Coexistence and decoupling of bulk and edge states in disordered two-dimensional topological insulators

Yan-Yang Zhang,[1,2] Man Shen,[1,3] Xing-Tao An,[4] Qing-Feng Sun,[5] Xin-Cheng Xie,[5] Kai Chang,[1,2] and Shu-Shen Li[1,2]

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

[2]*Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

[3]*Department of Physics and Hebei Advanced Thin Film Laboratory, Hebei Normal University, Hebei 050024, China*

[4]*School of Sciences, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China*

[5]*International Center for Quantum Materials and School of Physics, Peking University, Beijing 100871, China*

(Received 20 February 2014; revised manuscript received 4 July 2014; published 18 August 2014)

We investigate the scattering and localization properties of edge and bulk states in a disordered two-dimensional topological insulator when they coexist at the same Fermi energy. Due to edge-bulk backscattering (which is not prohibited *a priori* by topology or symmetry), Anderson disorder makes the edge and bulk states localized indistinguishably. Two methods are proposed to effectively decouple them and to restore robust transport. The first kind of decouple is from long-range disorder since edge and bulk states are well separated in k space. The second one is from an edge gating, owing to the edge nature of edge states in real space. The latter can be used to electrically tune a system between an Anderson insulator and a topologically robust conductor, i.e., a realization of a topological transistor.

DOI: 10.1103/PhysRevB.90.054205 PACS number(s): 72.20.−i, 71.23.−k, 73.20.−r, 73.40.−c

I. INTRODUCTION

What makes topological insulators (TIs) unusual are the boundary states carrying dissipationless currents. This novel property is characterized by a topological invariant (e.g., Chern number or Z_2 invariant) associated with the occupied bulk bands. In the presence of boundaries (surfaces or edges), a nontrivial topological invariant guarantees the existence of gapless boundary states connecting the conduction and valance bands [1–3]. In the bulk gap, the absence of backscattering between boundary states is protected by the topological order and some symmetry. Topological states in two-dimensional (2D) [4–7] and three-dimensional (3D) [8,9] systems have been experimentally observed recently.

However, the robust transport properties of boundary states are only valid without the interference of bulk states. Unfortunately, most of the 3D TIs found so far are actually metals, i.e., with Fermi energy in the bulk band coexisting with boundary states [1,2]. This is one of the biggest experimental obstacles to realizing boundary transport in TIs. It has been observed that the mix between boundary and bulk states will lead to remarkable backscattering [1,9–11]. This backscattering can destroy the perfect conducting of boundary states. Therefore, further understanding of boundary-bulk interplay is necessary. In Refs. [12,13], Fano-type rearrangements of bulk and boundary spectra arising from the mixing was discussed. Nevertheless, the scattering and transport properties due to disorder is not clear: How does such scattering happen? To what extent does it affect the robustness of boundary states? Most importantly, how can we avoid it?

In this paper, we systematically investigate the scattering between edge and bulk states in 2D TIs, in the presence of nonmagnetic impurities. We found that with the coexistence of edge and bulk states at the same Fermi energy, (1) Anderson disorder tends to localize them indistinguishably; (2) long-range impurities can effectively decouple them and restore perfectly conducting channels (PCCs); (3) a local voltage gating at the edge is also enough to effectively decouple them and leads to PCCs, therefore giving rise to a convenient way to switch the system between a localized Anderson insulator and a perfect conducting TI. This is a concrete proposal of a topological field effect transistor (FET) [14,15].

II. MODEL AND METHODS

We adopt the Kane-Mele–type [16] Hamiltonian defined on a honeycomb lattice

$$\begin{aligned} H = & t\sum_{\langle ij\rangle,\sigma} c^{\dagger}_{i\sigma}c_{j\sigma} + \lambda_v \sum_{i,\sigma} \xi_i c^{\dagger}_{i\sigma}c_{i\sigma} \\ & + i\frac{\lambda_{\mathrm{SO}}}{3\sqrt{3}} \sum_{\langle\langle ij\rangle\rangle,\sigma\sigma'} \nu_{ij} c^{\dagger}_{i\sigma} s^{z}_{\sigma\sigma'} c_{j\sigma'} \\ & - i\frac{2\lambda_{\mathrm{R}}}{3} \sum_{\langle\langle ij\rangle\rangle,\sigma\sigma'} \mu_{ij} c^{\dagger}_{i\sigma} (\boldsymbol{s}\times\hat{\boldsymbol{d}}_{ij})^{z}_{\sigma\sigma'} c_{j\sigma'}, \end{aligned} \tag{1}$$

which has been used to describe the electronic states in silicene [17–19]. The first term describes the nearest-neighbor (NN) hopping, where $c^{\dagger}_{i\sigma}$ creates an electron at site i with spin σ. The second term represents the staggered potential with $\xi_i = \pm 1$ for sublattice A (B). The third term is the intrinsic spin-orbital coupling (SOC) between the near-nearest-neighbor (NNN) sites, where $\boldsymbol{s} = (s_x, s_y, s_z)$ are the Pauli matrices for physical spins, and $\nu_{ij} = (\boldsymbol{d}_i \times \boldsymbol{d}_j)_z/|\boldsymbol{d}_i \times \boldsymbol{d}_j| = \pm 1$ with $\boldsymbol{d}_i$ and $\boldsymbol{d}_j$ the two NN bonds connecting NNN sites i and j. The fourth term is the NNN Rashba SOC, where $\mu_{ij} = \pm 1$ for the A and B sites, respectively, and $\hat{\boldsymbol{d}}_{ij} = \boldsymbol{d}_{ij}/|\boldsymbol{d}_{ij}|$ representing the unit vector of $\boldsymbol{d}_{ij}$ which connects NNN sites i and j. Finite Rashba SOC breaks the conservation of spin σ_z, but does not change the bulk gap, and therefore does not change the Z_2 invariant. We will take $\lambda_{\mathrm{R}} = 0$ in most of this paper. In Sec. VI, it will be shown that finite λ_{R} does not affect the results we have concluded. Hereafter, we adopt t

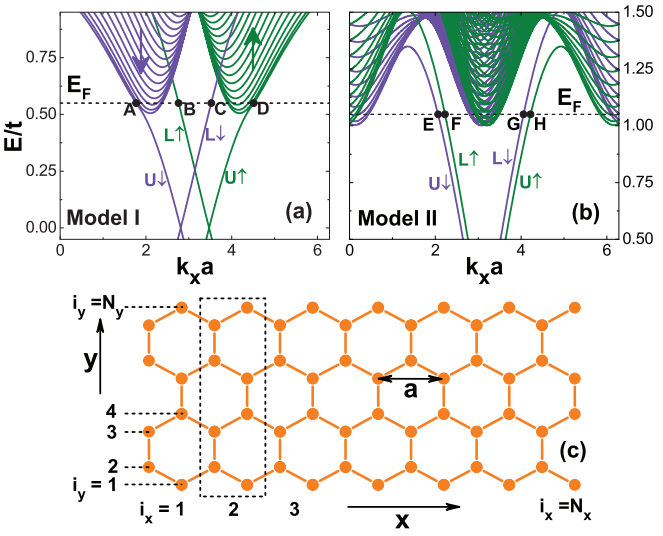

FIG. 1. (Color online) (a) Band structure of model I ($\lambda_{SO} = 0.8t$, $\lambda_v = 0.3t$, and $\lambda_R = 0$). (b) Band structure of model II ($\lambda_{SO} = 1.6t$, $\lambda_v = 0.1t$, and $\lambda_R = 0$). Both models are particle-hole symmetric so only conduction bands are plotted. Both (a) and (b) are from a zigzag nanoribbon with width $N_y = 80$. States with spin ↑ (↓) are plotted in green (purple). Edge states associated with lower (L) or upper (U) edge are also marked. (c) Schematic of the zigzag edge nanoribbon with length N_x and width N_y. The unit cell is marked by the black dashed line.

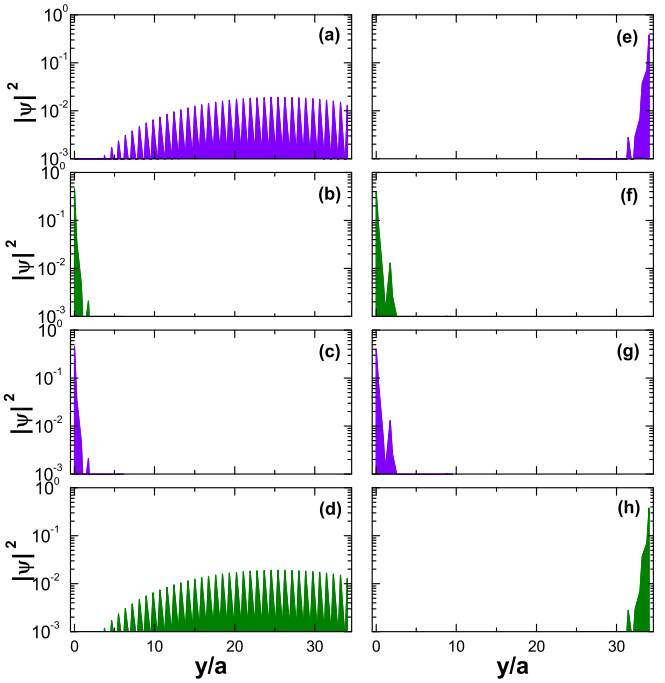

FIG. 2. (Color online) The square magnitudes of the wave functions $|\psi(y)|^2$ along the cross section of the $Ny = 80$ ribbon, for model I at $E_F = 0.55t$ (left column), and for model II at $E_F = 1.05t$ (right column), with (a) to (h) corresponding to states A to H in Figs. 1(a) and 1(b), respectively.

as the energy unit and lattice constant a [NNN distance, see Fig. 1(c)] as the length unit.

The Hamiltonian (1) respects time-reversal symmetry as well as particle-hole symmetry. Its electronic structure when $|\lambda_{SO}| < t$ has been well studied [18,19]. The bulk bands are gapped at Dirac points $K(K') = (\pm 4\pi/3, 0)$ with magnitude $\Delta_G = 2|\lambda_{SO} - \lambda_v|$. In the finite gap, the system is a topologically trivial (nontrivial) insulator if $\lambda_{SO} < \lambda_v$ ($\lambda_{SO} > \lambda_v$). In Fig. 1(a), we plot the typical band structure of a quasi-one-dimensional (1D) ribbon in the topological nontrivial phase. The model parameters are $\lambda_{SO} = 0.8t$, $\lambda_v = 0.5t$, and $\lambda_R = 0$, which will be referred as model I in the following. Within the bulk gap, $|E| < |\lambda_{SO} - \lambda_v|$, there are four gapless edge states, corresponding to lower (L) and upper (U) edges, with spin up (↑) and down (↓), respectively. In the presence of time-reversal symmetry (TRS) and sufficiently ribbon width, the backscattering between them is prohibited [20]. This is a typical model of the quantum spin Hall effect.

The interests in this work will be in the energy region of bulk bands. As shown in the band structure in Fig. 1(a), when $E > \lambda_{SO} - \lambda_v = 0.5t$, one pair of edge states (U↑ and U↓) at the upper edge merge into the continuum of bulk states, while another pair of edge states (L↑ and L↓) at the lower edge are still well separated from bulk states. In Figs. 2(a)–2(d), we plot their wave-function distributions in the real space, corresponding to states A, B, C, and D at $E_F = 0.55t$ in Fig. 1(a). Indeed, for states B (L↑) and C (L↓), their wave functions are extremely localized at the lower edge $y = 0$. On the other hand, states A (extension of U↓) and D (extension of U↑) spread over most of the bulk with strong oscillations [21], and lose their edge nature. Hereafter, states like A and D will be called "pseudo-edge states" and will be classified as bulk states. As will be seen in Table II in the following, their responses to disorder are similar with other ordinary bulk states. In other words, what we mean by "edge states" in the rest of this paper are those states like B and C, which are *well defined* and *distinguishable* from the bulk, both in real space and in k space. Our focus will be their scattering properties when such edge states are coexisting with bulk states at the same Fermi energy. Other 2D TI systems (e.g., see Fig. 11) with only bulk states (including pseudo-edge states) in the energy region of bulk bands are out of this context, which will be discussed in Sec. VII.

Model I has only edge states at the lower edge coexisting with the bulk. Other more "natural" configurations of edge states with Hamiltonian (1) can arise when $\lambda_{SO} > t$. Figure 1(b) is the nanoribbon band structure of a $N_y = 80$ ribbon for model II, with parameters $\lambda_{SO} = 1.6t$, $\lambda_{SO} = 0.1t$, and $\lambda_R = 0$. In this case, the bulk bands are gapped with the magnitude $2t$ at two valleys at $k_x = 0$ and π/a. Notice that now all the four edge states (L and U, ↑ and ↓, respectively) survive after the appearance of bulk states when $|E_F| > t$, as can be confirmed by their real-space distributions in Figs. 2(e)–2(h). Another distinct feature of model II is that both spin components coexist in the same valley, while for model I, the spins are opposite for two valleys [18,19].

Equipped with the basic background of models, we will investigate scattering from impurities. Nonmagnetic impurities are expressed by adding a term to the Hamiltonian as

$$H_I = \sum_{i,\sigma} V_i c^\dagger_{i\sigma} c_{i\sigma}, \tag{2}$$

where V_i is the random onsite potential, and is independent of spin to preserve TRS. Impurities induce scatterings between states at the Fermi energy. We investigate this in a standard geometry of quasi-1D ribbon infinitely extending in the x direction, which is divided into three parts: the left lead (source), the central region (sample), and the right lead (drain). The leads are clean and semi-infinite, with well-defined channels indexed by n at Fermi energy E_F. The width N_y of the ribbon is chosen to be sufficiently large to avoid finite-size effects [20]. With impurities in the sample, the channel-resolved transmission amplitudes $T_{mn} \equiv T_{m\leftarrow n}$ between right-going channels, i.e., from channel n in the left lead to channel m in the right lead, can be calculated by using the numerical Green's function methods introduced in Refs. [22,23]. Thus, the transport capability of channel n from the left lead can be defined as its transmission through the sample over all possible channels m in the right lead $T_n = \sum_m T_{mn}$. The total transmission $\sum_{mn} T_{mn}$ over all active channels is the conductance in units of e^2/h at zero temperature, with m and n running over all right-going channels [24]. The model parameters (except for disorder) are chosen to be identical for the sample and lead to reveal the intrinsic scattering behavior from disorder at this set of parameters.

In order to investigate the localization of edge (bulk) states, respectively, we define the edge (bulk) transmission as

$$T_{\text{bulk}} = \sum_m \sum_{n\in\text{bulk}} T_{mn}, \tag{3}$$

$$T_{\text{edge}} = \sum_m \sum_{n\in\text{edge}} T_{mn}, \tag{4}$$

where n only runs over *edge* (*bulk*) right-going channels in the left lead, and m runs over all right-going channels in the right lead. This T_{edge} (T_{bulk}) reflects the final fate of incident edge (bulk) states after going through the sample. Due to disorder, possible localization makes T's badly distributed over different disorder realizations. Thus, it is essential to investigate the geometric mean $T^{\text{typ}} = \exp\langle \ln T \rangle$ to give the "typical value" instead of the arithmetic mean $T^{\text{ave}} = \langle T \rangle$ [25].

III. ANDERSON DISORDER

In the presence of bulk-edge coexistence, a natural conjecture is that disorder will localize the bulk states and remain the edge states left robust. We will show that the case is not so simple. Let us start from Anderson disorder, where the random onsite potential V_i is independently and uniformly distributed between $(-W/2, W/2)$.

If the Fermi energy is in the bulk gap, as a typical 2D TI, it is well known that the electronic transport should be robust even with disorder. Indeed, as shown in the inset of Fig. 3(a) for model I, the transmission as a function of sample length N_x at $E_F = 0.05t$ is a robust plateau 2, carried by right-going edge states L↓ and U↑. On the other hand, in the presence of coexisting bulk states at $E_F = 0.55t$, as discussed above, there is only one right-going edge channel (L↓) left. Now, as shown in the main panel of Fig. 3(a), the edge and bulk transmissions are both decaying with transport length N_x, with almost the same decay rate.

This nonrobustness can also be seen from the transmission dependence on disorder strength W, as shown in Fig. 3(b):

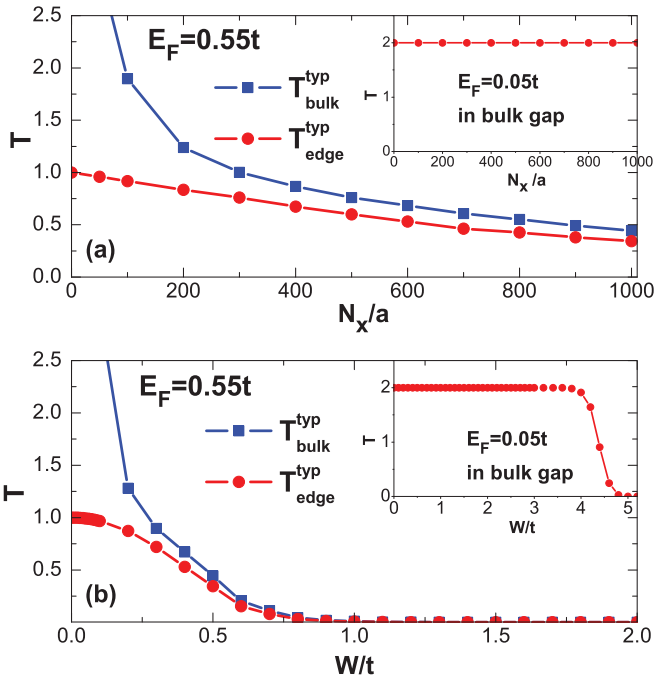

FIG. 3. (Color online) Model I. Typical values of bulk (blue square) and edge (red circle) transmissions with Anderson disorder. Main panels are for $E_F = 0.55t$ (in band) while insets are for $E_F = 0.05t$ (in bulk gap). (a) Transmissions as functions of sample length N_x with $W = 0.5t$. (b) Transmissions as functions of disorder strength W with $N_x = 1000$. Other model parameters are the same as in Fig. 1(a). Each data point is averaged over 1000 samples.

the in-band edge and bulk transmissions ($E_F = 0.55t$, main panel) collapse to zero almost at the same time at $W \sim 0.8t$, which are much earlier than the robust in-gap ($E_F = 0.05t$, inset; notice different scales of horizontal axes between the main and the inset panels) edge transmission at $W \sim 5t$. In this model, although the backscattering from the edge state L↓ (right-going) to the edge state L↑ (left-going) is still prohibited by TRS, the edge state L↓ still decays by leaking into the bulk and being reflected to left-going bulk channels.

Similar localization happens to model II, with all four edge states surviving after going into the bulk, as plotted in Fig. 4. In other words, although there exist pseudo-edge states [A and D in Fig. 1(a)] or even real edge states [B, C, E, F, G, and H in Figs. 1(a) and 1(b)], disorder does not necessarily make the bulk states localized while keeping the edge states extended. On the contrary, they are localized indiscriminately. This is the first important finding in this work. With the coexistence, the robustness of edge states has been destroyed by their leakage and backscattering into the bulk, which make themselves easily localized, as conventional electronic states in disordered 2D systems. This is consistent with previous analytical results that the coupling with bulk states will endow the edge states with a finite lifetime $\tau \sim \frac{1}{|g|^2 \nu(E_F)}$, where g is the coupling and ν is the bulk density of states [12].

IV. LONG-RANGE DISORDER

The above results show the disastrous consequence of edge-bulk scattering in disordered 2D TI by destroying the

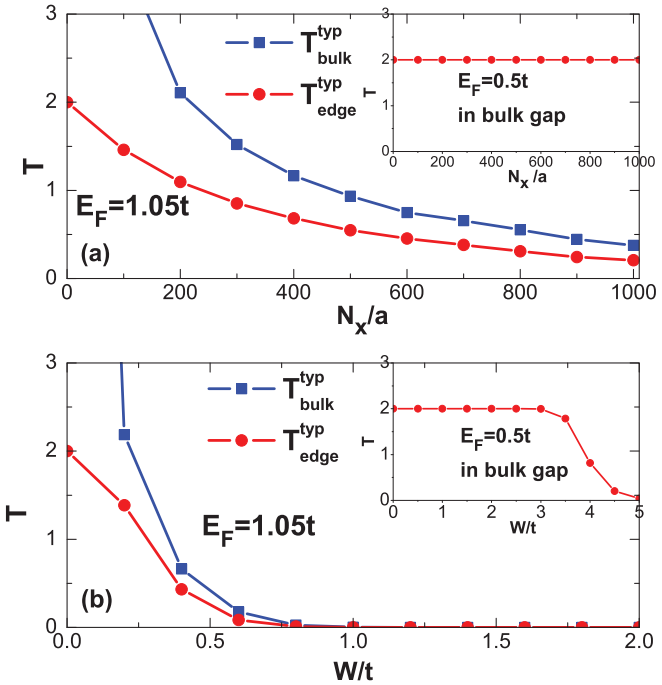

FIG. 4. (Color online) Similar with Fig. 3 but for model II. Main panels are for $E_F = 1.05t$ (in bulk band) while insets are for $E_F = 0.5t$ (in bulk gap). Others are the same with Fig. 3.

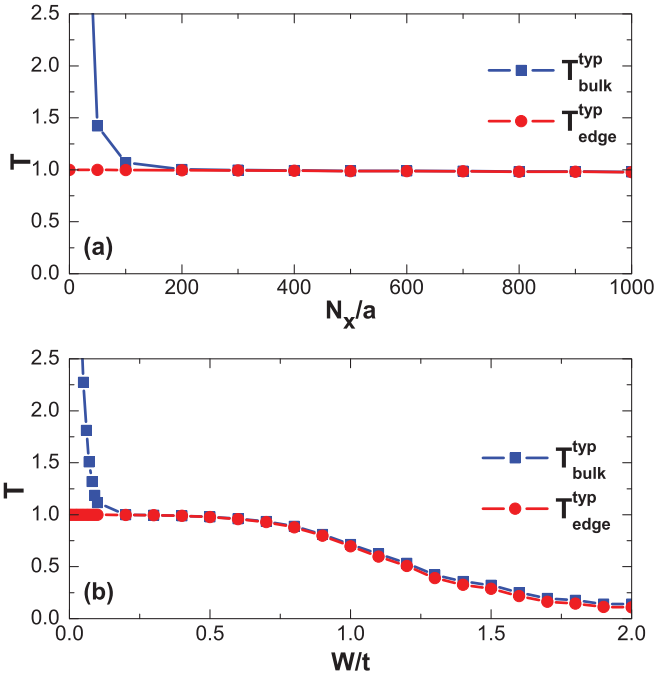

FIG. 5. (Color online) Model I with long-range impurities $\xi = 1.5a$, $n_I = 2.5\%$. (a) Transmissions as functions of sample length N_x with $W = 0.5t$. (b) Transmissions as functions of disorder strength W with $N_x = 1000a$. Others are the same with main panels in Fig. 3.

robustness of edge states remarkably. We are now in a position to look for means to avoid this. As can be seen in Figs. 1(a) and 1(b), the bulk and edge states are well separated in k space when they coexist. To effectively decouple them, it is useful to consider the long-range disorder, which suppresses scattering processes with large-momentum transfer. In the presence of long-range disorder, the onsite potential V_i is the sum of contributions from N_I impurities randomly centered at $\{\mathbf{r}_m\}$ among N sites $V_i = \sum_{m=1}^{N_I} U_m \exp[-|\mathbf{r}_i - \mathbf{r}_m|^2/(2\xi)]$, where U_m is randomly distributed within $(-W/2, W/2)$. The above Anderson disorder corresponds to short-range impurities with correlation length $\xi = +0$ and density $n_i \equiv N_I/N = 100\%$. Impurities with correlation length $\xi > a$ can be said to be long range, as widely investigated in graphene with substrate [26–28].

The transmissions for model I with long-range impurities ($\xi = 1.5a$ and $n_i = 2.5\%$) are plotted in Fig. 5. In Fig. 5(a), interestingly, the edge transmission (red circle) almost does not decay with length at all, restoring its robustness by effective decoupling from bulk states. This robustness of the edge channel can also be seen from its dependence on disorder strength W in Fig. 5(b). On the other hand, the bulk transmission (blue square) decays very fast before $N_x < 200$, reflecting the Anderson localization nature of the 2D bulk states. But, after that this bulk decay also ceases at another perfectly conducting channel with unit transmission. This residual PCC can be attributed to the imbalance between the number of left- and right-moving bulk channels [29], which is explained as follows.

The information of all channels at $E_F = 0.55t$ is listed in Table I. The long-range nature of disorder effectively decouples the edge channels (No. 8 for spin ↑ and No. 9 for spin ↓) from the bulk ones, which results in an imbalance between opposite channels in the bulk. For the bulk spin-↑ component, there are four right-going channels (Nos. 13 to 16) but three left-going channels (Nos. 10 to 12). While for the bulk spin-↓ component, there are four left-going channels (Nos. 1 to 4) but three right-going channels (Nos. 5 to 7). In brief, there is one more right- (left-) going bulk channel for spin-up (-down) component, and there is no scattering between states with opposite spins due to TRS. Thus, for a right-going setup here, the residual right-going spin-up channel (in the spin-↑ component) exhibits itself by the appearance of a PCC. This picture can be confirmed by the channel-resolved transmissions for a typical sample listed in Table II. Of the total transmission $\sum_n T_n = 2$, 1 is from the edge channel (No. 9) with spin ↓, and the rest 1 is mostly from the sum of four bulk

TABLE I. Model I. Information for 16 active channels at $E_F = 0.55t$ as plotted in Fig. 1(a): the directions of their group velocities [+ (−) for right (left) going] and the orientations of their spins. The numbers are sorted by ascending k_x. States A, B, C, and D in Fig. 1(a) correspond to channels No. 1, 8, 9, and 16, respectively.

n	1	2	3	4	5	6	7	8
Velocity	−	−	−	−	+	+	+	−
Spin	↓	↓	↓	↓	↓	↓	↓	↑
Edge								✓
No.	9	10	11	12	13	14	15	16
Velocity	+	−	−	−	+	+	+	+
Spin	↓	↑	↑	↑	↑	↑	↑	↑
Edge	✓							

TABLE II. Model I. Channel transmissions of the eight right-going channels, for a typical sample with $N_x = 1000$ used in Fig. 5, with long-range impurities at $W = 0.5t$. The channel numbers are same as in Table I.

n	5	6	7	9	13	14	15	16	Sum
T_n	0.000	0.000	0.000	0.999	0.035	0.123	0.590	0.252	1.999
Spin	↓	↓	↓	↓	↑	↑	↑	↑	
Edge				✓					

channels (Nos. 13, 14, 15, and 16) with spin ↑. Notice this PCC in the bulk is a collective effect, and it does *not* correspond to any *one* specific eigenchannel listed in Table I. For example, even the transmission of the pseudo-edge state, channel No. 16, is much less than 1, as other bulk channels. Similar PCCs in other quasi-1D systems with odd number of channels have been discussed in Refs. [28,30]. To sum up, for model I, the long-range disorder effectively decouples edge states from bulk, resulting in two perfectly conducting channels, one from the edge and one from the bulk. This originates from its special configuration of edge and bulk states: only one pair of edge states coexisting with bulk states.

Let us check the case of model II, with both pairs of edge states surviving with bulk states. At Fermi energy $E_F = 1.05$, there are totally 40 active channels (see Table III), much more than in model I. Now, for each spin component (↑ or ↓) of model II plotted in Fig. 1(b) there are 18 bulk channels and 2 edge channels, both with exactly half left-going and half right-going members, as listed in Table III. After the effective decoupling from long-range disorder, as shown in Fig. 6(a), the bulk states (now with even number of channels) are completely localized to zero transmission after transport length $N_x > 500$, even though there are much more bulk channels than in model I. Meanwhile, two right-going edge channels (L↓ and U↑), respectively, restore their robust transports, giving rise to the total transmission 2. The transmissions of a typical sample are listed in Table IV. Therefore, in the case of model II, after the edge and bulk states have been effectively decoupled, there is no "odd-channel" problem for the latter, which happened in model I. Thus, the situation is simpler and clearer: the bulk states are easily localized as in conventional 2D systems, while edge states regain their robustness.

TABLE III. Similar with Table I, but for model II at $E_F = 1.05t$. States E, F, G, and H in Fig. 1(b) correspond to channels No. 7, 8, 33, and 34, respectively.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Velocity	−	+	+	+	+	+	−	−	−	−	−	−	−	−
Spin	↓	↑	↑	↓	↓	↓	↓	↑	↓	↓	↓	↓	↓	↑
Edge							✓	✓						
n	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Velocity	−	−	−	−	−	−	+	+	+	+	+	+	+	+
Spin	↑	↓	↑	↑	↑	↑	↓	↓	↓	↓	↑	↓	↓	↑
Edge														
n	29	30	31	32	33	34	35	36	37	38	39	40		
Velocity	+	+	+	+	+	+	−	−	−	−	−	+		
Spin	↑	↑	↑	↑	↓	↑	↑	↑	↑	↓	↓	↑		
Edge					✓	✓								

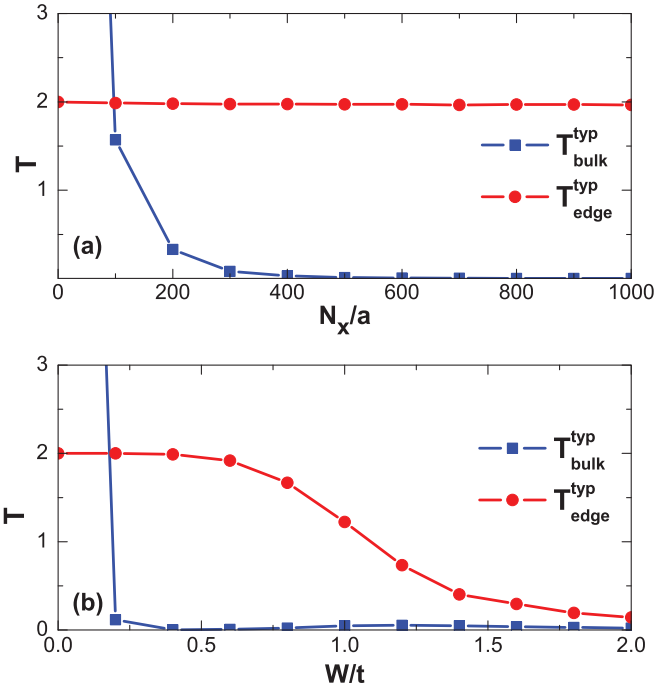

FIG. 6. (Color online) Model II with long-range impurities $\xi = 4a$, $n_I = 0.25\%$. Main panels are for $E_F = 1.05t$ (in band) while insets are for $E_F = 0.5t$ (in bulk gap). (a) Transmissions as functions of sample length N_x with $W = 0.5t$. (b) Transmissions as functions of disorder strength W with $N_x = 1000a$. Others are the same with Fig. 1(b).

V. EDGE GATING

The decoupling of edge and bulk states effectively recovers the robust conducting behavior of topological states. The above decoupling from long-range disorder is based on the k-space consideration. However, the correlation length ξ and density n_I of impurities depend strongly on the details of materials, and are hard to control. Now, we consider a real-space proposal. For model I at the intrinsic Fermi energy $E_F = 0.55t$ in Fig. 1(a), a simple way to discard the redundant bulk states is to gate the whole 2D sample into the bulk gap, by a top gate covering the whole sample sheet. However, we will show that an edge gating is also sufficient to create a decoupling between edge and bulk states and therefore results in PCCs. This seems unnecessary for a 2D system as discussed above, but is intuitive for a generalization to 3D TIs because a gate surrounding a 3D

TABLE IV. Transmissions for model II, for a typical sample with $N_x = 1000$ used in Fig. 6, with long-range impurities at $W = 0.5t$. The channel numbers are the same as in Table III.

n	33	34	Bulk	Sum
T_n	1.000	0.997	0.003	2.000
Spin	↓	↑	−	
Edge	✓	✓		

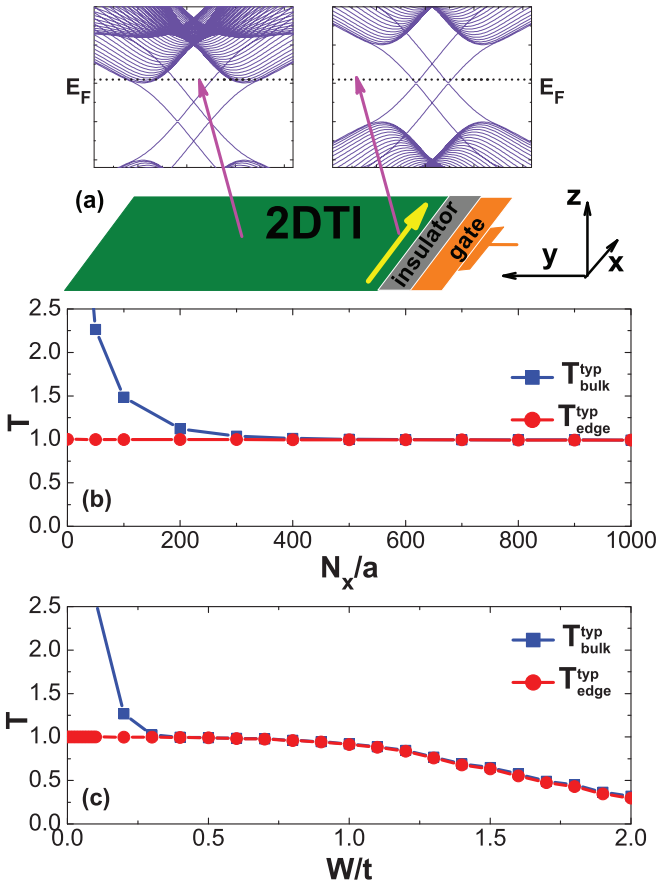

FIG. 7. (Color online) Model I with edge gating. (a) Illustration of the metal edge gate (orange)/insulator (gray) structure attached around one edge of the 2D TI (green), which induces a local potential so that the Fermi energy around the edge meets the bulk gap. The electronic transport is along the x direction. The yellow arrow represents the edge state. (b) and (c) are all the same with main panels of Figs. 3(a) and 3(b), but with an edge gating $V_G = -0.5t$ and $\xi_G = 3a$.

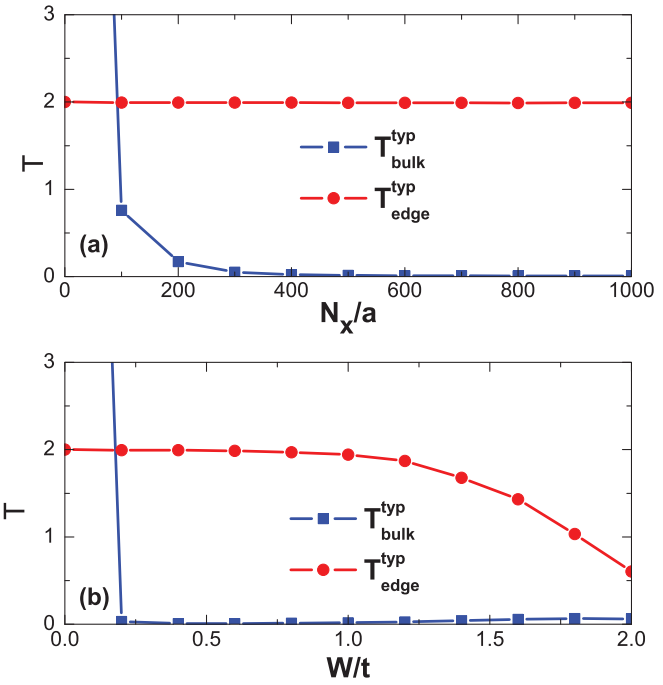

FIG. 8. (Color online) Model II with both edge gated. Same with main panels of Fig. 4, but with both edges gated as $E_G = 0.8t$ and $\xi_G = 4a$.

sample can only affect local carrier densities (therefore local chemical potentials) near the interface to the gate, instead of those in the whole bulk.

As illustrated in Fig. 7(a), consider that a local gate structure [31,32] is attached around the lower edge ($y = 0$) of the sample of model I, where the edge and bulk states coexist. This edge-gate voltage can only change the sample's local potential V around this edge, say, in an exponential decaying form as

$$V(x,y) = V_G \exp(-y/\xi_G), \tag{5}$$

where ξ_G is the decay length. If V_G can be tuned so that the Fermi energy meets the bulk gap around the edge, and ξ_G is larger than the penetration depth of the edge state in transverse direction, then there will be only edge state [the yellow arrow in Fig. 7(b)] in this gated region ($0 \leq y \lesssim \xi_G$) near this edge, in spite of the presence of bulk states outside it. Although, at different potentials the gated and ungated regions still belong to the same topological phase, with no additional edge states at the interface ($y \sim \xi_G$) between them. So far, edge and bulk states have been well separated in real space. This decoupling should also restore the robustness of transport even for Anderson disorder. This is confirmed from the transmissions shown in Fig. 7(c), where perfectly conducting edge and bulk channels can be seen. The transmissions converge to unity rapidly when E_G goes into the bulk gap and when ξ_G is just several lattice constants long. We also show the transmissions at fixed length $N_x = 1000a$ when varying disorder strength W. As other proposals of edge tuning in 2D TI [33–36], we have made this decoupling by taking advantage of the local nature of edge states.

For model II, with two pairs of edge states at both edges, edge gating on *both* edges is necessary. In Fig. 8, we plot the transmissions for model II with both edge gated. Similar to the results of long-range disorder shown in Fig. 6, this decoupling of edge states from bulk also restores the robust transport $T_{\text{edge}} = 2$ of two right-going channels, while it makes the bulk easily localized, i.e., $T_{\text{bulk}} \sim 0$.

Therefore, an edge gate can drive a disordered TI with bulk-edge coexistence from Anderson insulator [Figs. 3(a) and 4(a)] to a perfectly conductor [Figs. 7(b) and 8(a)]. This transition can be experimentally tuned to make a topological FET [14,15]. The on-off conductance ratio $G_{\text{on}}/G_{\text{off}} \sim \exp[-N_x/\lambda_G]/\exp[-N_x/\lambda_0]$, where λ_0 and λ_G are the localization length of zero gating and edge gating, respectively, as long as N_x is within the phase coherence length. Moreover, the ON state is very stable due to its topological origin. Since this tuning only requires an electric control near the boundary, it can be naturally generalized to 3D: A surface gating surrounding a 3D TI can decouple surface and bulk states coexisting at the same Fermi energy, so that the robust transport of surface states is restored while the bulk states can be localized by impurities. Recently, a 3D TI based FET with

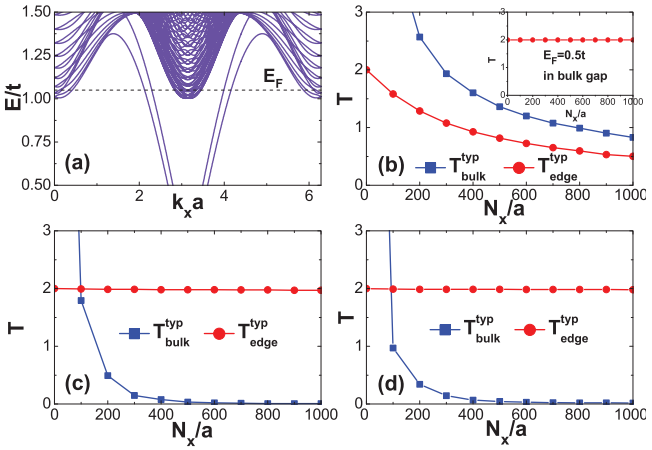

FIG. 9. (Color online) Model II with Rashba SOC $\lambda_R = 0.4t$. (a) The band structure of a zigzag nanoribbon. Transmissions at $E_F = 1.05t$ as functions of N_x for Anderson disorder (b), long-range disorder (c), and edge gating (d). Other parameters are the same with Figs. 1(b), 4(a), 6(a), and 8(a), respectively. The inset of (b) is at $E_F = 0.5t$ in the bulk gap.

boundary gating has been experimentally observed in [37], and our present work offers a possible theoretical explanation.

VI. FINITE RASHBA SOC

So far, we have neglected the Rashba SOC term, which makes the spin s_z a conserving quantity and decouples the system into two spin components $\uparrow$ and $\downarrow$. In this case, the quantum spin Hall system with $Z_2 = 1$ can be viewed as two Chern insulating systems with opposite Chern numbers, one for each spin component [1,18,19]. Therefore, our conclusions so far are also valid for Chern insulators. Finite Rashba SOC does not break time-reversal symmetry but breaks the conservation of s_z, making two spin components $\uparrow$ and $\downarrow$ coupled and the spin orientation process with momentum. Nevertheless, the system remains in the quantum spin Hall phase with $Z_2 = 1$ and with robust edge states, as long as the Rashba SOC does not close the bulk gap.

To check the validity of our conclusions for the case of finite Rasha SOC, in Fig. 9(a), we plot the nanoribbon band structure for model II with $\lambda_R = 0.4t$, where two spin components $\uparrow$ and $\downarrow$ are coupled remarkably. Despite a different spin configuration, similar to the version without Rashba SOC plotted in Fig. 1(b), two pairs of edge states survive with the bulk band when $|E_F| > t$. In the bulk gap $|E_F| < t$, also as a typical 2D TI system with $Z_2 = 1$, the edge transport is robust, as shown in the inset of Fig. 9(b). At Fermi energy $E_F = 1.05t$ with the coexistence of edge and bulk states, we repeat the calculations of transmissions (as functions of length N_x) for Anderson disorder, long-range disorder, and edge gating, and the results are shown in Figs. 9(b), 9(c), and 9(d), respectively. Interestingly, the results almost duplicate what has happened in the case of $\lambda_R = 0$ [compare with Figs. 4(a), 6(a), and 8(a), respectively]. It is now clear that all the above conclusions apply to the case of finite Rashba SOC, as long as the coexistence of edge and bulk states.

VII. PHASE DIAGRAM

In order to have a global view, in Fig. 10, we plot the total transmission of model I for a specific sample on the $W - E_F$ plane. For Anderson disorder [Fig. 10(a)], there is a clear transmission plateau $T = 2$ (green) in the bulk-gap region with $|E_F| < 0.5t$, originating from two pairs of edge states. Apart from this region, although there are still well-defined edge states coexisting with bulk states, the scattering between them makes the transmission decay rapidly to zero (blue), as discussed in Sec. III. In the presence of long-range disorder [Fig. 10(b)], on the other hand, the transmission plateau $T = 2$ extends itself into the bulk-band region, as indicated by the region enclosed by black dashed contours. Notice in these regions the transmission was $T \sim 0$ in Fig. 10(a). This reflects the restoring of robust transport from the decoupling of edge and bulk states. These two phase diagrams possess particle-hole symmetry, which originates from that of the Hamiltonian (1) and from the zero mean values of disorder potentials.

As for the case of edge gating, as illustrated in Fig. 7(b), an edge gate with E_G will push the local chemical potential at the edge to $E_F - E_G$. Therefore, with an edge gating $E_G = 0.5t$

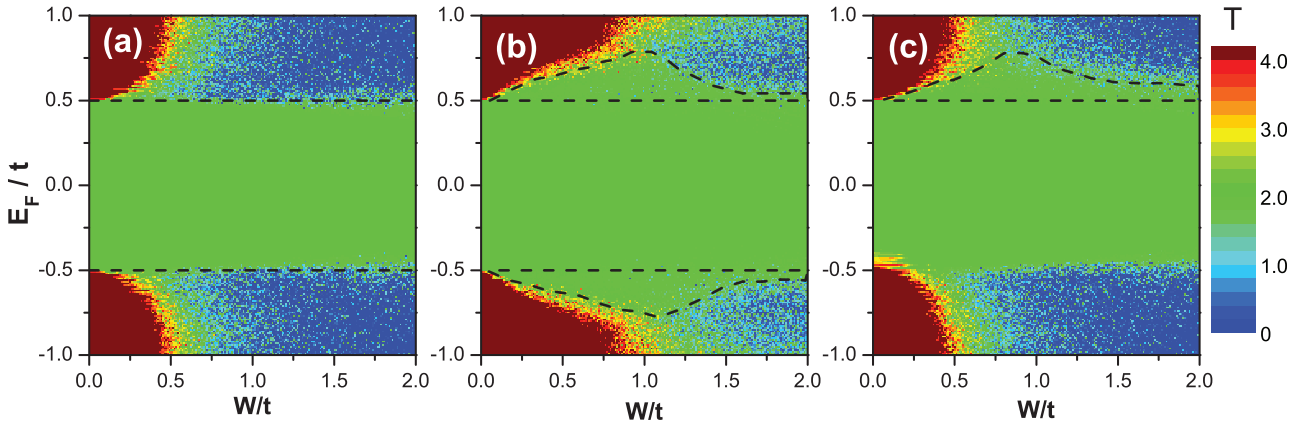

FIG. 10. (Color online) Total transmission as a function of W and E_F for model I, of a sample with the size $N_x \times N_y = 500 \times 120$. (a) Anderson disorder the same with Fig. 3. (b) Long-range disorder the same with Fig. 5. (c) Edge gating the same with Fig. 7.

[Fig. 10(c)], it can push the edge potential back to the bulk gap if Fermi energy is in the conduction band ($E_F > 0.5t$). However, if Fermi energy is in the valence band ($E_F < -0.5t$), this same edge gate will push the edge potential deeper into the valence band, so it cannot decouple edge and bulk states. This is reflected in Fig. 10(c) that the robust transmission plateau (green) extends itself only into the conduction band while remains the same in the valence band as in the ungated case plotted in Fig. 10(a).

Let us discuss a few words about the size and energy dependencies. The results in Fig. 10 are calculated from samples with width $N_y = 120$, instead of $N_y = 80$ in all the previous calculations in this paper. This means that the above physical pictures are independent of the sample size as long as it exceeds the threshold of finite-size effect [20]. In Figs. 10(b) and 10(c), with Fermi energy deeper into the bulk band, the transmission plateau (green) is narrower in the W direction, or, states are more easily to be localized. This is also consistent with the argument from the effective theory that, with fixed coupling (whether large or small) from the bulk, the lifetime of edge states is inversely proportional to the density of states of the bulk [12].

Before closing, it is interesting to compare our results with those related works from other 2D TI models. Another popular model of 2D TI is the Bernevig-Hughes-Zhang (BHZ) model, describing quantum spin Hall effects in HgTe/CdTe quantum wells [2]. In previous numerical experiments, starting from a topological nontrivial phase at clean limit, there may also develop a transmission plateau $T = 2$ with the increasing of Anderson disorder strength, at some Fermi energy in the bulk band [38–40]. This is one type of so-called topological Anderson insulator (TAI). We reproduce the transmission diagram of the BHZ model in Fig. 11(b). At $E_F = 0.02$ eV (black lines in Fig. 11) in the bulk band, the transmission plateau $T = 2$ appears after $W > 0.1$ eV. In the following, we will demonstrate that this has different physics from our above results summarized in Fig. 10.

First, now there are no "real" edge states for the BHZ model at this Fermi energy in the clean limit. As can be seen in Fig. 11(a) and insets, when $E > 0.01$ eV, the edge states from the bulk gap merge into the continuum of the bulk band and lose their edge nature, like states A and D in Fig. 1(a). This is why we did not choose the BHZ model in this work. In a most related work [12], in order to construct the coexistence of bulk and edge states, same as we defined above, one more orbital was deliberately added to the Haldane (spinless version of the KM model) model. We found that appropriate parameters of the Kane-Mele model itself can also attain the same purpose, as plotted in Figs. 1(a) and 1(b).

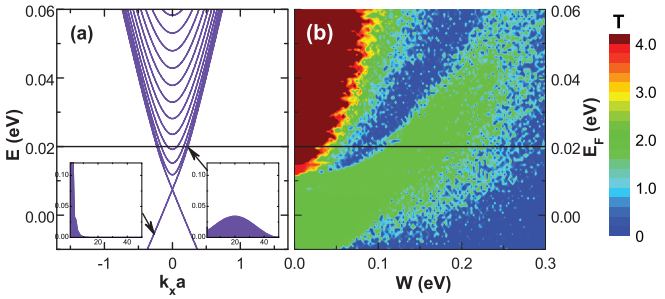

FIG. 11. (Color online) Bernevig-Hughes-Zhang model. (a) The band structure of a $N_y = 50$ ribbon. Insets are real-space distributions of wave functions at indicated states. (b) Total transmission as a function of W and E_F of a $N_x \times N_y = 500 \times 50$ sample, with Anderson disorder. The model parameters are the same with Fig. 2(c) in Ref. [38].

Second, as can be seen from the $W - E_F$ phase diagram in Fig. 11(b), starting from a topologically nontrivial phase at $W = 0$, this type of TAI originates from the asymmetrical renormalizations to two band edges from disorder [41,42]. In other words, with increasing W, the transmission plateau 2 is growing [from the bulk-gap region $(-0.02, 0.02)$ eV at $W = 0$] upward towards positive energy direction, because the model itself lacks particle-hole symmetry. Therefore, at some appropriate E_F in the bulk conduction band, one will meet this transmission plateau with increasing W. This seemingly "disorder-induced transmission plateau" at some definite Fermi energy will not happen in our case of particle-hole-symmetric system with Anderson disorder, shown as in Fig. 10(a). Thus, in the case of the BHZ model, two methods in our paper (long-range disorder and edge gating) to decouple bulk and edge states are not applicable any more since there are no edge states for one to decouple at all. As a result, this model's responses to disorders are different from the two models we considered above. For example, it was found that long-range disorder gives rise to a shrink of the $T = 2$ transmission plateau [43], instead of an expansion as shown in Fig. 10(b).

In brief, at a definite Fermi energy, TAI is to induce a topological edge transport by disorder, when there were no edge states at clean limit; on the other hand, our work is to restore the robustness of edge states, when they have been there at the clean limit but were not robust due to backscattering from the coexisting bulk.

VIII. CONCLUSIONS

As a summary, we studied the edge and bulk transmissions of 2D TI when they coexist at the same Fermi energy. Anderson disorder tends to localize edge and bulk states indiscriminately because of unprotected backscattering between them. In order to decouple them and restore the robust transport of edge states, we introduced long-range disorder and edge gating, based on momentum-space and real-space considerations, respectively. After effective decoupling of bulk and edge states, disorder can easily localize the bulk states, while keeping the robust transport of edge states, if the bulk has even number of transport channels. By edge gating, one can electrically tune the disordered system between an Anderson localized state and a robust extended state, making it a topological FET.

ACKNOWLEDGMENT

This work was supported by NSFC (Grants No. 11374294, No. 11204294, and No. 11274364) and 973 Program Project No. 2013CB933304.

[1] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[2] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
[3] S.-Q. Shen, *Topological Insulators: Dirac Equation in Condensed Matters* (Springer, Berlin, 2012).
[4] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X. L. Qi, and S. C. Zhang, Science **318**, 766 (2007).
[5] A. Roth, C. Brüne, H. Buhmann, L. W. Molenkamp, J. Maciejko, X.-L. Qi, and S.-C. Zhang, Science **325**, 294 (2009).
[6] C.-Z. Chang, J.-S. Zhang, X. Feng, J. Shen, Z.-C. Zhang, M.-H. Guo, K. Li, Y.-B. Ou, P. Wei, L.-L. Wang, Z.-Q. Ji, Y. Feng, S.-H. Ji, X. Chen, J.-F. Jia, X. Dai, Z. Fang, S.-C. Zhang, K. He, Y.-Y. Wang, L. Lu, X.-C. Ma, and Q.-K. Xue, Science **340**, 167 (2013).
[7] L.-J. Du, I. Knez, G. Sullivan, and R.-R. Du, arXiv:1306.1925.
[8] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nat. Phys. **5**, 398 (2009).
[9] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008).
[10] Y.-Y. Li, G. Wang, X.-G. Zhu, M.-H. Liu, C. Ye, X. Chen, Y.-Y. Wang, K. He, L.-L. Wang, X.-C. Ma, H.-J. Zhang, X. Dai, Z. Fang, X.-C. Xie, Y. Liu, X.-L. Qi, J.-F. Jia, S.-C. Zhang, and Q.-K. Xue, Adv. Mater. **22**, 4002 (2010).
[11] S. Kim, M. Ye, K. Kuroda, Y. Yamada, E. E. Krasovskii, E. V. Chulkov, K. Miyamoto, M. Nakatake, T. Okuda, Y. Ueda, K. Shimada, H. Namatame, M. Taniguchi, and A. Kimura, Phys. Rev. Lett. **107**, 056803 (2011).
[12] D. L. Bergman and G. Refael, Phys. Rev. B **82**, 195417 (2010).
[13] Y.-T. Hsu, M. H. Fischer, T. L. Hughes, K. Park, and E.-A. Kim, Phys. Rev. B **89**, 205438 (2014).
[14] Q.-K. Xue, Nat. Nanotechnol. **6**, 197 (2011).
[15] L. Andrew Wray, Nat. Phys. **8**, 705 (2012).
[16] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[17] C.-C. Liu, H. Jiang, and Y. Yao, Phys. Rev. B **84**, 195430 (2011).
[18] M. Ezawa, New J. Phys. **14**, 033003 (2012).
[19] M. Ezawa, Phys. Rev. Lett. **109**, 055502 (2012).
[20] B. Zhou, H. Z. Lu, R. L. Chu, S. Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).
[21] M. Pang and X.-G. Wu, Chin. Phys. B **23**, 077302 (2014).
[22] T. Ando, Phys. Rev. B **44**, 8017 (1991).
[23] P. A. Khomyakov, G. Brocks, V. Karpan, M. Zwierzycki, and P. J. Kelly, Phys. Rev. B **72**, 035450 (2005).
[24] Y. Imry and R. Landauer, Rev. Mod. Phys. **71**, S306 (1999).
[25] K. Slevin, P. Markoš, and T. Ohtsuki, Phys. Rev. Lett. **86**, 3594 (2001).
[26] T. Ando and Nakanishi, J. Phys. Soc. Jpn. **67**, 1704 (1998).
[27] A. Rycerz, J. Tworzydło, and C. W. J. Beenakker, Europhys. Lett. **79**, 57003 (2007).
[28] K. Wakabayashi, Y. Takane, and M. Sigrist, Phys. Rev. Lett. **99**, 036601 (2007).
[29] S. D. Wang, Z.-Z. Sun, G. Xiong, S. Yin, and X. R. Wang, J. Phys. A: Math. Gen. **37**, 1337 (2004).
[30] Y. Takane, S. Iwasaki, Y. Yoshioka, M. Yamamoto, and K. Wakabayashi, J. Phys. Soc. Jpn. **78**, 034717 (2009).
[31] A. F. Young and Philip Kim, Nat. Phys. **5**, 222 (2009).
[32] N. Stander, B. Huard, and D. Goldhaber-Gordon, Phys. Rev. Lett. **102**, 026807 (2009).
[33] Y.-T. Zhang, Q.-F. Sun, and X. C. Xie, J. Appl. Phys. **109**, 123718 (2011).
[34] X.-T. An, Y.-Y. Zhang, J.-J. Liu, and S.-S. Li, New J. Phys. **14**, 083039 (2012).
[35] H.-C. Li, L. Sheng, and D. Y. Xing, Phys. Rev. Lett. **108**, 196806 (2012).
[36] H.-C. Li, L. Sheng, R. Shen, L. B. Shao, B.-G. Wang, D. N. Sheng, and D. Y. Xing, Phys. Rev. Lett. **110**, 266802 (2013).
[37] H. Zhu, C. A. Richter, E. Zhao, J. E. Bonevich, W. A. Kimes, H.-J. Jang, H. Yuan, H. Li, A. Arab, O. Kirillov, J. E. Maslar, D. E. Ioannou, and Q. Li, Sci. Rep. **3**, 1757 (2013).
[38] J. Li, R.-L. Chu, J. K. Jain, and S.-Q. Shen, Phys. Rev. Lett. **102**, 136806 (2009).
[39] H. Jiang, L. Wang, Q.-F. Sun, and X. C. Xie, Phys. Rev. B **80**, 165316 (2009).
[40] Y.-X. Xing, L. Zhang, and J. Wang, Phys. Rev. B **84**, 035110 (2011).
[41] C. W. Groth, M. Wimmer, A. R. Akhmerov, J. Tworzydło, and C. W. J. Beenakker, Phys. Rev. Lett. **103**, 196805 (2009).
[42] Y.-Y. Zhang and S.-Q. Shen, Phys. Rev. B **88**, 195145 (2013).
[43] A. Girschik, F. Libisch, and S. Rotter, Phys. Rev. B **88**, 014201 (2013).