cite{2}, the surface-states in Bi$_2$Se$_3$ display a considerable robustness towards decreasing $L$ and become measurably gapped around $L \approx 10 \text{ nm}$. In this way, the topological surface state remains intact for a wider range of widths $L$. Secondly, the gapping between the surface states features a qualitatively distinct dependence on $L$ in Bi$_2$Se$_3$ compared to HgTe, namely an oscillatory decay with $L$ which stems from the material parameters that give an eigenvalue pair $\Lambda_\alpha$ that are complex conjugates. Both of these effects can be experimentally detected in a two-terminal geometry by varying the width and the chemical potential of the junction.

Acknowledgments

J.L. and A.S. were supported by the Research Council of Norway, Grants No. 158518/431 and No. 158547/431 (NANOMAT), and Grant No. 167498/V30 (STORFORSK). T.Y. acknowledges support by JSPS.

1 For a review, see e.g. M. König et al., J. Phys. Soc. Jap. 77, 031007 (2008) and references therein.
2 M. König et al., Science 318, 766 (2007).
3 D. Hsieh et al., Nature 452, 970 (2008); D. Hsieh et al., Science 323, 919 (2009).
4 B. A. Bernevig et al., Science 314, 1757 (2006).
5 C. L. Kane and E. J. Mele, Phys. Rev. Lett. 95, 146802 (2005); ibid 95, 226801 (2005).
6 S. Murakami, Phys. Rev. Lett. 97, 236805 (2006).
7 B. A. Bernevig and S. C. Zhang, Phys. Rev. Lett. 96, 106802 (2006).
8 L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. 98, 106803 (2007); L. Fu and C. L. Kane, Phys. Rev. B 76, 045302 (2007); J. E. Moore and L. Balents, Phys. Rev. B75, 121306(R) (2007).
9 X.-L. Qi, T. Hughes, and S.-C. Zhang, Phys. Rev. B 78, 195424 (2008).
10 X.-L. Qi et al., Science 323, 1184 (2009).
11 T. Yokoyama, Y. Tanaka, and N. Nagaosa, Phys. Rev. Lett. 102, 166801 (2009); T. Yokoyama, Y. Tanaka, N. Nagaosa, arXiv:0907.2810
12 Y. Tanaka, T. Yokoyama, N. Nagaosa, Phys. Rev. Lett. 103, 107002 (2009).
13 S. Mondal et al., arXiv:0908.2019
14 L. Fu and C. L. Kane, Phys. Rev. Lett. 100, 096407 (2008).
15 L. Fu and C. L. Kane, Phys. Rev. Lett. 102, 216403 (2009).
16 J. Nilsson, A. R. Akhmerov, and C.W. J. Beenakker, Phys. Rev. Lett. 101, 120403 (2008).
17 A. R. Akhmerov, J. Nilsson, and C.W.J. Beenakker, Phys. Rev. Lett. 102, 216404 (2009).
18 H. Zhang et al., Nature Phys. 5, 438 (2009)
19 Y. Xia et al., Nature Phys. 5, 398 (2009). 
20 B. Zhou et al., Phys. Rev. Lett 101, 246807 (2008).
21 C.-R. Hu, Phys. Rev. Lett. 72, 1526 (1994).
22 H.-Z. Lu et al., arXiv:0908.3120
23 C.-X. Liu et al., arXiv:0908.3654