

This is the accepted manuscript made available via CHORUS. The article has been published as:

Interplay of bulk and edge states in transport of two-dimensional topological insulators

R. W. Reinthaler and E. M. Hankiewicz

Phys. Rev. B **85**, 165450 — Published 30 April 2012

DOI: [10.1103/PhysRevB.85.165450](https://doi.org/10.1103/PhysRevB.85.165450)

Interplay of bulk and edge states in transport of two-dimensional topological insulators

R. W. Reinthaler and E. M. Hankiewicz

Faculty of Physics and Astrophysics, University of Würzburg, Würzburg, Germany

(Dated: April 9, 2012)

We study transport in two-terminal metal/quantum spin-Hall insulator (QSHI)/metal junctions. We show that the conductance signals originating from the bulk and the edge contributions are not additive. While for a long junction the transport is determined by the edge states contribution, for a short junction, the conductance signal is built from both bulk and edge states in the ratio which depends on the width of the sample. Further, in the topological insulator regime the conductance for short junctions shows a non-monotonic behavior as a function of the sample length. Surprisingly this non-monotonic behavior of conductance can be traced to the formation of an "effectively propagating" solution which is robust against scalar disorder. Our predictions should be experimentally verifiable in HgTe QWs and Bi₂Se₃ thin films.

I. INTRODUCTION

Topological states of matter are characterized by bulk invariants: Chern numbers^{1,2} or by a Z_2 invariant³, depending if the time reversal symmetry is broken or conserved. Due to the bulk-boundary correspondence^{4,5} this classification can be translated to the existence of topologically protected states at the edges of the system. In the case of the Z_2 insulator the edge states are formed by time reversed modes, so called Kramers' partners and they are helical⁶, i.e. Kramers' partners counter-propagate along a given edge of the sample. Depending if the number of Kramers' pairs on the edge is even or odd, the system is topologically trivial or non-trivial. Although the topological invariant is the property of the bulk, experiments usually detect edge states properties. Indeed the confirmation that HgTe/CdTe quantum wells (QWs) are two-dimensional topological insulators⁷ was provided through the measurement of quantized edge conductance⁸⁻¹¹.

One benefit of HgTe/CdTe QWs is that the topological order can be controlled by the thickness of the HgTe layer: below the critical thickness of $d_c \approx 6.3$ nm the system is a trivial insulator, whereas above d_c the system behaves as topological insulator. This gives the possibility to observe the topological phase transition directly. Therefore it is of great interest to find further experimental measurable indicators which distinguish the two regimes. Since the topological invariant is a bulk property, it seems natural that the bulk conductivity could also carry information about topological properties of the system. Indeed it was proposed in Ref. 12 that a non-monotonic conductance as a function of geometrical aspect ratio in metal/HgTe/metal junctions with a characteristic maximum describes the topological insulator regime. The conductance maximum occurs when incoming metallic bulk states tunnel through a short and wide two-dimensional topological insulator and is robust against scalar disorder on the order of the energy gap¹³. However, so far this analysis was limited to the periodic

boundary conditions (PBC) and neglected the existence of edge states, which can significantly modify the behavior in experimentally relevant setups. Here we analyze carefully the properties of the single QSHI/metal interface as well as the interplay between edge states and bulk states in metal/QSHI/metal junctions. To do so we apply a generalized wave matching method based on Fourier modes, like it was e.g. used to analyze two interface tunneling junctions in HgTe¹⁴. Zhang et al. restricted themselves to a low energetic behavior around the gap, which is dominated by the linear dispersion of the bulk and the edge states. In contrast our studies combine a topological insulator with two highly doped metallic leads treated as highly doped HgTe with quadratic dispersion. We show that the conductance signals originating from the bulk and the edge contributions are not additive. While for a long junction the transport is determined by the edge states contribution, for a short junction, the conductance signal is built from both bulk and edge states in the ratio which depends on the width of the sample. In particular, for the short and wide junctions, the longitudinal conductance shows a non-monotonic behavior with the characteristic maximum as a function of the sample length in the topological insulator regime. Surprisingly this maximum is built from two counter-propagating evanescent waves from the left and the right interfaces which cancel each other leading to the formation of an "effectively propagating" solution, robust against scalar disorder. Our predictions should be experimentally verifiable in HgTe QWs and Bi₂Se₃ thin films.

The rest of this paper is organized as follows: In the next Section we give a short introduction to the model. In Section III we analyze the single QSHI/metal interface, which helps us to understand the metal/QSHI/metal junction results in Section IV. We use tight-binding calculations in Section V to test the robustness of conductance maximum in the presence of scalar disorder. We finish the paper with conclusions.

II. BRIEF DESCRIPTION OF MODEL

Close to the Γ point ($k = 0$), HgTe quantum wells or Bi_2Se_3 thin films can be described by an effective 4×4 Dirac-like model^{7,15,16}. In the following we concentrate on the description of HgTe QWs while we will comment on Bi_2Se_3 thin films at the end of this section. In the absence of structure inversion asymmetry the Hamiltonian for HgTe QWs consists of two decoupled blocks, which are related by time reversal symmetry (TRS). Using the basis $(|E1+\rangle, |H1+\rangle, |E1-\rangle, |H1-\rangle)^T$, with (E1) and (H1) designating electron and heavy hole sub-bands and \pm denoting the Kramers' partner, the Hamiltonian reads ($k_{\pm} = k_x \pm ik_y$, $k = \sqrt{k_x^2 + k_y^2}$, $k_{\parallel} = (k_x, k_y)^T$)

$$\mathcal{H} = \begin{pmatrix} h(k_{\parallel}) & 0 \\ 0 & h^*(-k_{\parallel}) \end{pmatrix}, \quad h(k_{\parallel}) = \epsilon(k)\mathbb{I} + d_{\alpha}(k_{\parallel})\sigma_{\alpha}, \quad (1)$$

with $d_{\alpha}(k_{\parallel}) = (Ak_x, -Ak_y, \mathcal{M}(k))^T$. σ_{α} are the Pauli matrices and \mathbb{I} is the unit matrix in the subband space. Here $\epsilon(k) = C - Dk^2$ is the energy dispersion, $\mathcal{M}(k) = M - Bk^2$ is the k dependent gap and A gives the strength of the coupling between electron and heavy hole states. The parameters A , B , D and M depend on the quantum well width. Below (above) the critical thickness of the quantum well $M > 0$ ($M < 0$), the system describes the topologically trivial (non-trivial) regime. Due to the block diagonal form of Eq. (1) we will just consider the upper block in the following. The results for the lower block then follow by TRS. A generalization including a finite value of structure inversion asymmetry was presented in Ref. 15.

In the following we consider the transport along x -direction through a single interface between QSHI and normal metal as well as QSHI connected to two metallic leads (see inset to Fig. 1 and Fig. 3a, respectively). We set C to zero in the center region, so called quantum spin-Hall insulator regime, and to a large negative value in the leads, thereby effectively modeling the metal contacts by highly doped HgTe, cf. Fig. 3a for a schematic. So far similar systems were described within periodic boundary conditions (PBC)¹² where due to the periodicity in y -direction k_y is conserved and quantized i.e. $k_y = 2\pi n/W$, where $n = 0, \pm 1, \dots, \pm N_{\max}$ with N_{\max} given by the number of propagating modes. However PBC do not allow for a formation of edge states in the transport direction and the competition between edge and bulk transport was not explored so far. In order to include edge states one has to choose hard wall boundary conditions (HBC), i.e. the wave function is zero at the edges of the sample. Throughout this paper we use both boundary conditions to distinguish the transport carried by bulk and edge states. An analytical solution of Eq. (1) for HBC¹⁷ couples different k_y modes and a simple wave matching procedure like e.g. in Refs. 18 and 12 is no longer possible. To resolve this

issue we expand the states in terms of the Fourier modes $\phi_n(y) = \sqrt{2/W} \sin[n\pi y/W]$, whose orthogonality allow an independent matching for each mode. Therefore, for an infinite quasi-1D system with k_x conserved, the eigenvector can be expanded as follows:

$$\psi_m(x, y) = \exp[ik_x^m x] \sum_{n=1}^{\infty} \chi_n^m \phi_n(y) \quad (2)$$

Here the index m labels the k_x^m -eigenvalues, χ_n^m is the two component spinor corresponding to the upper block of Hamiltonian (1) for a given transverse mode n . To find the eigenvalues and eigenspinors we use the Schrödinger equation

$$(\mathcal{H} - E) \exp[ik_x^m x] \sum_{n=1}^{\infty} \chi_n^m \phi_n(y) = 0. \quad (3)$$

Multiplying from the left by $\int_0^W dy \phi_{n_2}^{\dagger}(y)$ and using the orthogonality of the sine functions yields:

$$\mathcal{H}^{\text{const}} \chi_{n_2}^m + \mathcal{H}^{k_x} k_x^m \chi_{n_2}^m + \mathcal{H}^{k_x^2} k_x^m \chi_{n_2}^m + \sum_{n_1} \int_0^W dy \phi_{n_2}^{\dagger} \mathcal{H}^{k_y} \phi_{n_1} \chi_{n_1}^m = 0, \quad (4)$$

where all terms of the Hamiltonian proportional to $k_y = -i\partial_y$ are collected in \mathcal{H}^{k_y} , all constant terms in $\mathcal{H}^{\text{const}}$ and all with k_x (k_x^2) in \mathcal{H}^{k_x} ($\mathcal{H}^{k_x^2}$). Additionally we introduced $\chi_{n_2}^m = k_x^m \chi_{n_2}^m$. Defining the vectors $\chi^m = (\chi_{n=1}^m, \chi_{n=2}^m, \dots)^T$ and $\chi'^m = (\chi_{n=1}^m, \chi_{n=2}^m, \dots)^T$, which are built by the 2 component spinors χ_n^m and $\chi_n'^m$, Eq. (4) can be written as matrix equation

$$\begin{pmatrix} \mathbb{I} & 0 \\ 0 & (H^{k_x^2})^{-1} \end{pmatrix} \begin{pmatrix} 0 & \mathbb{I} \\ H^{\text{const}} + H^{k_y} & H^{k_x} \end{pmatrix} \begin{pmatrix} \chi^m \\ \chi'^m \end{pmatrix} = k_x^m \begin{pmatrix} \chi^m \\ \chi'^m \end{pmatrix}. \quad (5)$$

So doubling the dimension of the system of equations by introducing χ'^m allows to reduce the problem of finding k_x^m to a linear eigenvalue equation^{19,20}. The sub-matrices in Eq. (5) are

$$H_{n_1 n_2}^{k_x^2} = \delta_{n_1 n_2} \text{diag} [\tilde{D}_+, \tilde{D}_-], \quad (6)$$

$$H_{n_1 n_2}^{k_x} = \delta_{n_1 n_2} \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, \quad (7)$$

$$H_{n_1 n_2}^{\text{const}} = \delta_{n_1 n_2} \text{diag} [\tilde{E}_+, \tilde{E}_-], \quad (8)$$

$$H_{n_1 n_2}^{k_y} = \begin{pmatrix} -(\frac{n\pi}{W})^2 \tilde{D}_+ \delta_{n_1 n_2} & iA\eta_{n_1 n_2} \\ -iA\eta_{n_1 n_2} & -(\frac{n\pi}{W})^2 \tilde{D}_- \delta_{n_1 n_2} \end{pmatrix}, \quad (9)$$

Further we used here $\tilde{D}_{\pm} = (D \pm B)$ and $\tilde{E}_{\pm} = C - E \pm M$. The only term coupling different modes is $\eta_{n_1 n_2} = \langle \phi_{n_1} | k_y | \phi_{n_2} \rangle$. The solutions k_x^m can be characterized as propagating

($k_x^m \in \mathbb{R}$) or evanescent ($\text{Im}(k_x) \neq 0$). For real k_x^m the propagation direction can be determined by the sign of

$$v^m = \int_0^W dy \psi_m^\dagger(x, y) [\partial_{k_x} h(k)]_{k_x \rightarrow k_x^m} \psi_m(x, y). \quad (10)$$

Analogously evanescent states with $\text{Im}(k_x) > 0$ ($\text{Im}(k_x) < 0$) are decaying to the right (left). In the following we will denote right (left) going solutions by the index m_R (m_L). In general the Fourier ansatz also works when the two blocks of Hamiltonian (1) are connected by the Rashba-spin orbit coupling like it was e.g. done in Ref. 14.

Having the solutions for a single system we can now consider QSHI/metal and metal/QSHI/metal junctions. The interfaces in Figs. 1 and 3a break translational symmetry in x -direction and an incoming state ψ_{in} can be scattered in all possible k_x -modes. We will now label the k_x eigenvalues as: $k_{x,1}^m$ for the part 1 stretching from $x = -\infty$ to $x = -L/2$, $k_{x,2}^m$ for $-L/2 < x < L/2$ and $k_{x,3}^m$ for $x \geq L/2$. If the incoming state is assumed as incident from the left in mode m , the state in the left lead takes the form

$$\Psi_1(x, y) = \psi_{m,1}(x, y) + \sum_{m_L} r_{m_L, m} \psi_{m_L,1}(x, y) \quad (11)$$

and that of the right lead is

$$\Psi_3(x, y) = \sum_{m_R} t_{m_R, m} \psi_{m_R,3}(x, y). \quad (12)$$

Here $|r_{m_L, m}|^2$ and $|t_{m_R, m}|^2$ are the probabilities that the incoming mode m is reflected into mode m_L of the left lead or transmitted into mode m_R of the right lead, respectively. Only in the central region of length L there are left and right outgoing states

$$\Psi_2(x, y) = \sum_{m_R} c_{m_R, m} \psi_{m_R,2}(x, y) + \sum_{m_L} d_{m_L, m} \psi_{m_L,2}(x, y). \quad (13)$$

The scattering amplitudes are computed by matching the wave functions $\Psi_i(x, y)$ and the associated currents $J_i^x(x, y) = [\partial_{k_x} h(k)]_{k_x \rightarrow -i\partial_x} \Psi_i(x, y)$ at the interfaces, i.e. $\Psi_1(-L/2, y) = \Psi_2(-L/2, y)$, $\Psi_2(L/2, y) = \Psi_3(L/2, y)$, $J_1^x(-L/2, y) = J_2^x(-L/2, y)$ and $J_2^x(L/2, y) = J_3^x(L/2, y)$. For setups like in Fig. 1, where we have only one interface ($L = 0$), the matching simplifies to $\Psi_1(0, y) = \Psi_3(0, y)$ and $J_1^x(0, y) = J_3^x(0, y)$. The total transmission then is

$$T = \sum_{\{m_R | k_{x,3}^{m_R} \in \mathbb{R}\}} \sum_{\{n_R | k_{x,1}^{n_R} \in \mathbb{R}\}} \frac{v_3^{m_R}}{v_1^{n_R}} |t_{m_R, n_R}|^2. \quad (14)$$

Moreover the knowledge of all scattering amplitudes determines the full state $\Psi(x, y)$ up to normalization. $\Psi(x, y)$ is the response to an incoming mode $\psi_{m,1}(x, y)$ from the left lead. Therefore it allows

us to calculate the non-equilibrium charge density $n(x, y) = |\Psi(x, y)|^2$ as well as the non-equilibrium current density, $\mathbf{J}(x, y) = \Psi^\dagger(x, y) \nabla_{\mathbf{k}} h(k) \Psi(x, y)$.

For actual computation the Fourier series has to be truncated at a sufficiently high mode N . The error of this approximation depends weakly on the width of the sample and on the Fermi energy. However, the asymmetric shape of an edge state is harder to approximate by sine functions than bulk states. Nevertheless a good approximation can be achieved. For example the maximal relative errors in the 99 lowest k_x^m eigenvalues at $W = 1000$ nm and $E_f = 0$ changes from 2 % at $N = 100$ to 0.6 % at $N = 200$. The higher modes decay very fast and have little influence on the transport. Unless otherwise specified we will use the following parameters: $A = 0.375$ nmeV, $B = -1.120$ nm²eV, $D = -0.730$ nm²eV, $M = -3.10$ meV. When we are referring to the normal regime we put $M = +3.10$ meV.

In Bi₂Se₃ thin films the overlap of surface states of opposite surfaces leads to hybridization. This finite size effect opens a gap at the Γ point of the Dirac spectrum of the surface states^{21,22}. This gap closes when the energetically lowest lying states, which have opposite parity, cross at certain thicknesses d_z^n ($n = 1, 2, \dots$) of the thin film¹⁶. At each crossing the parity changes leading to the oscillations between topologically trivial and non-trivial insulators as a function of the film thickness. Around the band crossing points a folding of the Hamiltonian of the 3D topological insulator^{23,24} to the lowest energetic sub-bands reproduces the model Eq. (1)^{16,25,26} for the 2D topological insulator. We choose for the calculations $d_z = 3$ nm with the parameters of 2D effective model as follows: $A = 0.406$ nmeV, $B = -0.250$ nm²eV, $D = 0.070$ nm²eV, $M = -22.5$ meV²⁵.

III. SINGLE INTERFACE BETWEEN QSHI AND METAL

Here we analyze the injection of the helical edge states from the quantum spin Hall insulator into a metallic lead. The latter is modeled by Eq. (1) with a high doping, i.e. large $|C|$. The used setup is sketched in the inset of Fig. 1. On the QSHI side, the Fermi energy is chosen to be zero while on the metal side E_f is shifted by C , with C characterizing an energetic height (strength) of the interface. Since we analyze properties of edge states, we choose HBC throughout this section. In Fig. 1 we show the transmission through such an interface as a function of C for different widths of the system $W = 250, 1000$ and 2000 nm. While for narrow samples strong interfaces introduce a significant backscattering due to an overlap of counter-propagating edge states, perfect transmission is observed for wide samples. The perfect transmission for wide samples can be explained analogously as for the topologically protected edge states in graphene²⁷, where

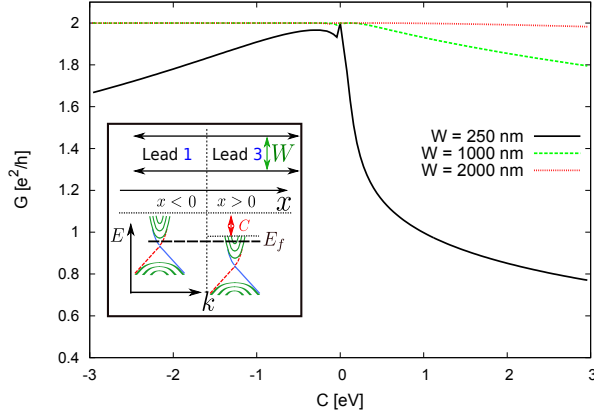


FIG. 1: The conductance as a function of the energetic height of the interface at a junction between a QSHI and a doped QSHI. A sketch of the system is shown in the inset. Perfect transmission is achieved in wide samples where counter propagating edge states for the same Kramers' partner (spin) do not couple. Negative (positive) C indicate a QSHI/n-type metal (QSHI/p-type metal) interface.

it was found that the counter propagating edge states for the same Kramers' partner are orthogonal. Indeed taking into account an exponential decay of the wave function for an edge state¹⁷ it can be easily confirmed that the overlap between edge states goes to zero for $W \rightarrow \infty$ in the presence of any x dependent potential:

$$\langle \Psi_{k_x} | V(x) | \Psi_{-k_x} \rangle = V(x) \int_0^W dy \Psi_{k_x}^\dagger \Psi_{-k_x} \xrightarrow{W \rightarrow \infty} 0. \quad (15)$$

By the Onsager- Casimir symmetry relation^{28,29} injecting electrons from the metal to QSHI should be also perfect for very wide samples. Let us mention for completeness, that a linear Rashba coupling¹⁵ does not increase the backscattering, because, although the direction of edge state polarization is rotated, there are still two orthogonal solutions for edge states at the same edge.

IV. METAL/QSHI/METAL JUNCTION

In this section we consider metal/QSHI/metal junctions with a finite length L of the QSHI as shown in Fig. 3a. So far such junctions have been only analyzed using PBC which neglect the topological edge states¹². However, although PBC excludes the existence of edge states in the transport direction, such junctions still allow for the distinction of different topological phases by studies of the longitudinal conductance of the evanescent bulk modes¹². While this longitudinal conductance rises monotonically with decreasing L in the topologically trivial regime ($M > 0$), the topologically non-trivial regime ($M < 0$) is characterized by a conductance maximum at

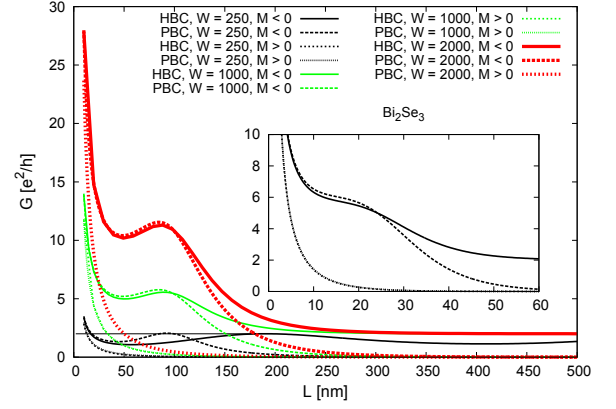


FIG. 2: The conductance through a metal/QSHI/metal junction, like depicted in Fig. 3a, as a function of the sample length L for different W (all values in nm). We compare the normal regime ($M = 3.10$ meV > 0) with the inverted one ($M = -3.10$ meV < 0) for HBC and PBC. Metallic leads are treated as highly doped HgTe with $C = -2.5$ eV. The inset shows the corresponding conductance signals for a Bi_2Se_3 thin film ($M = \pm 22.5$ meV) of width 250 nm at $C = -2.5$ eV.

L_{\max} . In the limit of $D = 0$ and $BC \gg A^2 \gg BM$ ($C < 0$) the position of the maximum can be predicted for the zero mode by

$$L_{\max} \approx \frac{A}{2|M|} \ln \left(\frac{2BC}{A^2} \right). \quad (16)$$

The purpose of this paper is to better understand the existence of the conductance maximum in the topologically non-trivial regime as well as to study the interplay of the bulk and edge contributions using hard wall boundary conditions. We use the technique described in Ref. 12 to solve the problem with PBC, where edge states are conceptually excluded, and the method described in Section II for a full solution with the hard wall boundary conditions. The comparison of results from PBC and HBC allows us to analyze the individual contributions of bulk and edge states to transport. In Fig. 2 we present the conductance as a function of the sample length L . First we observe that the conductance decreases exponentially with the increase of L in the normal regime, independent of the boundary conditions. Indeed in this regime we expect only evanescent solutions and therefore the same results for PBC and HBC. We now focus on the topological insulator regime i.e. inverted regime with $M < 0$. In this regime, for HBC, both evanescent and edge states contributions to the conductance are present and the signal depends on the length and the width of the sample.

In the very narrow QSHI regime, see black line in Fig. 2, we find Fabry-Perot like oscillations in the transmission of the edge states for HBC. At large L , the differences between two oscillation maxima is given by π/k_{edge} , where k_{edge} is real wave vector of the edge states. The Fabry-Perot oscillations originates from the backscatter-

ing of edge states at the interfaces and therefore are only present for narrow samples with the finite overlap between counter propagating edge states. In this case the existence of edge states has a dramatic effect on the behavior of conductance as a function of L . The quantum confinement in narrow wires shifts the energy levels of the states to higher values. This also increases the corresponding imaginary k_x values, which leads to a faster decay of incoming electrons in the QSHI tunneling junction. The combination of the Fabry Perot oscillations of the edge states and the bulk modes changes the position and the shape of the conductance maximum, which does not coincide with the maximum generated by either the evanescent bulk states or the Fabry Perot oscillations alone. Therefore Eq. (16) can not be used to find a maximum of conductance.

In contrast, for a wide and long QSHI part, the conductance is dominated by the edge states contribution and equals $2e^2/h$. For wide and medium length of QSHI, PBC and HBC give very different results while for short and wide QSHI the transport is dominated by tunneling evanescent modes³⁵. This is a very interesting result since naively one would expect that the edge state and bulk contributions to the conductance are additive for wide QSHI.

The same behavior can be found for Bi_2Se_3 thin films. In the inset of Fig. 2 we show the conductance for a 250 nm wide strip. Instead of a maximum we observe a plateau like behavior in the inverted regime. However, increasing further the C parameter restores the conductance maximum. Despite the narrow width the signal of HBC and PBC coincides around the plateau. This can be explained by the larger gap in Bi_2Se_3 , which decreases the overlap of the edge states¹⁷.

Since we showed in the last section that the transmission through a sufficiently wide metal-QSHI interface is perfectly quantized, the scattering from the interfaces cannot be the reason for the absence of the edge state contribution. Therefore to understand the non-additive behavior of bulk and edge contributions and the formation of the conductance maximum, we plot in Fig. 4 the evolution of the charge density for different sample lengths. In Fig. 4, we take into account only scattering of the first incoming HBC mode with $n=1$, like it was described in Sec. II after Eq. (14), and use again the parameters for HgTe QWs. Since the lowest modes give the largest contribution to the conductance maximum as shown in Fig. 3b, it is justified to analyze first incoming mode signal to grasp the important details of the conductance maximum. For plotting the charge density we use the following normalization:

$$\int_0^W dy \Psi^\dagger(x = -\frac{L}{2}, y) \Psi(x = \frac{L}{2}, y) = 1. \quad (17)$$

Let us first concentrate on the Figs. 4a, 4c and 4d, where $L \neq L_{\text{max}}$. Due to the large C parameter a strong reflection occurs at the left interface leading to standing waves in the left lead, see e.g. Fig. 4d. Further we observe that

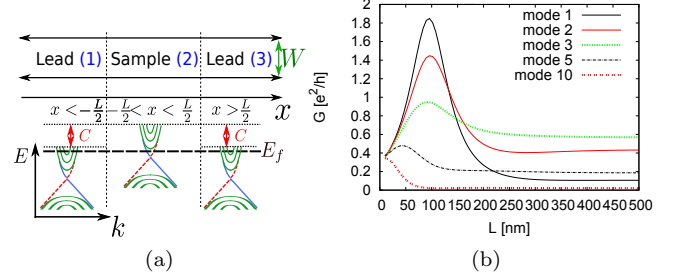


FIG. 3: (a) A schematic of the setup, consisting of the two metallic leads and the central QSHI at zero Fermi energy. The energy shift C is also shown in the band structures of the individual segments. (b) The conductance as a function of L for single modes ($W = 1000$ nm and $C = -2.5$ eV) in HBC. The maximum is mainly carried by the low modes.

the density is peaked inside the QSHI sample. The incoming states from the left lead are purely metallic bulk ones since for high energies (high C parameter) there is no influence of the edge states. The lowest metallic mode is characterized by a single maximum in the middle of the sample. In the short QSHI part (see Fig. 4a), the metallic states are not able to adjust to the shape of edge states, which are strongly confined to the edges of the wire, and so they keep their form and tunnel through the sample. Therefore for short L the conductance signals from HBC and PBC should coincide as indeed found in Fig. 2. When the length of QSHI part is comparable to the decay length of the evanescent modes, (see Fig. 4c), electrons are more likely transmitted through the sample by scattering from the bulk states into the edge states. For longer QSHI parts (see Fig. 4d) transport is driven by edge states.

It is interesting to compare Figs. 4a, 4c and 4d with Fig. 4b, where we choose the length of QSHI part corresponding to the peak in conductance i.e. $L \approx L_{\text{max}}$. One can see that the charge density is much larger in the QSHI part in comparison with other cases where $L \neq L_{\text{max}}$. Additionally we do not observe standing waves in the left lead, which indicates a weaker interface reflection. It also becomes clear that at the position of the maximum the edge states are not yet populated, which is the reason for a coincidence of conductance signals for PBC and HBC in Fig. 2 for L_{max} and large widths.

To further investigate the origin of the conductance maximum, when bulk states dominate transport, we analyze the effective wave vector in the direction of propagation for the QSHI part

$$k_x^{\text{eff}}(x) = \int_0^W dy \Psi_2^\dagger(x, y) (-i\partial_x) \Psi_2(x, y), \quad (18)$$

where $\Psi_2(x, y)$ is the full non-equilibrium scattering solution from Eq. (13). Our analysis will focus on the lowest

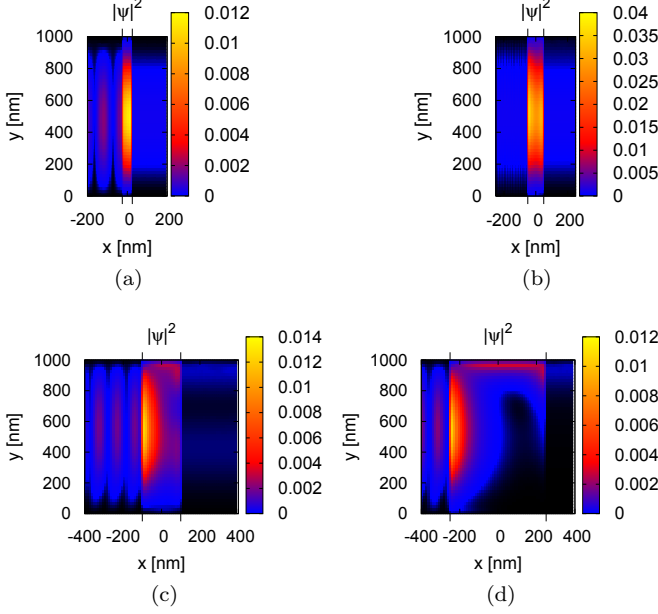


FIG. 4: Evolution of charge density in a metal/QSHI/metal junction as a function of the length of the sample. The charge response is the response to the first incoming mode, cf. Sec. II after Eq. (14). The interfaces at $\pm L/2$ are indicated by two black lines. In the plot, only upper edge state is seen since we restricted our calculations to the one spin direction. We used $W = 1000$ nm and $C = -2.5$ eV. (a) $L = 50$ nm, (b) $L = 100$ nm $\approx L_{\max}$, (c) $L = 200$ nm and (d) $L = 500$ nm. HBC have been used for calculation.

modes in PBC and HBC, which give the largest contribution to the signal. In Fig. 5 we present the imaginary as well as the real (inset) part of $k_x^{\text{eff}}(x)$ for different lengths and the lowest incoming mode. The lowest modes for PBC and HBC are $n=0$ and $n=1$, respectively. In general there are contributions of evanescent bulk and propagating edge modes, which have imaginary and real k_x , respectively. In this way a combination trivially leads to a complex k_x^{eff} . However, in the regime of short and wide samples the edge states are not populated, like we show for example in Fig. 4. In particular for $L = L_{\max} = 100$ nm and $W = 1000$ nm the edge states do not contribute (PBC and HBC give the same results for the lowest mode), and the real parts of k_x^{eff} in Fig. 5 arise from purely evanescent modes combined with the complex amplitudes in $\Psi_2(x, y)$. Indeed one can see in Fig. 5 that the real part of k_x^{eff} is non-zero and decays exponentially away from the two interfaces. The imaginary part instead has a more sophisticated behavior with several crossings of zero. In particular $\text{Im} k_x^{\text{eff}}(x)$ shows an antisymmetric behavior in respect to the sample middle for the length corresponding to the conductance maximum i.e. $L = 100$ nm $\approx L_{\max}$ for lowest modes. This suggests that the imaginary part of k_x^{eff} vanishes after

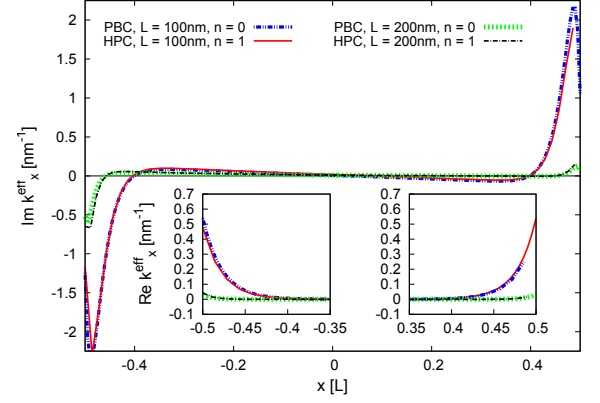


FIG. 5: The imaginary and real (inset) part of k_x^{eff} value as a function of the position x for different lengths of QSHI part, boundary conditions and incoming modes n . We used $W = 1000$ nm and $C = -2.5$ eV. The conductance maximum corresponds to the length $L_{\max} \approx 100$ nm.

averaging over x-direction. We define $\langle k_x^{\text{eff}} \rangle$ as:

$$\langle k_x^{\text{eff}} \rangle = \int_{-L/2}^{L/2} dx k_x^{\text{eff}}(x) \quad (19)$$

Fig. 6 shows $\langle k_x^{\text{eff}} \rangle$ as a function of the insulator length for PBC excluding any contribution of propagating edge states. The dashed red/dark gray curve in Fig. 6 shows that $\text{Im} \langle k_x^{\text{eff}} \rangle$ vanishes indeed at L_{\max} for the zero mode. The vanishing evanescent part of the effective wave vector coincides with a maximum of the real part of $\langle k_x^{\text{eff}} \rangle$ (solid red/dark gray line). This means that the transport is driven by two counter-propagating evanescent waves from the left and the right interfaces with the same amplitudes. In average the two evanescent contributions cancel each other. This leads to the resonant behavior reflected in the zero imaginary part of $\langle k_x^{\text{eff}} \rangle$ in Fig. 6, which we call "effectively propagating" solution. Of course this is not a propagating solution in the usual sense, since k_x^{eff} is still a complex number at each point inside the sample. Away from L_{\max} , $\text{Re} \langle k_x^{\text{eff}} \rangle$ decreases while $\text{Im} \langle k_x^{\text{eff}} \rangle$ has a finite value for the zero mode. Indeed Fig. 5 shows that the signal for $\text{Im} k_x^{\text{eff}}$ and $L = 2 * L_{\max}$ is not antisymmetric. The inset to Fig. 6 shows that the higher modes contribute less to the conductance maximum, but for $n = 1$ there is still a $\text{Re} \langle k_x^{\text{eff}} \rangle$ maximum around L_{\max} (solid green/light gray line in Fig. 6). This maximum for $n = 1$ is smaller in comparison to $n = 0$ and is shifted the same way as respective conductance maxima for $n = 0$ and $n = 1$ (see inset to Fig. 6). $\text{Im} \langle k_x^{\text{eff}} \rangle$ for $n = 1$ has also a zero value at L_{\max} . In contrary, for $n = 5$ which does not contribute to the maximum, the $\text{Im} \langle k_x^{\text{eff}} \rangle$ (black dashed line) is constant and the $\text{Re} \langle k_x^{\text{eff}} \rangle$ (black solid line) drops to zero. Therefore higher modes (modes with $n > 4$ for our set of parameters) behave almost like modes in

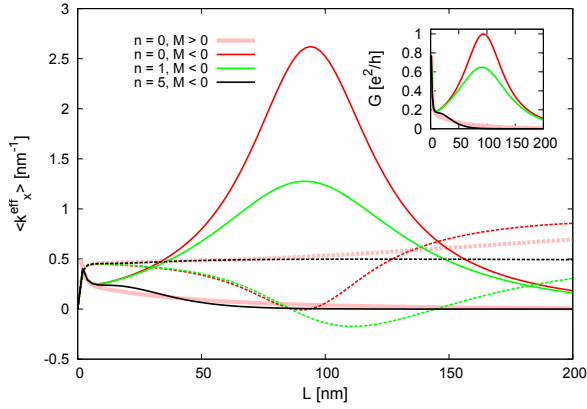


FIG. 6: $\langle k_x^{\text{eff}} \rangle$ as a function of the length for different incoming modes. Solid and dashed lines show real and imaginary parts of $\langle k_x^{\text{eff}} \rangle$, respectively. For clarity, only the calculations for PBC are shown. We see that for low modes the position of the conductance maximum (see inset) coincides with a maximum in the real part and a zero of the imaginary part of $\langle k_x^{\text{eff}} \rangle$. For higher modes this behavior gradually disappears. We used $W = 1000$ nm and $C = -2.5$ eV. The maximum lies at $L_{\text{max}} \approx 100$ nm.

the topologically-trivial regime (compare rose/light gray thick lines with black lines in Fig. 6) and have evanescent character.

Another way to find differences between the topologically trivial and non-trivial regimes is by analyzing the conductance as a function of the Fermi energy for a sample of fixed length. For PBC the two regimes show the same behavior as long as the Fermi energy in the sample lies within conduction or valence band¹². Only within the gap the QSHI can be distinguished from the normal insulator by a conductance peak around $E_f = 0$. For HBC we find the same behavior as within PBC in regimes where the bulk states dominate the transport, i.e. in wide and short QSHI junctions. When we increase the length of the QSHI above the decay length of the evanescent modes, the conductance is governed by the edge state contribution. In this limit within PBC the signal goes to zero while for HBC the conductance is $2e^2/h$ in wide wires where Fabry-Perot oscillations are negligible. In narrow and short samples however the transport is mediated by a mixture of edge and bulk states. Exactly this latter limit is shown in Fig. 7 where the Fermi energy dependence of the conductance is presented. In Fig. 7, the bulk gap is indicated by two black vertical lines. It is larger than $2|M|$ due to quantum confinement. Additionally the dashed vertical lines present the energy range, in which the overlap of the edge states produces a gap in the system¹⁷. We will refer to it as mini-gap from here on. One can notice that outside of the bulk gap (solid lines) the signals of the normal and the inverted regimes behave similar, like

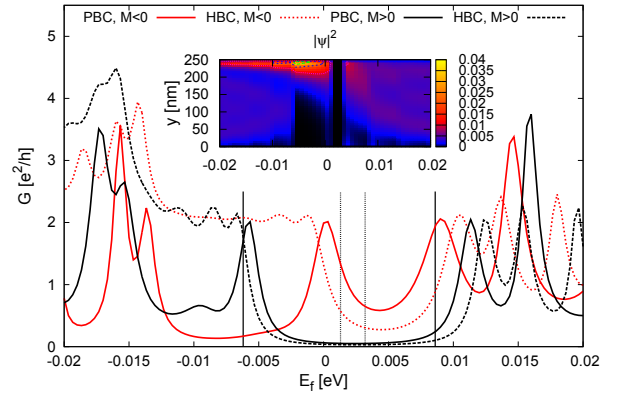


FIG. 7: The conductance as a function of the Fermi energy in the central part of the junction. The Fermi energy in the leads is $C = -2.5$ eV. The sample is 250 nm wide and 100 nm long. The vertical lines indicate the bulk gap (solid) and the mini-gap (dashed).

The inset shows the density of states for spin-up electrons as a function of the Fermi energy and the y -coordinate in an infinite wire and for the inverted regime. The mini-gap can be identified by the lack of any propagating solution.

it was found for wide structures^{12,13}. Within the gap however and for the PBC, a clear conductance peak in the inverted regime around $E_f = 0$ is seen, while in the normal regime the transport is almost completely suppressed. For HBC one observes a similar behavior for the normal regime, but the existence of the edge states changes the signal drastically for $M < 0$. Within the gap and for energies smaller than zero, we find the quantized conductance of the edge states. Around zero and for positive energies the signal is strongly reduced and resembles that of the bulk modes. This behavior can be understood by looking at the density of states in an infinite wire in the inverted regime. In the inset to Fig. 7 we show the corresponding $|\psi|^2$ as a function of the y -coordinate and the Fermi energy. As before we consider just one of the blocks, so that only one edge state (spin-up) is shown. The mini-gap as well as the bulk gap can be easily seen in this contour plot. We find that the edge state spreads over the whole sample shortly before the mini-gap opens. We have seen that the overlap of the edge state crucially influences the transport in metal/QSHI/metal junctions. Therefore let us note that the penetration depth of the edge state is in general a function of the energy^{17,30,31}, like it is shown in the inset of Fig. 7. For D finite and smaller than zero this behavior is asymmetric. For negative energies the edge state is strongly localized at the upper sample border. The small penetration depth leads to the fact that the edge state exists also at energies outside of the energy gap. For our parameters it merges to the valence band at energies of about $E_f = -80$ meV. Up to this energy the strong localization ensures a good quantization which we observe in the conductance. In contrast,

on the other side of the mini-gap ($E_F > 3$ meV) the edge state is widely spread, so that a strong backscattering occurs. Moreover it merges quickly to the bulk states around $E_f = 10$ meV. Therefore we believe that at these energies the transport is dominated by evanescent bulk states. Hence in this regime the conductance resembles the shape of the PBC signal.

V. DISORDER

In this section we want to analyze the robustness of the conductance signal in the presence of scalar disorder. Disorder breaks the translational invariance of the sample which makes the application of the wave matching routine impossible. We go around this problem by using tight-binding calculations within the Landauer-Büttiker formalism^{32–34} In the tight-binding approach we discretize the sample using a finite grid of say $N_x \times N_y$ points while the hard wall boundary conditions are naturally implemented by missing bonds at the borders. In principle we could have performed the whole analysis, which we reported in this paper, using tight-binding calculations. However, to check the validity of the conclusions for the experimentally relevant setups one needs to simulate large structures. Further for a good approximation within the tight-binding model one needs a small lattice constant a_L , i.e. a small distance between neighboring sites, because many fast decaying evanescent modes have to be taken into account. As a consequence we would have to use large systems with up to $(2000 \times 200 \times 2)^2$ (width \times length \times degrees of freedom) lattice points, which need extraordinary amount of memory and CPU time.

To analyze the influence of disorder let us here vary the sample structure and the parameters, similarly as it was performed for PBC in Ref. 13. First we decrease the width of the sample to 150 nm. For PBC this reduction should not change the position of the maximum. However, for HBC the overlap of the edge state wave functions influences the signal in narrow wires drastically. The overlap can be reduced by increasing the bulk gap, i.e. the M parameter as it was shown in Ref. 17. Further Eq. (16) shows that a larger M parameter leads to a smaller L_{\max} . Therefore, in the following by increasing M by the factor of 10 to $M = -31.0$ meV, we expect the maximum around 10 nm instead of 100 nm. Indeed Fig. 8 shows the conductance maximum at about 12 nm. First we compare the tight-binding (TB) calculations in the absence of disorder (red/gray dashed line in Fig. 8) with the signal computed in the wave matching (WM) formalism (red/gray solid line). The signals of the two models almost coincide when we choose $N_y = 250$ grid points. The number of points in x -direction varies with L , since we fix the lattice constant to $a_L = 150$ nm/251. Therefore indeed the

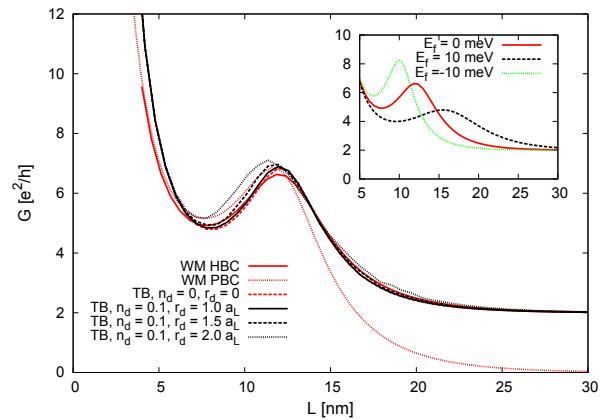


FIG. 8: The charge conductance as a function of the sample length. We compare the wave matching (WM) approach for HBC (solid red/gray line) and for PBC (dotted red/gray line) with the tight-binding calculations (dashed red/gray line) and find good coincidence. Adding scalar disorder with Gaussian correlations influences the signal slightly but does not destroy the conductance peak. We used $W = 150$ nm, $C = -2.5$ eV and $M = 31$ meV. The disorder strength was set to $W_d = 0.06$ eV for the black lines. The inset shows the influence of the Fermi energy E_f inside the sample region.

above described procedure does not change the results quantitatively but shifts the position of the conductance maximum to system sizes which can be easily calculated within tight binding model. Further the conductance for PBC (red/gray dotted line) coincide with the signal from tight-binding calculations indicating that the overlap between the edge states is small.

Scalar disorder is included by choosing randomly $N_d = n_d N_x N_y$ ($n_d \in [0, 1]$) lattice sites. For our analysis here we fixed $n_d = 0.1$. On each of the selected points we shift the energy randomly by $V_d \in [-W_d/2, W_d/2]$. The strongest disorder we can put without coupling the valence to the conduction band therefore is $W_d = 2|M|$. The disorder is modeled by a Gaussian-like function with the decay length r_d given in units of the lattice constant a_L . All signals including finite disorder have been averaged over 100 different disorder realization. In Fig. 8 we focus on the constant disorder strength $W_d = 0.06$ eV varying the decaying length of the Gaussian disorder from $r_d = 1a_L$ (solid black line) to $2a_L$ (dotted black line). For long sample lengths L , the disorder has no influence on the quantized conductance of the edge states, because the scalar disorder cannot couple the counter propagating edge channels. When evanescent waves contribute to the transport at smaller L the disorder has a visible influence on the signal. The influence of the disorder increases with an increase of disorder decay length. It influences the signal by shifting and increasing the maximum of conductance. To understand this effect we need to remember that

the transport for short QSHI lengths is dominated by evanescent modes. In this regime the extension of impurities can be comparable with the sample length introducing large fluctuations of energies over a relatively large part of the sample which can cause a shift of the conductance maximum. Indeed the inset to Fig. 8 shows the conductance signal shift to left or right depending on the Fermi energy inside the QSHI. Despite these small changes in the conductance behavior, the conductance signal is robust against disorder for HBC.

VI. CONCLUSION

We analyzed the interplay of edge and bulk states in metal/quantum spin-Hall insulator (QSHI)/metal junctions as a function of the size of QSHI part as well as the Fermi energy. We found that for short and wide QSHI part, the edge states are not populated due to the shape mismatch between the low modes existing in the highly doped metallic leads and the QSHI regime at zero Fermi energy. In this regime the transport occurs through the evanescent modes, and the topologically non-trivial insulator can be distinguished from the trivial insulator by the appearance of the distinct conductance maximum at small QSHI lengths or the conductance peak in the gap when the Fermi energy changes for the constant length. Moreover, we showed that the origin of this conductance maximum as a function of the length is caused by the

formation of an "effectively propagating" solution where two counter-propagating evanescent waves from the left and the right interfaces cancel each other.

In contrast, in the wide and long samples the transport is driven by spin edge channels while in the intermediate size regime the conductance signal is built from both bulk and edge states in the ratio which depends on the width of the sample. In this latter regime the combination of the Fabry-Perot oscillations (coming from the overlap of edge states) and bulk modes changes the position and the shape of the maximum. Further, our predictions are robust against scalar disorder. Therefore the measurements of systems with different geometrical aspect ratios should allow to distinguish the bulk and edge state contributions to the transport. Further the detection of the bulk conductance peak as a function of the Fermi energy or the length, for short and wide QSHI junctions, should be an alternative method to distinguish between topologically trivial and non-trivial insulators in HgTe QWs as well as Bi₂Se₃ thin films.

Acknowledgements

We acknowledge interesting discussions with B. Trauzettel, P. Recher, E. Novik, D.G. Rothe and P. Michetti. This work was funded through DFG Grant HA 5893/3-1. We thank Leibniz Rechenzentrum Munich for providing computing resources. EMH thanks Kavli Institute for Theoretical Physics for hospitality. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915.

-
- ¹ D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Phys. Rev. Lett.* **49**, 405 (1982).
 - ² M. Kohmoto, *Annals of Physics* **160**, 343 (1985).
 - ³ C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
 - ⁴ B. I. Halperin, *Phys. Rev. B* **25**, 2185 (1982).
 - ⁵ M. Z. Hasan and C. L. Kane, *Rev. Mod. Phys.* **82**, 3045 (2010).
 - ⁶ C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).
 - ⁷ B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, *Science* **314**, 1757 (2006).
 - ⁸ M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, *Science* **318**, 766 (2007).
 - ⁹ M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, *Journal of the Physical Society of Japan* **77**, 031007 (2008).
 - ¹⁰ A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, *Science* **325**, 294 (2009).
 - ¹¹ C. Brüne, A. Roth, H. Buhmann, E. M. Hankiewicz, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, *ArXiv:1107.0585* (2011).
 - ¹² E. G. Novik, P. Recher, E. M. Hankiewicz, and B. Trauzettel, *Phys. Rev. B* **81**, 241303 (2010).
 - ¹³ P. Recher, E. G. Novik, R. W. Reinthaler, D. G. Rothe, E. M. Hankiewicz, and B. Trauzettel, *Spintronics III* **7760**, 018 (2010).
 - ¹⁴ L. B. Zhang, F. Zhai, and K. Chang, *Phys. Rev. B* **81**, 235323 (2010).
 - ¹⁵ D. G. Rothe, R. W. Reinthaler, C.-X. Liu, L. W. Molenkamp, S.-C. Zhang, and E. M. Hankiewicz, *New Journal of Physics* **12**, 065012 (2010).
 - ¹⁶ C.-X. Liu, H. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, *Phys. Rev. B* **81**, 041307 (2010).
 - ¹⁷ B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu, *Phys. Rev. Lett.* **101**, 246807 (2008).
 - ¹⁸ T. Yokoyama, Y. Tanaka, and N. Nagaosa, *Phys. Rev. Lett.* **102**, 166801 (2009).
 - ¹⁹ Y.-C. Chang and J. N. Schulman, *Phys. Rev. B* **25**, 3975 (1982).
 - ²⁰ G. Liu, G. Zhou, and Y.-H. Chen, *Applied Physics Letters* **99**, 222111 (2011).
 - ²¹ J. Linder, T. Yokoyama, and A. Sudbø, *Phys. Rev. B* **80**, 205401 (2009).
 - ²² Y. Zhang, K. He, Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, et al., *Nature Physics* **6**, 584 (2010).
 - ²³ H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, *Nat Phys* **5**, 438 (2009).

- ²⁴ C.-X. Liu, X.-L. Qi, H. Zhang, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B **82**, 045122 (2010).
- ²⁵ H.-Z. Lu, W.-Y. Shan, W. Yao, Q. Niu, and S.-Q. Shen, Phys. Rev. B **81**, 115407 (2010).
- ²⁶ W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New Journal of Physics **12**, 043048 (2010).
- ²⁷ G. Metalidis and E. Prada, ArXiv:1012.4345 (2010).
- ²⁸ M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).
- ²⁹ M. Büttiker, IBM J. Res. Dev. **32**, 317 (1988).
- ³⁰ V. Krueckl and K. Richter, Phys. Rev. Lett. **107**, 086803 (2011).
- ³¹ F. Lu, Y. Zhou, J. An, and C.-D. Gong, ArXiv:1111.5550 (2011).
- ³² S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, 2007).
- ³³ E. M. Hankiewicz, L. W. Molenkamp, T. Jungwirth, and J. Sinova, Phys. Rev. B **70**, 241301 (2004).
- ³⁴ C. Brüne, A. Roth, E. G. Novik, M. König, H. Buhmann, E. M. Hankiewicz, W. Hanke, J. Sinova, and L. W. Molenkamp, Nat Phys **6**, 448 (2010).
- ³⁵ Please note, that the maximum in Fig. 2 cannot be predicted by Eq. (16) even for PBC, since we included a finite D parameter in our calculations.