

Controlling edge state transport in a HgTe topological insulator by superlattice effect

L.-Z. Lin,^{1,*} F. Cheng,² L. B. Zhang,³ D. Zhang,¹ and Wen Yang^{4,†}

¹*SKLSM, Institute of Semiconductors, Chinese Academy of Sciences, P.O. Box 912, Beijing 100083, China*

²*Department of Physics and Electronic Science, Changsha University of Science and Technology, Changsha 410004, China*

³*Department of Physics, Hunan Normal University, Hunan 410012, China*

⁴*Beijing Computational Science Research Center, Beijing 100089, China*

(Received 1 April 2013; published 24 June 2013)

We investigate theoretically the edge state transport in a HgTe topological insulator under periodic electrical modulation. We find constructive interference of the backscattering amplitudes, leading to the formation of superlattice minigaps and hence complete suppression of the edge state transmission. Consequently, the edge channel can be switched on/off by appropriately tuning the modulation amplitude via gate voltages, even for wide Hall bar with a small finite size effect. We also find efficient conversion between spin-up and spin-down edge channels by the gate-induced Rashba spin-orbit interaction.

DOI: [10.1103/PhysRevB.87.245311](https://doi.org/10.1103/PhysRevB.87.245311)

PACS number(s): 74.25.F-, 72.25.-b, 73.63.-b

I. INTRODUCTION

The topological insulator (TI) is a topologically distinct quantum state of the matter, where the spin-orbit coupling plays an essential role.^{1,2} Unlike the quantum Hall insulator (with broken time-reversal symmetry) or trivial insulators (with a zero \mathbb{Z}_2 topological invariant), the TI preserves the time-reversal symmetry and has nonzero \mathbb{Z}_2 topological invariants.³⁻⁷ At the interface between the TI and the trivial insulator, the transition from a nontrivial \mathbb{Z}_2 topology to a trivial topology dictates the closing of the insulating gap and hence the existence of an odd number of massless, helical Dirac fermions on the interface.⁸ For example, in a two-dimensional HgTe/CdTe or InAs/GaSb quantum well (QW) with an inverted band structure,^{9,10} the helical edge channels have been demonstrated experimentally in the Hall bar geometry.¹¹⁻¹³ These edge states show spin-momentum locking in a time-reversal invariant fashion: The right-moving electrons with spin pointing strictly up are connected to the left-moving electrons with spin pointing strictly down by time-reversal operation.¹⁴ Consequently, the spin-resolved edge states are robust against elastic scattering (e.g., by nonmagnetic impurities) that preserves the time-reversal symmetry.^{15,16} This offers great potential for a new generation of spintronic devices for low-power information processing.^{17,18} A crucial step in this development is the (preferentially *electrical*) manipulation, e.g., switch on or off, of the edge channel transport. However, since electrical modulation also preserves the time-reversal symmetry, it is not expected to have a nontrivial influence on the edge states. Recently it was proposed¹⁹ that in an inverted-band HgTe Hall bar, the edge channel transport can be switched off by narrowing the Hall bar by a single quantum point contact, where a stronger coupling between the edge states on opposite edges opens a sizable forbidden gap in the edge channel that blocks the electron transport within the energy window of the gap.

In this work we investigate theoretically the electron transport through periodic electric potential modulation in a HgTe Hall bar with a *constant width*. We find constructive interference of the backscattering by different modulation layers, leading to the formation of superlattice minigaps that block the edge state transmission. This effect can be further

enhanced by tuning the edge channel into resonance with the bulk states. The present mechanism is based on the superlattice minigap effect instead of the finite-size gap as utilized in a previous proposal.¹⁹ Consequently, our mechanism works well in the regime of a small finite-size gap and enables us to switch on/off the edge channel transport by appropriately tuning the modulation amplitude via gate voltages. In the presence of gate-induced Rashba spin-orbit interaction, efficient conversion between spin-up and spin-down channels is found.

The paper is organized as follows. In Sec. II we describe the model and the formalism of ballistic electron transport through the HgTe Hall bar, with the detailed derivations given in the Appendix. In Sec. III we present the numerical results to demonstrate the switch off and electron spin conversion of the spin-resolved edge channel transport at an appropriate modulation amplitude. Finally, the summary and conclusions are given in Sec. IV.

II. MODEL AND FORMALISM

We consider a narrow-gap HgTe QW with an inverted band structure in the Hall bar geometry, as shown schematically in Fig. 1(a). Through periodic electric gating, N periods of *tunable*, square-shaped potential modulation of height V_0 (or depth $|V_0|$ if $V_0 < 0$, see the shaded region in Fig. 1) along the transport direction, the x axis, is generated. The low-energy spectrum is well described by a four-band Hamiltonian, including two Γ_6 electron bands $|e\uparrow\rangle, |e\downarrow\rangle$ and two Γ_8 heavy hole bands $|hh\uparrow\rangle, |hh\downarrow\rangle$. The Hamiltonian $\hat{H}(x, \hat{k}_x, \hat{k}_y)$ for the above structure is the sum of the Hamiltonian $\hat{H}_{\text{QW}}(\hat{k}_x, \hat{k}_y)$ of the HgTe QW, the hard-wall confining potential $U(x, y) = 0$ for $y \in [0, W]$ and $U(x, y) = \infty$ elsewhere, the modulation potential $V(x)$, and the gate-induced Rashba spin-orbit interaction²⁰ (the linear Rashba term)

$$\hat{H}_{\text{RSOI}}(x, \hat{k}_x, \hat{k}_y) = i \frac{\alpha(x)\hat{k}_- + \hat{k}_-\alpha(x)}{2} |e\uparrow\rangle\langle e, \downarrow| + \text{H.c.}, \quad (1)$$

which leads to intrinsically nonlinear and even nonmonotonic Rashba spin splitting as a function of the in-plane momentum $\mathbf{k}_{\parallel} = (k_x, k_y)$, especially in narrow-gap QWs. Here the

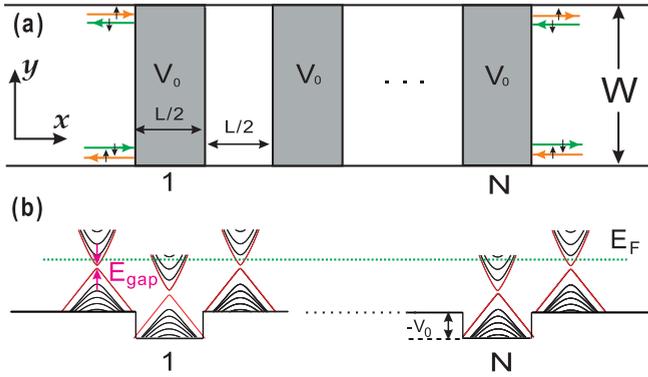


FIG. 1. (Color online) (a) HgTe QW with an inverted band structure in the Hall bar geometry, under N periods of potential modulation. (b) The bulk energy spectra of the HgTe QW in the regions with and without the potential modulation V_0 (tunable through electric gating), respectively. Here a small gap around E_{gap} is opened in the Dirac spectra of the edge states due to the finite-size effect, i.e., the finite width of the Hall bar. The green dashed line indicates the Fermi energy E_F .

piecewise constant functions $V(x) = V_0$ and $\alpha(x) = \alpha$ in the shaded regions of Fig. 1(a) and $V(x) = \alpha(x) = 0$ elsewhere. In the basis $|e\uparrow\rangle$, $|hh\uparrow\rangle$, $|e\downarrow\rangle$, and $|hh\downarrow\rangle$, the QW Hamiltonian⁹

$$\hat{H}_{\text{QW}}(\hat{k}_x, \hat{k}_y) = \begin{pmatrix} \varepsilon_{\hat{\mathbf{k}}} + M(\hat{\mathbf{k}}) & A\hat{k}_- & 0 & 0 \\ A\hat{k}_+ & \varepsilon_{\hat{\mathbf{k}}} - M(\hat{\mathbf{k}}) & 0 & 0 \\ 0 & 0 & \varepsilon_{\hat{\mathbf{k}}} + M(\hat{\mathbf{k}}) & -A\hat{k}_+ \\ 0 & 0 & -A\hat{k}_- & \varepsilon_{\hat{\mathbf{k}}} - M(\hat{\mathbf{k}}) \end{pmatrix},$$

where $\hat{\mathbf{k}} = (\hat{k}_x, \hat{k}_y)$ is the in-plane momentum operator, $\varepsilon_{\hat{\mathbf{k}}} = C - D(\hat{k}_x^2 + \hat{k}_y^2)$, $M(\hat{\mathbf{k}}) = M - B(\hat{k}_x^2 + \hat{k}_y^2)$, $\hat{k}_{\pm} = \hat{k}_x \pm i\hat{k}_y$, and A , B , C , D , and M are parameters describing the band structure of the HgTe/CdTe QWs. The topological insulator and trivial insulator phases are distinguished by the different signs of the parameter M , which in turn is determined by the thickness of the HgTe/CdTe QWs.⁹

The experimentally measurable conductance at zero temperature across the Hall bar structure is given by the Landauer-Büttiker formula as

$$G(E_F) = \frac{e^2}{h} \sum_{m \in L, n \in R} T_{n \leftarrow m}, \quad (2)$$

where $T_{n \leftarrow m}$ ($m \in L$, $n \in R$) is the transmission probability from the m th forward propagating channel in the left lead (denoted as $\psi_{+,m}^{(L)}$) to the n th forward propagating channel in the right lead (denoted as $\psi_{+,n}^{(R)}$). These transmission probabilities can be calculated by applying the scattering matrix formalism^{21–25} to the HgTe topological insulator in the Hall bar geometry under periodic modulations. The details are given in the Appendix.

In this paper we focus on the control of the edge state transmission, so E_F lies in the bulk gap of the left and right leads and only intersects the four edge channels in the leads. The Rashba spin-orbit interaction \hat{H}_{RSOI} has little influence on the edge states since \hat{H}_{RSOI} preserves the time-reversal symmetry and hence can only couple distant edge states on

opposite sides of the Hall bar.²⁰ In particular, all four edge channels on the Fermi level have well-defined spin orientations [as shown in Fig. 1(b)]: the spin-up, forward propagating channel $\psi_{+,\uparrow}$ and its time-reversed counterpart, the spin-down, backward propagating channel $\psi_{-,\downarrow}$ on the upper edge; as well as the spin-down, forward propagating channel $\psi_{+,\downarrow}$ and its time-reversed counterpart, the spin-up, backward propagating channel $\psi_{-,\uparrow}$ on the lower edge. Due to mirror reflection symmetry about the plane $y = W/2$, we have

$$\begin{aligned} T_{\uparrow \leftarrow \uparrow} &= T_{\downarrow \leftarrow \downarrow} = T_{\text{sc}}, \\ T_{\downarrow \leftarrow \uparrow} &= T_{\uparrow \leftarrow \downarrow} = T_{\text{sf}}, \end{aligned}$$

i.e., the spin-conserving transmission probabilities for the spin-up channel (from $\psi_{+,\uparrow}^{(L)}$ to $\psi_{+,\uparrow}^{(R)}$) and the spin-down channel (from $\psi_{+,\downarrow}^{(L)}$ to $\psi_{+,\downarrow}^{(R)}$) are equal (denoted as T_{sc}). Similarly, the spin-flip transmission probabilities from $\psi_{+,\uparrow}^{(L)}$ to $\psi_{+,\downarrow}^{(R)}$ and that from $\psi_{+,\downarrow}^{(L)}$ to $\psi_{+,\uparrow}^{(R)}$ are also equal (denoted as T_{sf}). Therefore, the total conductance $G(E_F) = (2e^2/h)(T_{\text{sc}} + T_{\text{sf}})$ is simply proportional to the total transmission probability $T \equiv T_{\text{sc}} + T_{\text{sf}}$ of a single incident edge channel.

III. RESULTS AND DISCUSSIONS

In this section we present our numerical results for the transmission probabilities $T, T_{\text{sc}}, T_{\text{sf}}$ of a single incident edge state (on the Fermi level E_F in the bulk gap of the leads) across N periods of potential modulations with amplitude V_0 , which can be tuned by varying the gate voltages. The transmission probabilities are connected to the experimentally measurable conductance G by the relation $G = (2e^2/h)(T_{\text{sc}} + T_{\text{sf}})$. We consider a $W = 200$ nm wide Hall bar under N modulations with the period $L = 100$ nm. The parameters for the HgTe/CdTe QWs are $A = 364.5$ meV nm, $B = -686$ meV nm², $C = 0$, $D = -512$ meV nm², and $M = -10$ meV.⁹ In this case the energy spectrum of the lead has a finite-size gap near $E_{\text{gap}} = 7.5$ meV, while the energy spectrum of the modulated region [shaded area in Fig. 1(a)] has a finite-size gap near $E_{\text{gap}} + V_0 = V_0 + 7.5$ meV. In contrast to a previous work,²⁰ where the Rashba spin-orbit coupling \hat{H}_{RSOI} has a negligible effect on the edge states, in our system with periodic modulation, \hat{H}_{RSOI} is able to induce almost complete spin-flip transmission $T_{\text{sf}} \approx 1$ once the edge states are tuned into resonance with the bulk states in the modulated region [shaded area in Fig. 1(a)], as discussed below. To clearly distinguish the superlattice minigap effect (which switches on/off the edge state transport) and the spin-flip effect by Rashba spin-orbit coupling (which converts between spin-up and spin-down edge channels), we first focus on the former by dropping \hat{H}_{RSOI} (so that $T_{\text{sf}} = 0$ and $T = T_{\text{sc}}$), leaving the inclusion of \hat{H}_{RSOI} at the end of our discussion.

A. Constructive interference of backscattering

To begin with, we illustrate the effect of periodic modulation by plotting the edge state transmission $T = T_{\text{sc}}$ across the Hall bar structure under successively increasing number of modulations $N = 1, 2, \dots, 6$ in Fig. 2. The sudden block of the edge state transmission around $V_0 = 5$ meV originates from the finite-size gap: At $V_0 = 5$ meV the finite-size gap of

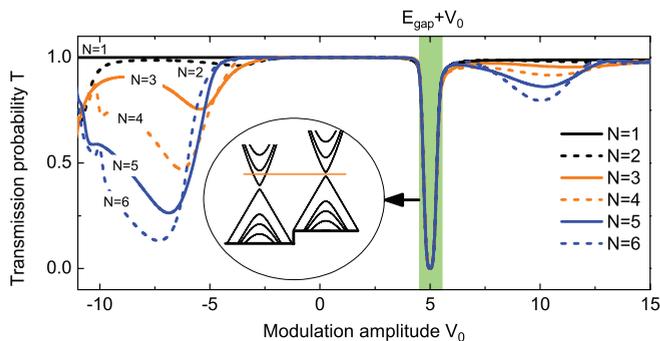


FIG. 2. (Color online) Transmission probability $T = T_{sc}$ of an incident edge state on the Fermi level $E_F = 12.5$ meV across successively increasing $N = 1, 2, \dots, 6$ periods of potential modulation as a function of modulation amplitude V_0 . Inset: For $V_0 = 5$ meV, the Fermi level E_F lies in the finite-size gap of the modulated region.

the modulated region [shaded area in Fig. 1(a)] locates around $E_{\text{gap}} + V_0 = 12.5$ meV, which coincides with the Fermi energy E_F , as sketched in the inset of Fig. 2. This mechanism has been used to confirm the intrinsic spin Hall effect²⁶ and to switch off the edge state transport by quantum point contact.²⁰ Here we propose a very different mechanism based on periodic modulation to control the edge state transport. The dramatic influence of this mechanism on the transmission can be clearly seen in Fig. 2. For a single modulation $N = 1$, the edge state transmission is nearly perfect $T \approx 1$ away from the finite-size gap, corresponding to a well quantized conductance $G = 2e^2/h$. With increasing N , successively pronounced dips gradually develop near $V_0 \approx 10$ meV and $V_0 \approx -8$ meV, respectively. These dips arise from the constructive interference of the backscattering amplitudes by different modulation layers when the incident wave vector k_x is an integer multiple of π/L : The extra phase $k_x \times 2L$ acquired in one round trip between successive modulation layers is an integer multiple of 2π , thus constructive interference occurs.

When the number N of modulation is further increased, we expect that the increasingly strong interferences of backscattering might give rise to superlattice minigaps near $V_0 \approx 10$ meV and $V_0 \approx -8$ meV that could block the edge state transport completely. To demonstrate this effect we show the contour plot of the transmission probability $T(V_0, N)$ as a function of the modulation amplitude V_0 and the number N of modulation in Fig. 3. In addition to the finite-size gap at $E_{\text{gap}} + V_0$ that blocks the edge state transmission for any modulation number N , the most striking feature is the appearance of two additional gaps (marked by Δ_{g1} and Δ_{g2}), where the edge state transmission decreases with increasing N and finally becomes blocked for large modulation numbers $N \gtrsim 15$ (for Δ_{g1}) or $N \gtrsim 8$ (for Δ_{g2}). As shown below, these two gaps Δ_{g1} and Δ_{g2} are both associated with the superlattice minigap effect, but their origin are quite different: Δ_{g1} comes from the finite-size gap in the leads, while Δ_{g2} results from the interference of the edge states with the bulk states in the modulated region.

B. Superlattice minigap from finite-size gap

To reveal the nature of the gap Δ_{g1} , we set $V_0 = 10$ meV at the center of the Δ_{g1} gap (see Fig. 3) and plot in Fig. 4(b)

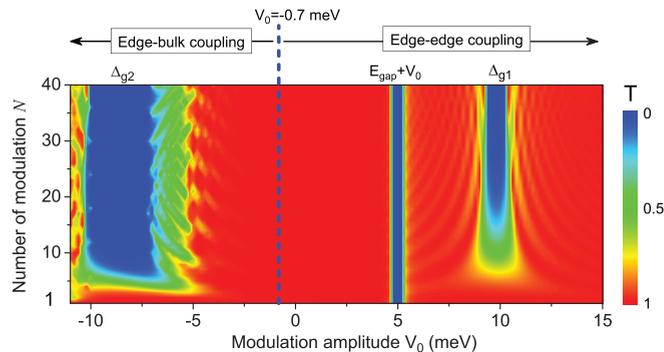


FIG. 3. (Color online) Contour plot of the transmission probability T of an incident edge state on the Fermi level $E_F = 12.5$ meV as a function of modulation amplitude V_0 and number of modulation N . For $V_0 < -0.7$ meV, the Fermi level enters the bulk energy bands in the modulation region [shaded area in Fig. 1(a)], as schematically shown in Fig. 1(b) (edge-bulk coupling). For $V_0 > -0.7$ meV, the Fermi level lies in the bulk gap (edge-edge coupling).

the transmission probability $T(E)$ across the Hall bar under $N = 30$ modulations as a function of the incident energy E . There are three transmission gaps. The first two are associated with the finite-size gap of the left lead and the modulated region, respectively: The first one around ~ 7.5 meV arises since $E_F = 7.5$ meV coincides with the finite-size gap around $E_{\text{gap}} = 7.5$ meV of the left lead, as sketched in the inset of Fig. 4(b); the second one around ~ 17.5 meV arises because $E_F = 17.5$ meV coincides with the finite-size gap around

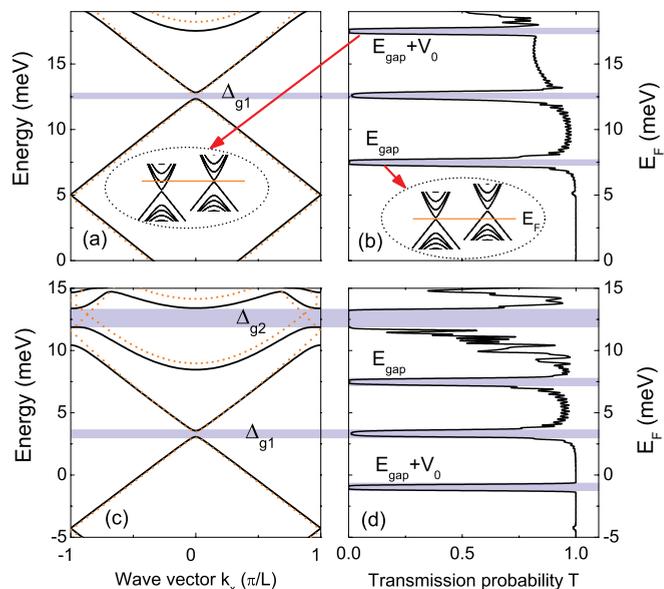


FIG. 4. (Color online) (a) and (c) The solid lines show the energy spectra of the Hall bar under an infinite number $N \rightarrow \infty$ of modulations; the dotted lines show the energy spectrum of the leads, shifted upward by $V_0/2$. (b) and (d) The transmission probability T of a single incident edge state with energy E across the Hall bar under $N = 30$ modulations. The modulation amplitude $V_0 = 10$ meV [corresponding to E_F in the bulk gap, as sketch in the inset of (b)] for (a), (b) and $V_0 = -8.5$ meV [corresponding to E_F entering the bulk energy spectra of the modulated region, as sketched in the inset of (d)] for (c), (d).

$E_{\text{gap}} + V_0 = 17.5$ meV of the modulated region [shaded area in Fig. 1(a)], as sketched in the inset of Fig. 4(a). Here we are interested in the third one, the minigap Δ_{g1} around ~ 12.5 meV, which comes from the constructive interference of backscattering by the periodic modulation, i.e., the superlattice minigap effect. This is confirmed by the complete coincidence (marked by the shaded area) of the Δ_{g1} gap in the transmission spectrum with the corresponding minigap of the superlattice energy spectrum in Fig. 4(a) for the Hall bar under an infinite number of periodic modulations $N \rightarrow \infty$.

To understand the origin of the superlattice minigap Δ_{g1} , we recall that the superlattice Hamiltonian is the sum of the lead Hamiltonian and the modulation potential $V(x) = V_0[1 + (-1)^j]/2$, where $j = 0$ for x in the modulated region [the shaded area in Fig. 1(a)] and $j = 1$ elsewhere. The constant part $V_0/2$ of $V(x)$ leads to a global upper shift of the lead energy spectrum, while the modulation part $V_0(-1)^j/2$ with period L weakly couples the $k_x = n\pi/L$ ($n = 1, 2, \dots$) edge states to the $k_x = -n\pi/L$ edge states with the same spin orientation on the opposite side of the Hall bar. This coupling could in principle opens up new minigaps at the boundaries $k_x = \pm\pi/L$ of the superlattice Brillouin zone. However, for energies in the bulk gap, the strong localization of the edge states (especially those below the Dirac point) significantly reduces this coupling and the resulting minigap is negligible, e.g., at $E \approx 5$ meV in Fig. 4(a). With both effects of $V(x)$ taken into account, the superlattice energy spectrum should largely coincide with the energy spectrum of the leads upper shifted by $V_0/2$, as confirmed by the good coincidence of the solid and dotted lines in Fig. 4(a).

To conclude, the minigap at $\Delta_{g1} \approx E_{\text{gap}} + V_0/2$ originates from the shift of the finite-size gap E_{gap} of the leads by the modulation potential. For $V_0 = 10$ meV, $\Delta_{g1} \approx 12.5$ meV coincides with the Fermi energy, so the transmission is blocked, as shown in Fig. 3. However, the minigap at Δ_{g1} is qualitatively different from the finite-size gap at E_{gap} . The former results from the constructive interference of backscattering from many modulation layers, so it exists and hence blocks the edge state transmission only when the number N of modulation is sufficiently large ($N \gtrsim 15$, see Fig. 3). By contrast, the latter comes from the finite-size effect of the Hall bar, so it always exists and blocks the transmission for any modulation number N .

C. Superlattice minigap from edge-bulk coupling

Now we investigate the nature of the gap Δ_{g2} by setting $V_0 = -8.5$ meV at the center of the Δ_{g2} gap (see Fig. 3). From Fig. 4(d) the transmission probability $T(E)$ for $N = 30$ shows four transmission gaps. The first two around ~ 7.5 and ~ -1 meV, respectively, are just the finite-size gap E_{gap} and $E_{\text{gap}} + V_0$ of the left lead and the modulated region. The third one is the superlattice minigap $\Delta_{g1} = E_{\text{gap}} + V_0/2 \sim 3.25$ meV, which originates from the finite-size gap $E_{\text{gap}} \sim 7.5$ meV of the left lead shifted downwards by $|V_0|/2 = 4.25$ meV by the modulation potential, as discussed in the previous subsection. In this subsection we focus on the fourth gap $\Delta_{g2} \sim 12.5$ meV, which is also associated with the superlattice minigap effect, as confirmed by the coincidence (marked by the shaded area) of the gap at Δ_{g2}

in the transmission spectrum with the corresponding minigap of the superlattice energy spectrum in Fig. 4(c). However, the origin of the gap at Δ_{g2} is very different from the gap at Δ_{g1} . Instead of coming from the shift of the finite-size gap in the leads, the Δ_{g2} gap is a truly “new” minigap resulting from the superlattice effect: As sketched in Fig. 1(b), the Fermi level E_F lies in the bulk gap of the unmodulated regions, but it enters the bulk energy spectrum of the modulated regions [shaded area in Fig. 1(a)]. Therefore, the resonant coupling between the edge state in the unmodulated region and the bulk states in the modulated region by the modulation potential greatly enhances the superlattice minigap effect and opens a relatively large gap at Δ_{g2} . Therefore, by tuning the modulation amplitude with gate voltages, the edge state transport can be switched on/off without relying on the finite-size gap²⁰ or breaking the time-reversal symmetry by introducing magnetic impurities or magnetic fields.

D. Spin conversion by gate-induced Rashba spin-orbit coupling

So far we have focused on the superlattice minigap effect by dropping the Rashba spin-orbit coupling \hat{H}_{RSOI} . In a realistic experimental setup, however, the electric gating above the HgTe QWs not only induces the modulation potential $V(x)$, but also gives rise to Rashba spin-orbit coupling \hat{H}_{RSOI} . In this subsection we present the edge state transmission including the Rashba spin-orbit coupling. Now the spin-flip transmission probability T_{sf} does not vanish and the total transmission probability $T = T_{\text{sc}} + T_{\text{sf}}$. These two probabilities for an incident edge state on the Fermi level $E_F = 12.5$ meV across the Hall bar under $N = 30$ modulations are plotted in Fig. 5. First, we consider the linear Rashba term [as described by Eq. (1)], we confirm from Fig. 5(a) that inclusion of \hat{H}_{RSOI} does not change the essential features of the previous results, including the finite-size gap around $V_0 \sim 5$ meV (corresponding to the finite-size gap at $E_{\text{gap}} + V_0 \approx 12.5$ meV of the modulated region coinciding with the Fermi level $E_F = 12.5$ meV), the superlattice minigap Δ_{g1} around $V_0 \sim 10$ meV originating from the finite-size gap of the leads (corresponding to the minigap at $\Delta_{g1} = E_{\text{gap}} + V_0/2 \approx 12.5$ meV coinciding with the Fermi level E_F), and the superlattice minigap Δ_{g2} around $V_0 \sim 8$ meV, which comes from the resonant coupling between the edge channel and the bulk states. Note that although the Rashba spin-orbit coupling has negligible influence on the edge states, it does influence the bulk states and hence the associated superlattice minigap Δ_{g2} . Second, in the presence of \hat{H}_{RSOI} (which behaves like an in-plane magnetic field), the spin-up and spin-down bulk states become strongly mixed and have slightly different Fermi wave vectors. Consequently, once the Fermi level E_F enters the bulk energy bands in the modulated region (edge-bulk coupling regime), the bulk states mediate efficient conversion between the spin-up edge channel and the spin-down edge channel, e.g., maximal conversion $T_{\text{sf}} \approx 1$ is achieved at $V_0 \sim -3$ meV. (Furthermore, we find that the number of modulation with the Rashba spin-orbit coupling must satisfy $N \gtrsim 28$, hence the totally spin-flip will be achieved.) This allows the realization of a spin-flip device by adjusting the gate voltages. By contrast, when the Fermi level E_F lies in the bulk gap, \hat{H}_{RSOI} only induces a weak coupling between edge states on opposite sides of the Hall bar, so it

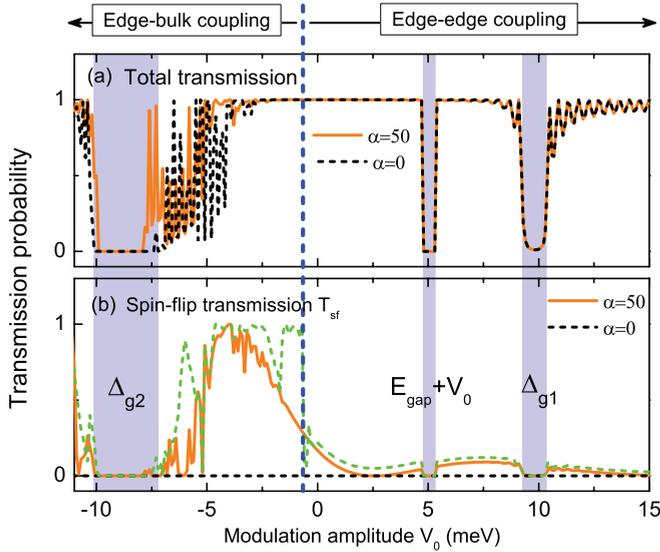


FIG. 5. (Color online) (a) Total transmission probability T and (b) the spin-flip transmission probability T_{sf} of an incident edge state on the Fermi level $E_F = 12.5$ meV across the Hall bar under $N = 30$ modulation periods. The gate-induced Rashba spin-orbit coupling strength (the linear Rashba term) is taken as $\alpha = 50$ meV nm (orange, solid curve) and $\alpha = 0$ (black, dashed curve), respectively. The green dashed curve indicates the spin-flip transmission probability T_{sf} with including the nonlinear Rashba terms. The vertical, blue dashed line distinguishes the edge-edge coupling regime (E_F in the bulk gap) and the edge-bulk coupling regime (E_F entering the bulk energy bands of the modulated region).

has a small influence on the edge state transmission. However, additionally to the linear Rashba term, there are other nonlinear Rashba terms present for HgTe QWs as investigated in Ref. 27. Therefore, we plot the edge state transmission probability T_{sf} in Fig. 5(b) (see the green dashed curve) with the nonlinear Rashba term. One can see that maximal conversion $T_{sf} \approx 1$ is achieved at $V_0 \sim -4, -3, -2.5$, and -1 meV (with $N \gtrsim 25$); the efficiency of the spin-flip processes will be increased with consideration of the nonlinear Rashba terms. That is because the quadratic Rashba terms will couple the electron and heavy hole while the cubic term will couple the heavy hole with different spins.

IV. CONCLUSIONS

In summary, we have investigated theoretically the edge state transport in an inverted-band HgTe topological insulator in the Hall bar geometry under gate-induced periodic potential modulation. We find constructive interference of the backscattering amplitudes by different modulation layers, leading to two different kinds of superlattice minigaps that block the edge state transmission: One comes from the shift of the finite-size gap in the lead energy spectrum, the other comes from the enhanced superlattice minigap effect by the resonant coupling between the edge state and the bulk states. In the presence of gate-induced Rashba spin-orbit coupling, we find efficient conversion between the spin-up and spin-down edge channels. This modulation induced switch on/off of the edge state transmission and spin conversion offers the possibility to

control the spin-resolved edge state transport by tuning the gate voltages, without relying on the finite-size effect or introducing magnetic impurities or magnetic fields. It may pave the way to constructing low-power information processing devices using the topological edge states.

ACKNOWLEDGMENTS

This work is supported by the NSFC Grants No. 11274036 and No. 11004017, and the Scientific Research Fund of Hunan Provincial Education Department 12B010.

APPENDIX: TRANSMISSION FROM SCATTERING MATRIX FORMALISM

In this Appendix we give the detailed derivation of the reflection and transmission coefficients of an incident eigenstate on the Fermi level E_F across the Hall bar structure (Fig. 1) using the scattering matrix formalism.²¹⁻²⁵ For this purpose we divide the entire Hall bar into $2N + 1$ segments separated by $2N$ interfaces x_1, \dots, x_{2N} , as sketched in Fig. 6. In particular, the left lead corresponds to the 0th segment and the right lead corresponds to the $2N$ th segment. In each segment, $V(x)$ and $\alpha(x)$ take constant values: $V(x) = V_0$ and $\alpha(x) = \alpha$ in the modulated region (shaded area) and $V(x) = \alpha(x) = 0$ elsewhere. The essential idea is to first solve the Schrödinger equation at the Fermi energy E_F in each segment and then match the wave functions of different segments at the $2N$ interfaces.

For the first step, we solve the Schrödinger equation inside each segment. In the j th segment, the total 4×4 Hamiltonian $\mathbf{H}^{(j)}(\hat{k}_x, \hat{k}_y)$ is independent of x , so the four-component eigenmode $\psi^{(j)}(x, y) = e^{ik(x-x_j)}\phi^{(j)}(y)$ (for $j = 0$, we define $x_0 \equiv x_1$), where $\phi^{(j)}(y)$ is obtained from the Schrödinger equation $E_F \phi^{(j)}(y) = \mathbf{H}^{(j)}(k, \hat{k}_y)\phi^{(j)}(y)$, with \hat{k}_x replaced by its eigenvalue k . The hard-wall boundary conditions along the upper edge $y = W$ and lower edge $y = 0$ enables the expansion of $\phi^{(j)}(y)$ using the j -independent complete basis $\varphi_n(y) = \sqrt{2/W} \sin(n\pi y/W)$ ($n = 1, 2, \dots, N_{\text{cut}}$, with a sufficiently large N_{cut} to ensure convergence) as $\phi^{(j)}(y) = \sum_n \mathbf{c}_n^{(j)} \varphi_n(y)$, where $\mathbf{c}_n^{(j)}$ is a four-component vector. This gives an eigenvalue problem

$$\mathbf{H}^{(j)}(k)\mathbf{C}^{(j)} = E_F \mathbf{C}^{(j)} \quad (\text{A1})$$

with dimension $M \equiv 4N_{\text{cut}}$, where $\mathbf{C}^{(j)} \equiv [\mathbf{c}_1^{(j)}, \dots, \mathbf{c}_{N_{\text{cut}}}^{(j)}]^T$ is a M -component vector and $\mathbf{H}^{(j)}(k)$ is a $M \times M$ square matrix with $\mathbf{H}_{m,n}^{(j)}(k) \equiv \langle \varphi_m | \mathbf{H}^{(j)}(k, \hat{k}_y) | \varphi_n \rangle$. For a given real k there are M eigenvalues $E^{(j)}(k)$ and eigenvectors $\mathbf{C}^{(j)}(k)$. Conversely, for a given real energy E_F , there are $2M$ solutions

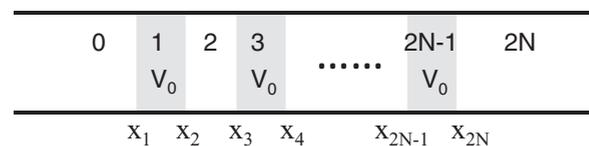


FIG. 6. Division of the Hall bar into $2N + 1$ segments (denoted by $j = 0, 1, \dots, 2N$) by the N periods of modulation potential. The interfaces coordinates are denoted by x_1, x_2, \dots, x_{2N} .

$k^{(j)}(E_F)$ (which are in general complex) and $\mathbf{C}^{(j)}(E_F)$. To find these solutions we expand $\mathbf{H}^{(j)}(k) = \mathbf{H}_0^{(j)} + \mathbf{H}_1^{(j)}k + \mathbf{H}_2^{(j)}k^2$ into power of k and define $\tilde{\mathbf{C}}^{(j)} \equiv k\mathbf{C}^{(j)}$, then Eq. (A1) can be written as

$$\begin{bmatrix} 0 & 1 \\ E_F - \mathbf{H}_0^{(j)} & -\mathbf{H}_1^{(j)} \end{bmatrix} \begin{bmatrix} \mathbf{C}^{(j)} \\ \tilde{\mathbf{C}}^{(j)} \end{bmatrix} = k \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{H}_2^{(j)} \end{bmatrix} \begin{bmatrix} \mathbf{C}^{(j)} \\ \tilde{\mathbf{C}}^{(j)} \end{bmatrix}.$$

Solving this $2M$ -dimensional generalized eigenvalue problem gives $2M$ solutions $\{k_m^{(j)}, \mathbf{C}_m^{(j)}\}$ ($m = 1, 2, \dots, 2M$), corresponding to $2M$ eigenmodes $\psi_m^{(j)}$ on the Fermi level E_F . Next we introduce the velocity operator $\mathbf{v}^{(j)}(\hat{k}_x) = -i[x, \mathbf{H}^{(j)}(\hat{k}_x, \hat{k}_y)]$ along the transport direction (x axis):

$$\mathbf{v}^{(j)}(\hat{k}_x) = \begin{bmatrix} -2D_+\hat{k}_x & A & i\alpha^{(j)} & 0 \\ A & -2D_-\hat{k}_x & 0 & 0 \\ -i\alpha^{(j)} & 0 & -2D_+\hat{k}_x & -A \\ 0 & 0 & -A & -2D_-\hat{k}_x \end{bmatrix} \quad (\text{A2})$$

(with $D_{\pm} \equiv D \pm B$) and classify these $2M$ solutions into M forward-going solutions $\{k_{+\alpha}^{(j)}, \mathbf{C}_{+\alpha}^{(j)}\}$ (corresponding to M forward-going eigenmodes $\psi_{+\alpha}^{(j)}$) and M backward-going solutions $\{k_{-\alpha}^{(j)}, \mathbf{C}_{-\alpha}^{(j)}\}$ (corresponding to M backward-going eigenmodes $\psi_{-\alpha}^{(j)}$). The forward-going eigenmodes include forward propagating states (with a real $k^{(j)}$ and positive velocity $\langle \psi^{(j)} | \mathbf{v}^{(j)} | \psi^{(j)} \rangle$) and forward decaying states (with $\text{Im}k^{(j)} > 0$). The backward-going eigenmodes include backward propagating states (with a real $k^{(j)}$ and negative velocity $\langle \psi^{(j)} | \mathbf{v}^{(j)} | \psi^{(j)} \rangle$) and backward decaying states (with $\text{Im}k^{(j)} < 0$).

For the second step, we calculate the scattering state ψ emanating from the α th forward propagating incident eigenmode $\psi_{+\alpha}^{(0)}$ of the left lead on the Fermi level E_F . In the j th segment, the scattering state can be expanded as the linear combination of local eigenmodes on the Fermi level:

$$\psi^{(j)} = \sum_{\beta} a_{\beta \leftarrow \alpha}^{(j)} \psi_{+\beta}^{(j)} + \sum_{\beta} b_{\beta \leftarrow \alpha}^{(j)} \psi_{-\beta}^{(j)},$$

subjected to the boundary conditions $a_{\beta \leftarrow \alpha}^{(0)} = \delta_{\alpha, \beta}$ in the 0th segment (i.e., only one incident eigenmode $\psi_{+\alpha}^{(0)}$ in the left lead) and $b_{\beta \leftarrow \alpha}^{(2N)} = 0$ in the $2N$ th segment (i.e., no incident eigenmodes in the right lead). Altogether there are $M + M + (2N - 1) \times 2M = 4NM$ unknown coefficients. These coefficients, especially the reflection and transmission coefficients $\{b_{\beta \leftarrow \alpha}^{(0)}\}$ and $\{a_{\beta \leftarrow \alpha}^{(2N)}\}$ can be obtained (see the next paragraph) by matching ψ and $\hat{v}\psi$ at the $2N$ interfaces x_1, \dots, x_{2N} , which gives $2N \times 2M = 4NM$ equations. Then the transmission probability from the α th forward propagating eigenmode $\psi_{+\alpha}^{(0)}$ in the left lead to the β th forward propagating eigenmode $\psi_{+\beta}^{(2N)}$ in the right lead is given by $T_{\beta \leftarrow \alpha} = (v_{\beta}^{(2N)}/v_{\alpha}^{(0)})|a_{\beta \leftarrow \alpha}^{(2N)}|^2$, where $v_{\alpha}^{(j)} = \langle \psi_{+\alpha}^{(j)} | \mathbf{v}^{(j)} | \psi_{+\alpha}^{(j)} \rangle$ is the velocity of the α th forward propagating eigenmode in the j th segment. The zero-temperature conductance follows from the Landauer-Büttiker

formula as

$$G(E_F) = \frac{e^2}{h} \sum_{\alpha \in L, \beta \in R} T_{\beta \leftarrow \alpha}, \quad (\text{A3})$$

where the sum runs over all spin-resolved forward propagating eigenmodes in the left (L) and right (R) leads.

Now we give the explicit procedures for solving the unknown coefficients in the scattering matrix formalism.^{21–25} At the $(j + 1)$ th interface, the continuity conditions $\psi^{(j+1)}|_{x_{j+1}} = \psi^{(j)}|_{x_{j+1}}$ and $\mathbf{v}^{(j+1)}\psi^{(j+1)}|_{x_{j+1}} = \mathbf{v}^{(j)}\psi^{(j)}|_{x_{j+1}}$ give $2M$ equations that connect the M forward-going coefficients $\mathbf{A}_{\alpha}^{(j)} \equiv [a_{1 \leftarrow \alpha}^{(j)}, a_{2 \leftarrow \alpha}^{(j)}, \dots, a_{M \leftarrow \alpha}^{(j)}]^T$ and the M backward-going coefficients $\mathbf{B}_{\alpha}^{(j)} \equiv [b_{1 \leftarrow \alpha}^{(j)}, b_{2 \leftarrow \alpha}^{(j)}, \dots, b_{M \leftarrow \alpha}^{(j)}]^T$ in the j th segment to those in the $(j + 1)$ th segment:

$$\begin{bmatrix} \mathbf{A}_{\alpha}^{(j)} \\ \mathbf{B}_{\alpha}^{(j)} \end{bmatrix} = \mathbf{M}(j, j+1) \begin{bmatrix} \mathbf{A}_{\alpha}^{(j+1)} \\ \mathbf{B}_{\alpha}^{(j+1)} \end{bmatrix}, \quad (\text{A4})$$

where

$$\mathbf{M}(j, j+1) = \begin{bmatrix} \gamma_{+}^{(j)} & 0 \\ 0 & \gamma_{-}^{(j)} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{P}_{+}^{(j)} & \mathbf{P}_{-}^{(j)} \\ \mathbf{Q}_{+}^{(j)} & \mathbf{Q}_{-}^{(j)} \end{bmatrix}^{-1} \times \begin{bmatrix} \mathbf{P}_{+}^{(j+1)} & \mathbf{P}_{-}^{(j+1)} \\ \mathbf{Q}_{+}^{(j+1)} & \mathbf{Q}_{-}^{(j+1)} \end{bmatrix}$$

is the $2M \times 2M$ dimensional transfer matrix from the j th segment to the $(j + 1)$ th segment and $\gamma_{\pm}^{(j)}$, $\mathbf{P}_{\pm}^{(j)}$, $\mathbf{Q}_{\pm}^{(j)}$ are all $M \times M$ square matrices: $\gamma_{\pm}^{(j)} = \text{diag}\{e^{ik_{\pm,1}^{(j)}L_j}, \dots, e^{ik_{\pm,M}^{(j)}L_j}\}$ is diagonal with $L_j \equiv x_{j+1} - x_j$ being the length of the j th segment (see Fig. 6), the β th column of $\mathbf{P}_{\pm}^{(j)}$ is $\mathbf{C}_{\pm, \beta}^{(j)}$, and the β th column of $\mathbf{Q}_{\pm}^{(j)}$ is $\mathbf{V}^{(j)}(k_{\pm, \beta}^{(j)})\mathbf{C}_{\pm, \beta}^{(j)}$, where $\mathbf{V}^{(j)}(k) = \text{diag}\{\mathbf{v}^{(j)}(k), \dots, \mathbf{v}^{(j)}(k)\}$ consists of N_{cut} identical 4×4 blocks $\mathbf{v}^{(j)}(k)$ on its diagonal, with $\mathbf{v}^{(j)}(k)$ obtained from Eq. (A2) by replacing \hat{k}_x with k .

Equation (A4) can be iterated to give

$$\begin{bmatrix} \mathbf{A}_{\alpha}^{(0)} \\ \mathbf{B}_{\alpha}^{(0)} \end{bmatrix} = \mathbf{M}(0, 2N) \begin{bmatrix} \mathbf{A}_{\alpha}^{(2N)} \\ \mathbf{B}_{\alpha}^{(2N)} \end{bmatrix},$$

where $\mathbf{M}(0, 2N) = \mathbf{M}(0, 1)\mathbf{M}(1, 2) \dots \mathbf{M}(2N - 1, 2N)$ is the total transfer matrix. Since $(\mathbf{A}_{\alpha}^{(0)})_{\beta} = \delta_{\alpha, \beta}$ and $\mathbf{B}_{\alpha}^{(2N)} = 0$ are known, the above equation can be rearranged to give the reflection and transmission coefficients $\mathbf{B}_{\alpha}^{(0)}$ and $\mathbf{A}_{\alpha}^{(2N)}$:

$$\begin{bmatrix} \mathbf{A}_{\alpha}^{(2N)} \\ \mathbf{B}_{\alpha}^{(0)} \end{bmatrix} = \mathbf{S}(0, 2N) \begin{bmatrix} \mathbf{A}_{\alpha}^{(0)} \\ \mathbf{B}_{\alpha}^{(2N)} \end{bmatrix},$$

where the scattering matrix

$$\mathbf{S}(i, j) = \begin{bmatrix} \mathbf{S}_{11}(i, j) & \mathbf{S}_{12}(i, j) \\ \mathbf{S}_{21}(i, j) & \mathbf{S}_{22}(i, j) \end{bmatrix}$$

is connected to the transfer matrix

$$\mathbf{M}(i, j) \equiv \begin{bmatrix} \mathbf{M}_{11}(i, j) & \mathbf{M}_{12}(i, j) \\ \mathbf{M}_{21}(i, j) & \mathbf{M}_{22}(i, j) \end{bmatrix}$$

through [omitting the arguments (i, j) for brevity]

$$\mathbf{S}_{11} = \mathbf{M}_{11}^{-1}, \quad (\text{A5a})$$

$$\mathbf{S}_{12} = -\mathbf{M}_{11}^{-1}\mathbf{M}_{12}, \quad (\text{A5b})$$

$$\mathbf{S}_{21} = \mathbf{M}_{21}\mathbf{M}_{11}^{-1}, \quad (\text{A5c})$$

$$\mathbf{S}_{22} = -\mathbf{M}_{21}\mathbf{M}_{11}^{-1}\mathbf{M}_{12} + \mathbf{M}_{22}. \quad (\text{A5d})$$

However, in the presence of evanescent eigenmodes, the exponentially decaying term $\boldsymbol{\gamma}_+^{(j)}$ and the exponentially growing term $\boldsymbol{\gamma}_-^{(j)}$ coexist in the transfer matrix formalism, making the calculation of $\mathbf{M}(0, 2N)$ and hence $\mathbf{S}(0, 2N)$ from

Eqs. (A5) numerically unstable. An efficient way to avoid this problem is to calculate the scattering matrix from the iterative relations^{21–24} starting from $\mathbf{S}(0, 0) = 1$:

$$\mathbf{S}_{11}(0, i+1) = (1 - \mathbf{M}_{11}^{-1}\mathbf{S}_{12}\mathbf{M}_{21})^{-1}\mathbf{M}_{11}^{-1}\mathbf{S}_{11},$$

$$\mathbf{S}_{12}(0, i+1) = (1 - \mathbf{M}_{11}^{-1}\mathbf{S}_{12}\mathbf{M}_{21})^{-1}\mathbf{M}_{11}^{-1}(\mathbf{S}_{12}\mathbf{M}_{22} - \mathbf{M}_{12}),$$

$$\mathbf{S}_{21}(0, i+1) = \mathbf{S}_{22}\mathbf{M}_{21}\mathbf{S}_{11}(0, i+1) + \mathbf{S}_{21},$$

$$\mathbf{S}_{22}(0, i+1) = \mathbf{S}_{22}\mathbf{M}_{21}\mathbf{S}_{12}(0, i+1) + \mathbf{S}_{22}\mathbf{M}_{22},$$

where $\mathbf{S}_{ij} \equiv \mathbf{S}_{ij}(0, i)$ and $\mathbf{M}_{ij} \equiv \mathbf{M}_{ij}(i, i+1)$ on the right-hand side.

*lzlin@semi.ac.cn

†wenyang@csrc.ac.cn

¹M. Z. Hasan and C. L. Kane, *Rev. Mod. Phys.* **82**, 3045 (2010).

²X.-L. Qi and S.-C. Zhang, *Rev. Mod. Phys.* **83**, 1057 (2011).

³C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).

⁴D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, *Phys. Rev. Lett.* **97**, 036808 (2006).

⁵L. Fu, C. L. Kane, and E. J. Mele, *Phys. Rev. Lett.* **98**, 106803 (2007).

⁶L. Fu and C. L. Kane, *Phys. Rev. B* **76**, 045302 (2007).

⁷J. E. Moore and L. Balents, *Phys. Rev. B* **75**, 121306 (2007).

⁸C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).

⁹B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, *Science* **314**, 1757 (2006).

¹⁰C. Liu, T. L. Hughes, X.-L. Qi, K. Wang, and S.-C. Zhang, *Phys. Rev. Lett.* **100**, 236601 (2008).

¹¹M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, *Science* **318**, 766 (2007).

¹²A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, *Science* **325**, 294 (2009).

¹³I. Knez, R.-R. Du, and G. Sullivan, *Phys. Rev. Lett.* **107**, 136603 (2011).

¹⁴M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, *J. Phys. Soc. Jpn.* **77**, 031007 (2008).

¹⁵C. Xu and J. E. Moore, *Phys. Rev. B* **73**, 045322 (2006).

¹⁶J. J. Zhu, D. X. Yao, S. C. Zhang, and K. Chang, *Phys. Rev. Lett.* **106**, 097201 (2011).

¹⁷J. E. Moore, *Nature (London)* **464**, 194 (2010).

¹⁸K. Chang and W. K. Lou, *Phys. Rev. Lett.* **106**, 206802 (2011); Z. H. Wu, F. M. Peeters, and K. Chang, *Appl. Phys. Lett.* **98**, 162101 (2011); J. S. Sheng and K. Chang, *Phys. Rev. B* **74**, 235315 (2006).

¹⁹L. B. Zhang, F. Cheng, F. Zhai, and K. Chang, *Phys. Rev. B* **83**, 081402(R) (2011).

²⁰L. B. Zhang, F. Zhai, and K. Chang, *Phys. Rev. B* **81**, 235323 (2010).

²¹D. Y. K. Ko and J. C. Inkson, *Phys. Rev. B* **38**, 9945 (1988).

²²H. Xu, *Phys. Rev. B* **47**, 9537 (1993).

²³H. Xu, *Phys. Rev. B* **50**, 8469 (1994).

²⁴H. Xu, *Phys. Rev. B* **52**, 5803 (1995).

²⁵L. Zhang, P. Brusheim, and H. Q. Xu, *Phys. Rev. B* **72**, 045347 (2005).

²⁶C. 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).

²⁷D. G. Rothe, R. W. Reinthaler, C.-X. Liu, L. W. Molenkamp, S.-C. Zhang, and E. M. Hankiewicz, *New J. Phys.* **12**, 065012 (2010).