Equivalent expression of $\mathbb{Z}_2$ topological invariant for band insulators using the non-Abelian Berry connection

Rui Yu,[1] Xiao Liang Qi,[2] Andrei Bernevig,[3] Zhong Fang,[1] and Xi Dai[1]

[1]*Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100080, China*

[2]*Department of Physics, Stanford University, Stanford, California 94305, USA*

[3]*Department of Physics, Princeton University, Princeton, New Jersey 08540, USA*

(Received 12 January 2011; revised manuscript received 2 June 2011; published 8 August 2011)

We introduce an expression for the $\mathbb{Z}_2$ topological invariant of band insulators using the non-Abelian Berry connection. Our expression can identify the topological nature of a general band insulator *without* any of the gauge-fixing problems that plague the concrete implementation of previous invariants. This expression can be derived from the "partner switching" of the Wannier function center during time-reversal pumping and is thus equivalent to the Z_2 topological invariant proposed by Kane and Mele. Using our expression, we have recalculated the Z_2 topological index for several topological insulator material systems and obtained consistent results with the previous studies.

DOI: 10.1103/PhysRevB.84.075119 PACS number(s): 71.10.Pm, 71.15.Mb

I. INTRODUCTION

Topological invariants play a very important role in the classification of band insulators. The studies on integer quantum Hall effect (IQHE) show that 2D band insulators without time-reversal symmetry can be classified by the Chern number—an integer describing the topological structure of a set of fully occupied Bloch bands without Kramers degeneracy. Systems with nonzero Chern numbers exhibit IQHE.[1,2]

A similar idea can also be applied to band insulators with time-reversal symmetry. A $\mathbb{Z}_2$ topological invariant has been proposed by Kane and Mele to characterize the time-reversal invariant band insulators in two dimensions.[3,4] According to this new topological invariant, all the two-dimensional (2D) band insulators with time-reversal invariance can be divided into two classes. The normal insulators with even $\mathbb{Z}_2$ numbers and topological insulators with odd $\mathbb{Z}_2$ numbers.[5–7] The 2D topological insulators will exhibit a quantum spin Hall effect (QSHE),[3,8] which is characterized by the presence of helical edge states.[3,8–14] Interestingly, the $\mathbb{Z}_2$ topological invariant can also be generalized to the three-dimensional (3D) band insulators with time-reversal symmetry.[5,6,15] In this case, there are four independent $\mathbb{Z}_2$ topological numbers: one strong topological index and three weak topological indices.[15–19] The 3D time-reversal invariant band insulators can be classified as normal insulators, weak topological insulators (WTI), and strong topological insulators (STI) according to the values of these four $\mathbb{Z}_2$ topological indices. Among them, STI attract much attention due to its unique Dirac-type surface states and robustness against disorder.[20–31] The helical spin structure of the Dirac-type surface states has been experimentally verified by the standing-wave structure in scanning tunneling microscope (STM) images around an impurity scattering center and measured directly by spin-resolved angle-resolved photoemission spectra (ARPES).[32–40] The Dirac-type 2D electron gas living on the surface of STI or at the interface between STI and normal insulators provides a new playground for spintronics and quantum computing.

Since the $\mathbb{Z}_2$ invariant characterizes whether a system is topologically trivial or nontrivial, its computation is essential to the field of topological insulators. For band insulators with extra spacial inversion symmetry, the $\mathbb{Z}_2$ topological numbers can be easily computed as the product of half of the parity (Kramers pairs have identical parities) numbers for all the occupied states at the high symmetry points.[16] The situation becomes complicated in the general case where spacial inversion symmetry is absent. At present, numerically there are three different ways to judge whether a band insulator without inversion symmetry is a topological insulator or not: (i) Compute the $\mathbb{Z}_2$ numbers using the integration of both Berry's connection and curvature over half of the Brillouin zone (BZ). In order to do so, one has to set up a mesh in the k space and calculate the corresponding quantities on the lattice version of the problem.[18,41,42] Since the calculation involves the Berry connection, one has to numerically fix the gauge on the half BZ, which is not easy for the realistic wave functions obtained by first-principle calculation. (ii) Start from an artificial system with spacial inversion symmetry, and then smoothly "deform" the Hamiltonian towards the realistic one without inversion symmetry. If the energy gap never closes at any points in the BZ during the "deformation" process, the realistic system must share the same topological nature with the initial reference system whose $\mathbb{Z}_2$ number can be easily counted by the inversion eigenvalue formula. Unfortunately, making sure that the energy gap remains open on the whole BZ is very difficult numerically, especially in 3D. (iii) Directly calculate the surface states. For most of the topological insulator materials, the first-principle calculation for the surface states is numerically heavy. Therefore it is valuable to develop a mathematically equivalent way to calculate the $\mathbb{Z}_2$ numbers of a band insulator, which satisfies the following conditions: first, it should use only the periodic bulk system; second, it should not require any gauge-fixing condition—thereby greatly simplifying the calculation; third, it should be easily applied to general systems lacking spacial inversion symmetry.

In the present paper we propose an equivalent expression for the $\mathbb{Z}_2$ topological invariant using the U($2N$) non-Abelian Berry connection. Based on this expression, we further propose a numerical method to calculate the $\mathbb{Z}_2$ topological number for general band insulators, without choosing a gauge-fixing

condition. The main idea of the method is to calculate the evolution of the Wannier function center directly during a "time-reversal pumping" process, which is a $\mathbb{Z}_2$ analog to the charge polarization.[43,44] We derive that the center of the Wannier function for the effective one-dimensional (1D) system can be expressed as the U(1) phase of the eigenvalues of a matrix obtained as the product of the U($2N$) Berry connection along the "Wilson loop." The $\mathbb{Z}_2$ topological numbers can be expressed as the number of times mod 2 of the partner switching of these phases during a complete period of the time-reversal pumping process. Using this method, we have recalculated the $\mathbb{Z}_2$ topological numbers for several topological insulator systems, including strained HgTe, Bi, Sb, and Bi_2Se_3, and found the "partner switching" patterns, which differentiate between topologically trivial and nontrivial behavior. The rest of paper will be organized as follows: in Sec. II we derive the mathematical form of the $\mathbb{Z}_2$ numbers through the "Wilson loop"; we apply our method to various topological insulator systems in Sec. III; we prove the equivalence of our methods and the Z_2 number proposed by Fu and Kane[7] in the appendices.

II. THE FORMALISM

We assume that the Hamiltonian of a band insulator with both time reversal and translational symmetry can be expressed in terms of a complete set of local basis $\phi_\alpha(\mathbf{r}-\mathbf{R}_i)$, where $\mathbf{R}_i$ denotes the position of the ith lattice site and α denotes the index of the local basis. We have

$$H=\sum_{\alpha\beta}\sum_{ij}h_{ij}^{\alpha\beta}|\alpha i\rangle\langle\beta j|+\text{H.c.}, \tag{1}$$

where $h_{ij}^{\alpha\beta}=\int\phi_\alpha^*(\mathbf{r}-\mathbf{R}_i)H_{\text{LDA}}(\mathbf{r})\phi_\beta(\mathbf{r}-\mathbf{R}_j)dr^3$, with $H_{\text{LDA}}(\mathbf{r})$ being the Hamiltonian obtained by the first-principle calculation, i.e., the local-density approximation (LDA).

Therefore the Bloch eigenstate of the above tight-binding Hamiltonian can be expressed as

$$|\Psi_{n\mathbf{k}}\rangle=\sum_\alpha g_{n\alpha}(\mathbf{k})|\alpha\mathbf{k}\rangle, \tag{2}$$

where $|\alpha\mathbf{k}\rangle=\frac{1}{\sqrt{N_{\text{cell}}}}\sum_i|\alpha i\rangle e^{i\mathbf{k}\cdot\mathbf{R}_i}$ is the Bloch sum of the corresponding local basis, n is the band index, and $g_{n\alpha}(\mathbf{k})$ can be viewed as the wave function of the Bloch state in the momentum space. For convenience we denote the periodic part of the Bloch eigenstate by $|n,\mathbf{k}\rangle$, where $|n,\mathbf{k}\rangle=e^{-i\mathbf{k}\cdot\mathbf{r}}|\Psi_{n\mathbf{k}}\rangle$, with $\mathbf{r}$ the position operator.

We will first focus on the 2D system. The topological nature of a 3D insulator can later be determined by looking at these effective 2D systems with one of the momenta fixed at $k_i=0$ and $k_i=\pi$ $(i=x,y,z)$. In Ref. 7 it was shown that the topological invariant in the 2D topological insulator can be described by an adiabatic pumping of "time-reversal polarization." Each wave vector k_y defines a one-dimensional subsystem, for which the time-reversal polarization is defined by splitting the bands in the system into two groups which are time-reversal partners of each other, and calculating the net charge polarization of one group of bands. The charge polarization is related to the position of Wannier functions in the subsystem, and the time-reversal polarization can be understood as the adiabatic shift of the Wannier functions. This calculation requires one to make a global choice in splitting the occupied bands in the system into two groups, which may be difficult to do, in general.

The main idea of our formalism is to directly look at the evolution of Wannier function centers for these effective 1D systems of fixed k_y in the subspace of occupied states. Fixing k_y, maximally localized Wannier functions (MLWF) in the one-dimensional subsystem can be obtained as eigenstates of the position operator projected into occupied subspace.[45] For a 1D lattice system with a periodic condition, the position operator is defined as

$$\hat{X}=\sum_{i\alpha}e^{-i\delta k_x\cdot\mathbf{R}_i}|\alpha i\rangle\langle\alpha i|, \tag{3}$$

where $\delta k_x\equiv\frac{2\pi}{N_x a_x}$, N_x is the number of real-space unit cells along the x direction, a_x is the lattice constant, α is the orbital and spin index, and $\mathbf{R}_i$ labels the unit cell. We note that the periodic boundary condition is used here. In the limit of an infinite lattice $\delta k_x\to 0$ and we can instead define the Hermitian position operator $\hat{x}=\sum_{i\alpha}R_i|\alpha i\rangle\langle\alpha i|=i\frac{\partial}{\partial k_x}$, but it is more convenient for us to use the periodic boundary condition for the purpose of numerical calculation. The operator $\hat{X}$ is a unitary operator with all the eigenvalues being $e^{-i\delta k_x\cdot\mathbf{R}_i}$, whose phase represents the position. The eigenvalue of the position operator can be viewed as the center of MLWF formed by the bands included in the operator $\hat{X}$.[45] Because the local basis set α is assumed to be complete, such MLWFs are always well defined. As pointed out by Fu and Kane, the $\mathbb{Z}_2$ topological invariant can be determined by looking at the evolution of the Wannier function center for the effective 1D system with fixed k_y in the subspace spanned by the occupied bands only. The projection operator for the occupied subspace can be defined as

$$\hat{P}_{k_y}=\sum_{n\in o,k_x}|\Psi_{n\mathbf{k}}\rangle\langle\Psi_{n\mathbf{k}}|, \tag{4}$$

where o in the summation means the occupied bands. Therefore we should consider the eigenvalue of the projected position operator defined as

$$\begin{aligned}\hat{X}_P(k_y)&=\hat{P}_{k_y}\hat{X}\hat{P}_{k_y}\\&=\sum_{nm\in o}\sum_{k_xk_x',i\alpha}e^{-i\delta k_x\cdot\mathbf{R}_i}|\Psi_{nk_x,k_y}\rangle\langle\Psi_{nk_x,k_y}|\\&\quad\times|\alpha i\rangle\langle\alpha i|\Psi_{mk_x',k_y}\rangle\langle\Psi_{mk_x',k_y}|\\&=\sum_{nm\in o}\sum_{k_xk_x',i}\frac{1}{N_{\text{cell}}}e^{i(k_x+\delta k_x-k_x')\cdot\mathbf{R}_i}|\Psi_{nk_x,k_y}\rangle\langle\Psi_{mk_x',k_y}|\\&\quad\times\left[\sum_\alpha g_{n\alpha}^*(k_x)g_{m\alpha}(k_x')\right]\\&=\sum_{k_xk_x'}\delta(k_x+\delta k_x-k_x')\sum_{nm\in o}|\Psi_{nk_x,k_y}\rangle\langle\Psi_{mk_x',k_y}|\\&\quad\times\left[\sum_\alpha g_{n\alpha}^*(k_x)g_{m\alpha}(k_x')\right].\end{aligned} \tag{5}$$

The above operator can be written in a more suggestive matrix form:

$$\hat{X}_P(k_y) = \begin{bmatrix} 0 & F_{0,1} & 0 & 0 & 0 & 0 \\ 0 & 0 & F_{1,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & F_{2,3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & F_{N_x-2,N_x-1} \\ F_{N_x-1,0} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{6}$$

where

$$F^{nm}_{i,i+1}(k_y) = \sum_\alpha g^*_{n\alpha}(k_{x,i},k_y) g_{m\alpha}(k_{x,i+1},k_y) \tag{7}$$

are the $2N \times 2N$ matrices spanned in $2N$-occupied states and $k_{x,i} = \frac{2\pi i}{N_x a_x}$ are the discrete k points taken along the x axis.

In the following we will further prove that for a lattice Hamiltonian equation (1) and Eq. (7) is equal to the inner product of the periodic parts of the Bloch functions. We note that for a Hamiltonian defined on the lattice, the local bases $|\alpha i\rangle$ can be factorized into the "internal freedom" part and spacial part,

$$|\alpha i\rangle = |\alpha\rangle \delta(\mathbf{r} - \mathbf{R}_i), \tag{8}$$

where α is the combined index for spin and orbital, $|\alpha\rangle = (0,0,\ldots,1,\ldots,0,0)^T$ is the basis vector with only the α component nonzero. Under this definition we have

$$\begin{aligned} \langle n,k|m,k'\rangle &= \langle \Psi_{nk}|e^{ikr}e^{-ik'r}|\Psi_{m,k'}\rangle \\ &= \sum_{\alpha,\beta} \frac{1}{N_{\text{cell}}} g^*_{n\alpha}(k) g_{m\beta}(k') \\ &\quad \times \sum_{j_1,j_2} \langle \alpha j_1|e^{i(k-k')r}|\beta j_2\rangle e^{i(k'R_{j_2}-kR_{j_1})} \\ &= \sum_{\alpha,\beta} \frac{1}{N_{\text{cell}}} g^*_{n\alpha}(k) g_{m\beta}(k') \\ &\quad \times \sum_{j_1,j_2} \langle \alpha|\beta\rangle \delta_{j_1 j_2} e^{i(k-k')R_{j_2}} e^{i(k'R_{j_2}-kR_{j_1})} \\ &= \sum_{\alpha,\beta} g^*_{n\alpha}(k) g_{m\beta}(k') \delta_{\alpha\beta} \\ &= \sum_{\alpha} g^*_{n\alpha}(k) g_{m\alpha}(k') \end{aligned} \tag{9}$$

and equivalently, we have

$$F^{mn}_{i,i+1}(k_y) = \langle m,k_{x,i},k_y|n,k_{x,i+1},k_y\rangle. \tag{10}$$

The eigenproblem of $\hat{X}_P(k_y)$ can be solved by the transfer matrix method. We can define a product of $F_{i,i+1}$ as

$$D(k_y) = F_{0,1}F_{1,2}F_{2,3}\cdots F_{N_x-2,N_x-1}F_{N_x-1,0}. \tag{11}$$

$D(k_y)$ is a $2N \times 2N$ matrix which has $2N$ eigenvalues:

$$\lambda^D_m(k_y) = |\lambda^D_m| e^{i\theta^D_m(k_y)}, \quad m = 1,2,\ldots,2N,$$

where $\theta^D_m(k_y)$ is the phase of the eigenvalues:

$$\theta^D_m(k_y) = \text{Im}\big[\log \lambda^D_m(k_y)\big]. \tag{12}$$

The eigenvalues of $D(k_y)$ are gauge invariant under a U($2N$) transformation of $|n,\mathbf{k}\rangle$. We can easily prove that the eigenvalue of projected position operator $\hat{X}_P(k_y)$ can be simply related to the eigenvalue of the above D matrix by

$$\lambda^P_{m,n} = N_x\sqrt{\lambda^D_m} = N_x\sqrt{|\lambda^D_m|}e^{i(\theta^D_m+2\pi n)/N_x}, \tag{13}$$

where $n = 1,2,\ldots,N_x$. We can further prove that the D matrix is unitary and all the $|\lambda^D_m|$ equals one.

For infinitesimal $\delta k = \frac{2\pi}{N_x a_x} \ll 2\pi$, we have

$$\begin{aligned} F^{mn}_{i,i+1} &= \langle m,k_{x,i},k_y|n,k_{x,i+1},k_y\rangle \\ &= \delta_{mn} + \langle m,k_{x,i},k_y|(|n,k_{x,i+1},k_y\rangle - |n,k_{x,i},k_y\rangle) \\ &= \delta_{mn} - iA^{mn}_{i,i+1}\delta k \\ &\approx e^{-iA^{mn}_{i,i+1}\delta k}, \end{aligned} \tag{14}$$

where $A^{mn}_{i,i+1} = i\frac{\langle m,k_{x,i},k_y|(|n,k_{x,i+1},k_y\rangle - |n,k_{x,i},k_y\rangle)}{\delta k}$ is the non-Abelian U($2N$) gauge field. Hence we have that $D(k_y)$ is

$$\begin{aligned} D(k_y) &= \left[\prod_{i=0}^{N_x-1} F_{i,i+1}\right] = \left[\prod_{i=0}^{N_x-1} e^{-iA_{i,i+1}\delta k}\right] \\ &= \left\{P\exp\left[\int_{C_{k_y}} -iA(k)\,dk\right]\right\}. \end{aligned} \tag{15}$$

This is just the U($2N$) [*not* SU($2N$)] Wilson loop, where the contour C_{k_y} is a contour at fixed k_y which goes across the BZ in k_x, i.e., goes from $k_x = -\pi$ to $k_x = \pi$, through $k_x = 0$.

The evolution of the Wannier function center for the effective 1D system with k_y can be easily obtained by looking at the phase factor θ^D_m. To see it more clearly, we fold θ axes and glue $\theta = -\pi$ line and $\theta = \pi$ line together. Then the Wannier function centers live on a cylinder surface, as shown in Fig. 1. At $k_y = 0$, the eigenvalues of the D matrix appear in degenerate pairs due to time-reversal symmetry, which results in pairs of Wannier centers sitting at $k_y = 0$. When k_y moves away from the origin, the Wannier center pairs split and recombine at $k_y = \pi$ (where they again have to be degenerate due to time reversal), as shown in Fig. 1. Because θ^D_m is a phase factor, when two θ^D_m's meet together, they may differ by integer times of 2π. Therefore the evolution of each Wannier center pair will enclose the whole cylinder an integer times, which can be viewed as the winding number of the Wannier center pair. The $\mathbb{Z}_2$ topological number is related to the summation of the winding numbers for all the pairs. If it is odd, then the $\mathbb{Z}_2$ topological number is odd. It seems that the total winding number of the Wannier center pairs should generate an integer class Z instead of $\mathbb{Z}_2$. To clarify this point, let us look at the evolution of a Wannier center pair with winding number 4π. In that particular case, as shown in Fig. 1(c), the pair of Wannier centers must have an extra "accidental" degeneracy between $k_y = 0$ and π, which is not protected by any symmetry and can be removed by "deforming" the Hamiltonian slightly to make the crossing of the levels become an anticrossing, as shown in Fig. 1(c). The "deformation" process will thus change the total winding number by 4π and make it 0. Therefore only the total winding number mod 2 is a topological invariant.

Equation (11) can be viewed as the discrete expression of the Wilson loop for the U($2N$) non-Abelian Berry connection. It is obviously invariant under the U($2N$) gauge transformation

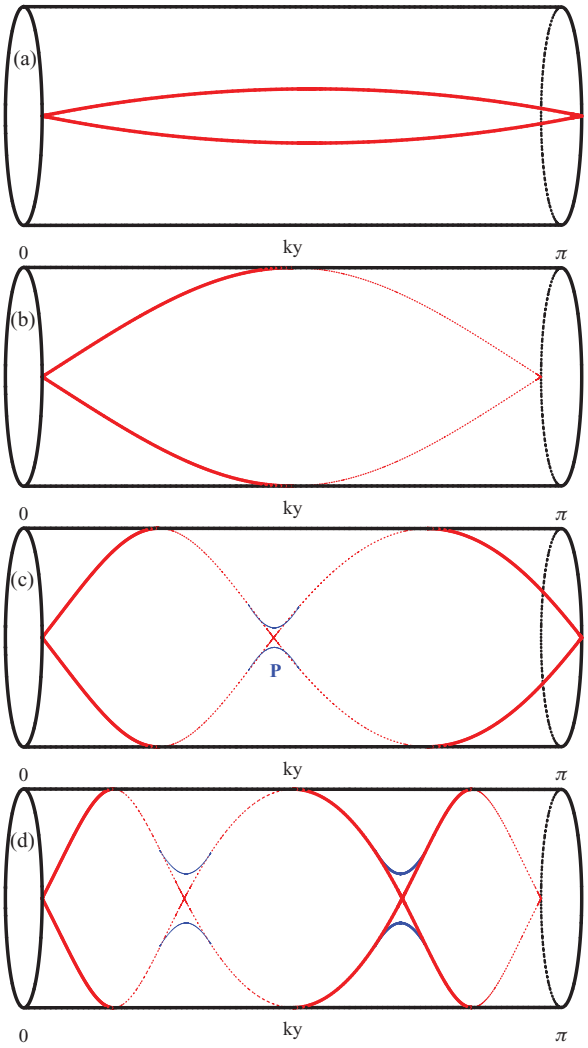

FIG. 1. (Color online) Schematic plots of the Wannier center curves: (a) for the trivial case, the Wannier center winding the cylinder zero times; (b) the Wannier center winding the cylinder one time; (c) the Wannier center winding the cylinder twice, and the cross point labeled by P is not protected by time-reversal symmetry and it is usually eliminated by some perturbation terms; and (d) the Wannier center winding the cylinder three times, which is topologically equal to the case in (b).

and thus can be calculated directly from the wave functions obtained by first-principle method without choosing any gauge-fixing condition, which is the biggest advantage of the present form of the $\mathbb{Z}_2$ invariance. From the discussion above, one can see intuitively that the evolution of Wannier function obtained in this way agrees with the conclusion of Fu and Kane, although they are obtained in different approaches. We also find a rigorous proof of the equivalence of the topological invariant obtained in the current approach with the $\mathbb{Z}_2$ number proposed by Kane and Mele,[4] which is presented in the appendix.

III. NUMERICAL RESULTS

In this section, we will implement the method presented above and explicitly compute the $\mathbb{Z}_2$ invariant for a series of systems. For each particular system, we calculate the evolution of the θ^D defined in Eq. (12) as the function of k_y from zero to π. The winding number of the Wannier center pairs defined in the above section can be checked in an equivalent way which is much simpler in practice. We first draw an arbitrary reference line parallel to the k_y axis, then compute the $\mathbb{Z}_2$ number by counting how many times the evolution lines of the Wannier centers cross the reference line.

A. BHZ model

Bernevig, Hughes, and Zhang (BHZ) showed that for an appropriate range of well thickness, the HgTe/CdTe quantum well exhibits an inverted subband structure. In this inverted regime, the system exhibits a 2D QSHE.[13] BHZ introduce a simple four-band tight-binding model to describe this effect:

$$H_{\text{eff}}(k_x,k_y) = \begin{bmatrix} H(\mathbf{k}) & 0 \\ 0 & H^*(-\mathbf{k}) \end{bmatrix}, \tag{16}$$

where $H(\mathbf{k}) = \varepsilon(\mathbf{k}) + d_i(\mathbf{k})\sigma_i$, $d_1 + id_2 = Aa^{-1}[\sin k_x a + i \sin k_y a]$, $d_3 = -2Ba^{-2}[2 - \frac{M}{2B} - \cos k_x a - \cos k_y a]$, and $\varepsilon(\mathbf{k}) = C - 2Da^{-2}[2 - \cos k_x a - \cos k_y a]$. The constants A,B,C,D are given in the caption of Fig. 2. Below we take the lattice constant $a = 1$. The basis order is $|E_1+\rangle,|H_1+\rangle$, $|E_1-\rangle,|H_1-\rangle$ and the relevant subbands $|E_1\pm\rangle$ and $|H_1\pm\rangle$ are two sets of Kramers' partners under the presence of time-reversal symmetry. H_{eff} consists of two decoupled blocks which are two copies of the massive Dirac Hamiltonian with a k-dependent mass $M(k)$. The coupling terms between blocks$|E1+\rangle,|H1+\rangle$ and $|E1-\rangle,|H1-\rangle$ will be induced by the breaking of inversion symmetry and were ignored in their original paper. When the effect of inversion symmetry breaking is taken into account, additional terms must be included in the

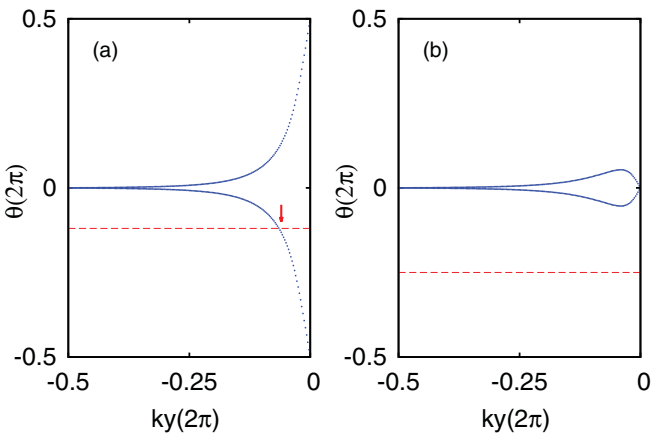

FIG. 2. (Color online) Wannier centers for the BHZ model. (a) For the quantum spinn Hall phase ($A = -13.68$ eV Å, $B = -16.9$ eV Å^2, $C = -0.0263$ eV, $D = -0.514$ eV Å^2, and $M = -2.058$ eV, $\Delta = 1.20$ eV), the Wannier center crosses the reference line (red dashed line) once (odd number of times). (b) For the normal insulating phase ($A = -14.48$ eV Å, $B = -18.0$ eV Å^2, $C = -0.018$ eV, $D = -0.594$ eV Å^2, $M = 2.766$ eV, and $\Delta = 1.20$ eV), the Wannier center crosses the reference line (red dashed line) zero (even number of) times.

effective model describing the mixing between the two blocks, which reads

$$H' = \begin{bmatrix} 0 & 0 & 0 & \Delta \\ 0 & 0 & -\Delta & 0 \\ 0 & -\Delta & 0 & 0 \\ \Delta & 0 & 0 & 0 \end{bmatrix}. \quad (17)$$

The BHZ model has phase transition from normal insulator to topological insulator [quantum spin Hall (QSH) phase] when M changes sign from positive to negative. We then apply our method to calculate the evolution pattern of the Wannier centers based on the above model Hamiltonian and show the results in Fig. 2.

The corresponding results for the topological insulator phase in the BHZ model are shown in Fig. 2(a) with the parameters taken from Ref. 13, as listed in the figure caption. When moving from $k_y = 0$ to π we see that the two evolution lines of Wannier centers enclose the cylinder once, and equivalently these evolution lines cross the reference line (the red dashed line) only once. By contrast, for the normal insulator phase, as shown in Fig. 2(b), the two evolution lines never cross the reference line. We emphasize that the reference line can be moved to somewhere else, but the even and odd properties of the crossing numbers between the evolution lines and reference line will never change, which determines the $\mathbb{Z}_2$ topological invariance. Therefore in the BHZ toy model for the topological insulator, the $\mathbb{Z}_2$ number calculated by our method is consistent with the previous conclusion. Next we will apply the method to more realistic models of other topological insulator materials.

B. CdTe and HgTe

The CdTe and HgTe materials have similar zinc-blende structures without bulk inversion symmetry. CdTe has a normal electronic structure, where the conduction bands (Γ_6) have the s-like character and the valance bands have the p-like character(Γ_8) throughout the whole BZ. For HgTe, the band structure is inverted in a small area near the Γ point, where the s-like Γ_6 bands sink below the p-like Γ_8 bands. The band inversion at the Γ point changes the topological nature of the band structure and makes the HgTe a topological insulator if a true energy gap is opened by the lattice distortion.[46] (As pointed out in Ref. 46, when the uniaxial strain is applied along the [001] direction for HgTe by choosing the ratio of lattice constants c/a to be 0.98, an energy gap $\sim$0.05 eV is opened at the Γ point.) We then use the tight-binding model[47] to calculate the pattern of the Wannier center evolution θ defined in Eq. (12) and show the results in Fig. 3.

We investigate the topological properties of a 3D bulk system by checking two planes in the k space, namely, the planes with $k_z = 0$ and $k_z = \pi$. It is clear that in the HgTe system, for $k_z = 0$ the evolution lines cross the reference line (red dashed line) once [as shown in Fig. 3(a)], while for $k_z = \pi$ they never cross it [as shown in Fig. 3(b)]. The above results indicate that for HgTe the effective 2D systems with $k_z = 0$ and π are effectively 2D topological insulator and normal insulator, respectively, which determines HgTe to be a strong 3D topological insulator.[15] A similar analysis can be also applied to CdTe and the results are shown in Figs. 3(c) and 3(d), indicating CdTe to be a normal insulator.

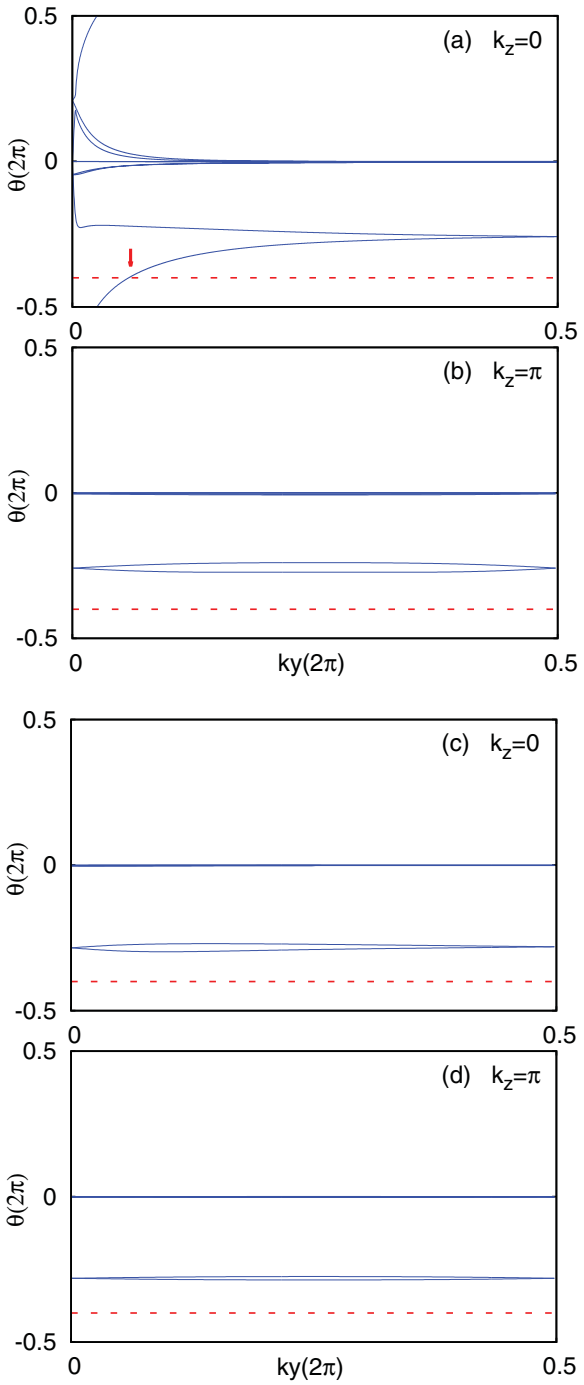

FIG. 3. (Color online) The evolution lines of Wannier centers for HgTe [(a and (b)] and CdTe [(c) and (d)]. For the HgTe system, the evolution lines cross the reference line an odd number of times in the $k_z = 0$ plane and an even number of times in the $k_z = \pi$ plane, indicating HgTe is a strong topological insulator. For CdTe, the evolution lines cross the reference line zero times for both $k_z = 0$ and π planes, indicating CdTe is a normal insulator.

C. Bi_2Se_3 system

Recently, the tetradymite semiconductors Bi_2Te_3, Bi_2Se_3, and Sb_2Te_3 have been theoretically predicted and experimentally observed to be topological insulators with a bulk band gap as large as 0.3 eV in Bi_2Se_3.[20,21,23,25,26,39] The surface states

in Bi_2Se_3 have been found by both ARPES[21,23] and STM,[36] which are consistent with the theoretical results.[20]

Since the Bi_2Se_3 family has inversion symmetry, the $\mathbb{Z}_2$ topological number can be easily calculated by the production of the parities at each high symmetry point in the BZ.[16] Below we apply our method to calculate the topological properties of this system, using the tight-binding model obtained in Ref. 20. We first perform the calculation for the Bi_2Se_3 without spin-orbit coupling (SOC); the results are shown in Figs. 4(a) and 4(d). It is clear that the evolution lines never cross the reference line for both $k_z = 0$ and π, indicating that the system is topologically trivial without SOC. When the realistic SOC is turned on, as shown in Figs. 4(c) and 4(d), the evolution lines cross the reference line once only in the case of $k_z = 0$ but not for $k_z = \pi$, indicating the Bi_2Se_3 bulk material is a 3D strong topological insulator.

D. Bi_2Te_3 slab system

As calculated by Liu *et al.*,[48] upon reducing the thickness of Bi_2Te_3 and Bi_2Se_3 films, the topological nature of the system alternates between topologically trivial and nontrivial behavior as a function of the layer thickness. Liu *et al.* pointed out that the 1QL Bi_2Te_3 slab is a trivial insulator and the 2QL Bi_2Te_3 slab is a 2D topological insulator.[48] We apply our method to these systems. The evolution patterns for the 1QL and 2QL Bi_2Te_3 slabs are obtained using the tight-binding Hamiltonian developed in Refs. 48 and 49, and the results are summarized in Fig. 5. In the 1QL slab system [Fig. 5(a)], the evolution pattern appears in a trivial manner, while that of the 2QL slab system is nontrivial [Fig. 5(b)]. This is consistent with the conclusion based on the parity counting.[48]

E. Bi and Sb system

Murakami pointed out that the $\mathbb{Z}_2$ topological number is odd in the 2D bilayer bismuth system.[50] We apply our method to this system using the tight-binding model developed in Ref. 51, which faithfully reproduces the bulk bismuth band structure. As shown in Fig. 6(a), the band structure of bilayer bismuth is topologically nontrivial, which is quite consistent with the previous conclusion.[50] After that we apply the same method to calculate bilayer Sb, which has a similar lattice structure to bismuth, but with relatively weak SOC. As plotted in Fig. 6(b), the evolution pattern of bilayer Sb shows clearly that it is in the normal insulator phase, which is also consistent with the parity counting.

F. Graphene system

Graphene consists of a honeycomb lattice of carbon atoms with two sites per unit cell. Kane and Mele introduced a tight-binding model which generalizes Haldane's model to include spin with time-reversal-invariant spin-orbit interactions:[4]

$$H = t \sum_{\langle i,j\rangle,\sigma} c^{\dagger}_{i,\sigma} c_{j,\sigma} + i\lambda_{\rm so} \sum_{\langle\langle i,j\rangle\rangle,\sigma\sigma'} v_{ij} c^{\dagger}_{i,\sigma} s^{z}_{\sigma\sigma'} c_{j,\sigma'} + i\lambda_R \sum_{\langle i,j\rangle,\sigma\sigma'} c^{\dagger}_{i,\sigma} (\mathbf{s}_{\sigma\sigma'} \times \hat{\mathbf{d}}_{ij})_z c_{j,\sigma'} + \lambda_v \sum_{i,\sigma} \xi_i c^{\dagger}_{i,\sigma} c_{i,\sigma}, \quad (18)$$

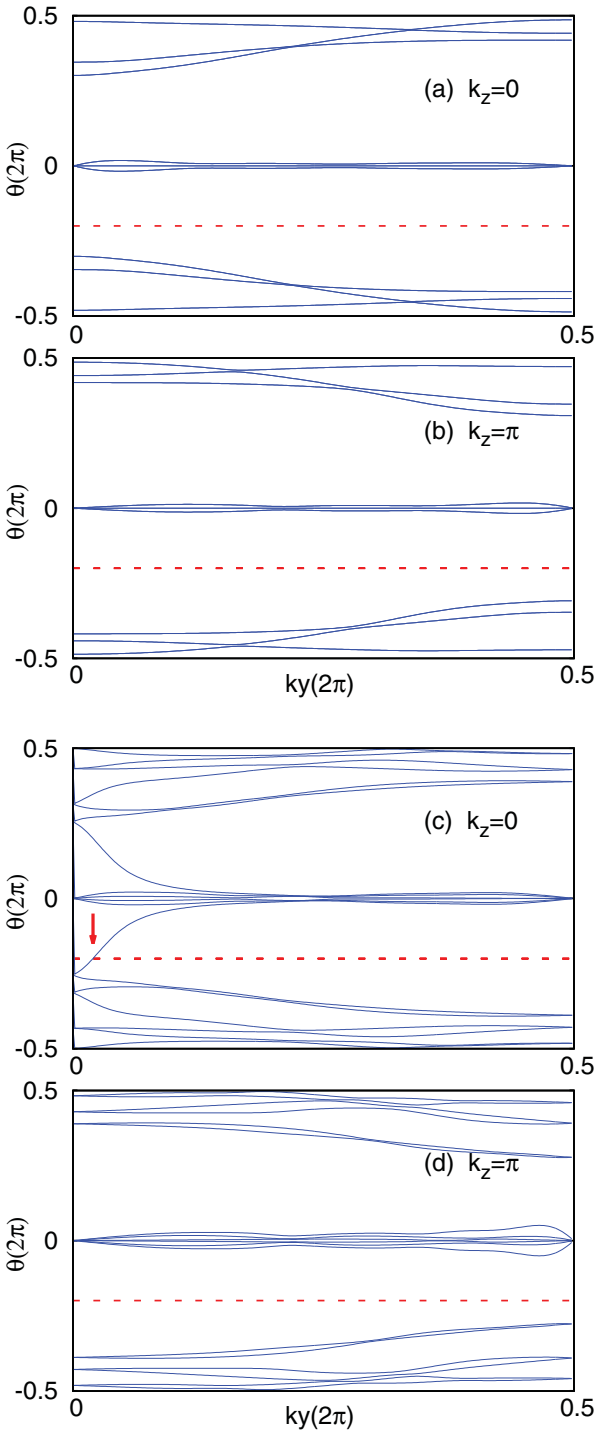

FIG. 4. (Color online) The evolution lines of Wannier centers for the Bi_2Se_3 system without [(a) and (b)] and with SOC [(c) and (d)]. If we turn off SOC the system is a normal insulator and the evolution lines never cross the reference line for both $k_z = 0$ and π planes as shown in (a) and (b), indicating the system to be a normal insulator. When the SOC is turned on the system is in a strong topological insulator phase, and the evolution lines cross the reference line an odd number of times in $k_z = 0$ and an even number of times in the $k_z = \pi$ plane as shown in (c) and (d), indicating the system is topologically nontrivial.

where i, j are the site indices and σ, σ' are the spin indices, and t is the nearest-neighbor hopping amplitude. In the second

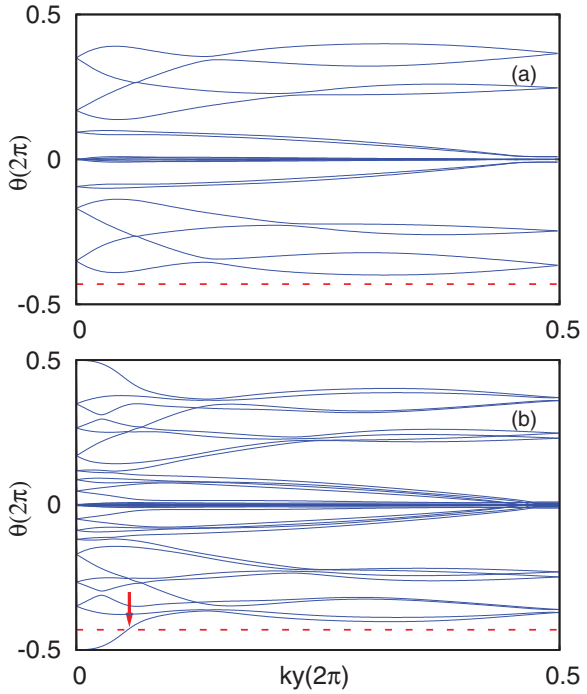

FIG. 5. (Color online) The evolution lines of Wannier centers for the 1QL and 2QL Bi_2Te_3 slabs. (a) The 1QL Bi_2Te_3 slab is in normal insulator phase. (b) The 2QL Bi_2Te_3 slab is in topological insulator phase.

term, λ_{so} is the strength of SOC between second neighbors, with $v_{ij} = (2/\sqrt{3})[\hat{\mathbf{d}}_1 \times \hat{\mathbf{d}}_2]_z = \pm 1$ depending on the relative orientation of the first-neighbor bond vectors $\hat{\mathbf{d}}_1$ and $\hat{\mathbf{d}}_2$ encountered by an electron hopping from site j to site i, and s_z is the z Pauli spin matrix. The third term is a nearest-neighbor Rashba term, which breaks the $z \to -z$ mirror symmetry, and can be generated by a perpendicular electric field or interaction with the substrate. The fourth term is a staggered sublattice potential, where ξ_i equals $+1$ and -1 on the A and B sites, respectively. In what follows we use t as the energy scale and fix $\lambda_{so} = 0.06t$ and $\lambda_R = 0.05t$. Varying the parameter λ_v allows us to switch from normal insulator to QSH phase. In the present study, we choose $\lambda_v = 0.1t$ for the QSH phase and $\lambda_v = 0.4t$ for the normal insulating phase. The calculated Wannier centers evolution patterns are shown in Fig. 7. It can be easily found that the evolution lines cut the reference line only once in Fig. 7(a) but not in Fig. 7(b), indicating the former is topologically nontrivial and the latter is trivial.

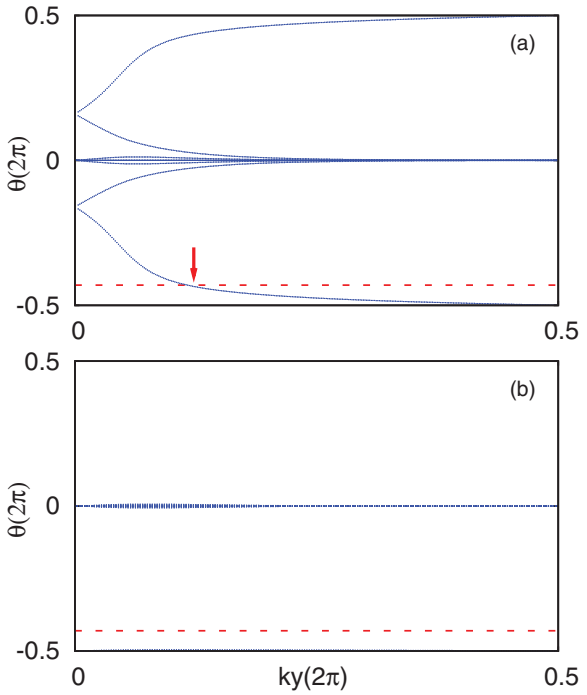

FIG. 6. (Color online) The evolution lines of Wannier centers for 2D bilayer Bi (a) and Sb (b) systems, indicating the bilayer Bi is topologically nontrivial but Sb is topologically trivial.

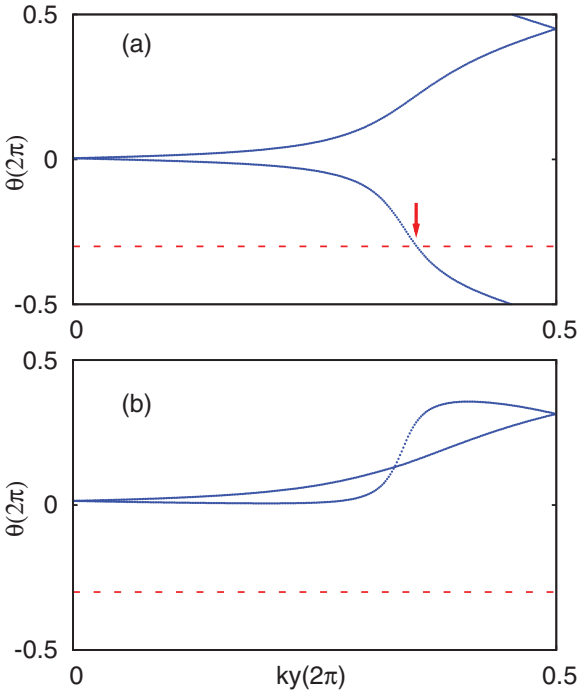

FIG. 7. (Color online) The evolution lines of Wannier centers for graphene in the (a) QSH phase $\lambda_v = 0.1t$ and (b) the normal insulating phase $\lambda_v = 0.4t$. In both cases $\lambda_{so} = 0.06t$ and $\lambda_R = 0.05t$.

In conclusion, we have proposed an equivalent expression for the $\mathbb{Z}_2$ topological invariance using the U($2N$) non-Abelian Berry connection. Based on this expression we calculated the evolution of the Wannier function center for several topological and normal insulating systems with or without inversion symmetry. We showed that for the nontrivial topological insulators, the Wannier function centers have partner switching patterns, topologically different from the normal (trivial) insulating systems. Additionally, we gave a proof that our method is equivalent to the $\mathbb{Z}_2$ number proposed by Fu and Kane.

Note added in proof. Recently, we noticed the paper by Soluyanov and Vanderbilt,[52] where the construction of Wannier functions for $\mathbb{Z}_2$ topological insulators are discussed from a different point of view. In this paper we addressed that the real construction of Wannier functions is not necessary, while only the "Wannier representation" and corresponding Berry connection evaluated along the "Wilson loop" are

essential keys in order to identify the topological nature. And at the same time, there is another paper by Ringel and Kraus,[53] which determines the $\mathbb{Z}_2$ invariant in a way that does not require any gauge fixing and leads to similar conclusions to ours.

ACKNOWLEDGMENTS

B.A.B. was supported by Princeton Startup Funds, Alfred P. Sloan Foundation, NSF DMR-095242 and NSF China 11050110420, and an MRSEC grant at Princeton University, NSF DMR-0819860. X.L.Q. is partly supported by Alfred P. Sloan Foundation. B.A.B. and X.L.Q. thank the Institute of Physics in Beijing, China for generous hosting. X.D. and Z.F. acknowledge support from NSF of China and from the 973 program of China (Grant No. 2007CB925000).

APPENDIX A: WILSON LOOP AND PFAFFIAN TOPOLOGICAL INVARIANT

In this appendix, we will prove that the U(2) Wilson loop is related to the Pfaffian topological invariant of time-reversal topological insulators. In a system with time-reversal invariant we can relate the bands at k and $-k$ through a unitary matrix B,

$$|n,-k\rangle = B^*_{nm}(k)\hat{T}|m,k\rangle, \tag{A1}$$

with $B(k)$ unitary and has the property

$$B(-k) = -B^T(k). \tag{A2}$$

We have the following relation between F_{k_1,k_2} matrices:

$$\begin{aligned} F^{mn}_{-k_2,-k_1} &= \langle m,-k_2|n,-k_1\rangle \\ &= \langle m',k_2|\hat{T}B_{mm'}(k_2)B^*_{nn'}(k_1)\hat{T}|n',k_1\rangle \\ &= B_{mm'}(k_2)B^*_{nn'}(k_2)F^{n'm'}_{k_1,k_2}, \end{aligned} \tag{A3}$$

or equivalently,

$$F_{-k_2,-k_1} = B(k_2)F^T_{k_1,k_2}B^\dagger(k_1). \tag{A4}$$

We now focus on the $k_y = 0$ or $k_y = \pi$ paths (say $k_y = 0$), each of which has k_x going from $-\pi$ to π, so that $k_y = -k_y$, and compute the finite difference:

$$\begin{aligned} D(k_y = 0) &= \prod_{i=0}^{N_x-1} F_{i,i+1} \\ &= F_{-(N_x/2)\Delta k,-[(N_x/2)-1]\Delta k}\cdots F_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta}. \end{aligned} \tag{A5}$$

All the F's in the above equation are considered at $k_y = 0$, and the k in the above equation expresses the k_x coordinate. By Eq. (A4), we have

$$\begin{aligned} F_{-\Delta k,0} &= B(\Delta k)F^T_{0,\Delta k}B^\dagger(0), \\ F_{-2\Delta k,-\Delta k} &= B(2\Delta k)F^T_{\Delta k,2\Delta k}B^\dagger(\Delta k),\ldots. \end{aligned} \tag{A6}$$

Hence the Wilson loop above becomes

$$\begin{aligned} D &= B\left(\frac{N_x}{2}\Delta k\right)F^T_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k}B^\dagger\left[\left(\frac{N_x}{2}-1\right)\Delta k\right] \\ &\quad\cdots B(2\Delta k)F^T_{\Delta k,2\Delta k}B^\dagger(\Delta k)B(\Delta k)F^T_{0,\Delta k}B^\dagger(0)F_{0,\Delta k} \\ &\quad\times F_{\Delta k,2\Delta k}F_{2\Delta k,3\Delta k}\cdots F_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k} \\ &= B\left(\frac{N_x}{2}\Delta k\right)F^T_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k} \\ &\quad\times F^T_{[(N_x/2)-2]\Delta k,[(N_x/2)-1]\Delta k} \\ &\quad\cdots F^T_{\Delta k,2\Delta k}F^T_{0,\Delta k}B^\dagger(0)F_{0,\Delta k}F_{\Delta k,2\Delta k}F_{2\Delta k,3\Delta k} \\ &\quad\cdots F_{[(N_x/2)-2]\Delta k,[(N_x/2)-1]\Delta k}F_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k}, \end{aligned} \tag{A7}$$

where we have used the fact that $B^\dagger(k)B(k) = I$ (unitary matrix). We see that all the intermediate B matrices vanish with the exception of the ones at the inversion symmetric points 0, $\frac{N_x}{2}\Delta k = \pi$. This is true for any time-reversal-invariant contour. Moreover, it is suggestive that the two leftover matrices $B^\dagger(0)$, $B(\frac{N_x}{2}\Delta k)$ should be brought together, so we must commute $B^\dagger(0)$ all across the matrix chain. The matrix $B(0)$ [and $B(\pi)$] has the property that it is unitary and antisymmetric Eq. (A2). We then know it has to be of the form

$$B(0) = e^{i\theta}\sigma_2. \tag{A8}$$

The matrix F_{k_1,k_2}, for $k_2 - k_1 \ll \pi$, as show in Eq. (14), has the following form:

$$F^{mn}_{k_1,k_2} = \delta_{mn} - iA^{mn}_{k_1,k_2}(k_2 - k_1), \tag{A9}$$

where $k_2 - k_1 = \Delta k$. We decompose the U(2) gauge field into its Abelian and non-Abelian parts:

$$A^{mn}_{k_1,k_2} = A^{\mathrm{U}(1)}_{k_1,k_2}\delta_{mn} + A^{\mathrm{SU}(2),i}_{k_1,k_2}(\sigma^i)_{mn}, \tag{A10}$$

where $i = 1,2,3$ and the double index implies summation. $A^{\mathrm{U}(1)}_{k_1,k_2}$, $A^{\mathrm{SU}(2),i}_{k_1,k_2}$ are numbers. We then have

$$\begin{aligned} (F^T_{k_1,k_2})B^\dagger(0) &= B^\dagger(0)\left[I - iA^{\mathrm{U}(1)}_{k_1,k_2}(k_2-k_1)I\right] \\ &\quad -iA^{\mathrm{SU}(2),i}_{k_1,k_2}(\sigma^i)^T(k_2-k_1)B^\dagger(0). \end{aligned} \tag{A11}$$

We aim to commute the matrix B with the Pauli matrices of the non-Abelian vector potential. We use the identities $\sigma^T_x\sigma_y = -\sigma_y\sigma_x$, $\sigma^T_y\sigma_y = -\sigma_y\sigma_y$, $\sigma^T_z\sigma_y = -\sigma_y\sigma_z$, and hence

$$\begin{aligned} A^{\mathrm{SU}(2),i}_{k_1,k_2}(\sigma^i)^T B^\dagger(0) &= e^{-i\theta}A^{\mathrm{SU}(2),i}_{k_1,k_2}(\sigma^i)^T\sigma_2 \\ &= -e^{-i\theta}\sigma_2 A^{\mathrm{SU}(2),i}_{k_1,k_2}\sigma^i = -B^\dagger(0)A^{\mathrm{SU}(2),i}_{k_1,k_2}\sigma^i. \end{aligned} \tag{A12}$$

Hence

$$\begin{aligned} (F^T_{k_1,k_2})B^\dagger(0) &= B^\dagger(0)\left[I + i\left(A^{\mathrm{U}(1)}_{k_1,k_2}I + A^{\mathrm{SU}(2),i}_{k_1,k_2}\sigma^i\right)(k_2-k_1)\right. \\ &\quad \left. -i2A^{\mathrm{U}(1)}_{k_1,k_2}I(k_2-k_1)\right]. \end{aligned} \tag{A13}$$

We also have

$$\begin{aligned} &I + i\left(A^{\mathrm{U}(1)}_{k_1,k_2}I + A^{\mathrm{SU}(2),i}_{k_1,k_2}\sigma^i\right)(k_2-k_1) - 2iA^{\mathrm{U}(1)}_{k_1,k_2}I(k_2-k_1) \\ &\quad = I + i(A_{k_1,k_2})^\dagger(k_2-k_1) - 2iA^{\mathrm{U}(1)}_{k_1,k_2}(k_2-k_1)I \\ &\quad \approx F^\dagger_{k_1,k_2}e^{-2iA^{\mathrm{U}(1)}_{k_1,k_2}(k_2-k_1)}, \end{aligned} \tag{A14}$$

where in the limit of $k_2 - k_1 \ll 2\pi$ (the case in all our terms) we neglect $(k_2 - k_1)^2$-order terms, and where A_{k_1,k_2} is the full U(2) field strength. We have also used the fact that A is an anti-Hermitian matrix $A^{mn}_{k_1,k_2}(k_2 - k_1) = -(A^{nm}_{k_1,k_2})^*(k_2 - k_1)$. After this long detour, we have proved

$$F^T_{k_1,k_2} B^\dagger(0) = B^\dagger(0) F^\dagger_{k_1,k_2} e^{-2iA^{\mathrm{U}(1)}_{k_1,k_2}(k_2-k_1)}. \tag{A15}$$

We then return to the U(2) Wilson loop

$$\begin{aligned} D &= B\left(\frac{N_x}{2}\Delta k\right) F^T_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k} F^T_{[(N_x/2)-2]\Delta k,[(N_x/2)-1]\Delta k} \\ &\quad \times F^T_{[(N_x/2)-3]\Delta k,[(N_x/2)-2]\Delta k} F^T_{[(N_x/2)-4]\Delta k,[(N_x/2)-3]\Delta k} \\ &\quad \cdots F^T_{\Delta k,2\Delta k} F^T_{0,\Delta k} B^\dagger(0) F_{0,\Delta k} F_{\Delta k,2\Delta k} F_{2\Delta k,3\Delta k} \\ &\quad \cdots F_{[(N_x/2)-2]\Delta k,[(N_x/2)-1]\Delta k} F_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k} \\ &= e^{-2i(A^{\mathrm{U}(1)}_{0,\Delta k} + A^{\mathrm{U}(1)}_{\Delta k,2\Delta k} + A^{\mathrm{U}(1)}_{2\Delta k,3\Delta k} + \cdots + A^{\mathrm{U}(1)}_{(N_x/2-1)\Delta k,(N_x/2)\Delta k})\Delta k} \\ &\quad \times B\left(\frac{N_x}{2}\Delta k\right) B^\dagger(0), \end{aligned} \tag{A16}$$

where we have used $U^\dagger_{0,\Delta k} U_{0,\Delta k} = U^\dagger_{\Delta k,2\Delta k} U_{\Delta k,2\Delta k} = I$, etc. Hence

$$D = e^{-2i[A^{\mathrm{U}(1)}_{0,\Delta k} + A^{\mathrm{U}(1)}_{\Delta k,2\Delta k} + \cdots + A^{\mathrm{U}(1)}_{[(N_x/2)-1]\Delta k,(N_x/2)\Delta k}]\Delta k} B(\pi) B^\dagger(0).$$

We note that the phase above is twice the U(1) (matrix, not traced) phase picked up from 0 to π. Let us redefine it by reexpressing it from $-\pi$ to , i.e., the full Abelian Berry for the interval considered. We have

$$\begin{aligned} A^{\mathrm{U}(1)}_{-k_2,-k_1}\Delta k &= \frac{i}{2}\mathrm{tr}[F_{-k_2,-k_1} - I] \\ &= \frac{i}{2}\mathrm{tr}\big[B(k_2) F^T_{k_1,k_2} B^\dagger(k_1) - I\big], \end{aligned} \tag{A17}$$

where we have used Eq. (A4). Since $k_2 - k_1 \ll 2\pi$, we can approximate $B(k_2) - B(k_1)$ as small and write $B(k_2) = B(k_1) + B(k_2) - B(k_1)$ to get

$$\begin{aligned} A^{\mathrm{U}(1)}_{-k_2,-k_1}\Delta k &= \frac{i}{2}\mathrm{tr}\big[B(k_1) F^T_{k_1,k_2} B^\dagger(k_1) - I \\ &\quad + [B(k_2) - B(k_1)] F^T_{k_1,k_2} B^\dagger(k_1)\big] \\ &= \frac{i}{2}\mathrm{tr}\big\{F^T_{k_1,k_2} - I + [B(k_2) - B(k_1)] F^T_{k_1,k_2} B^\dagger(k_1)\big\} \\ &= A^{\mathrm{U}(1)}_{k_1,k_2}\Delta k + \frac{i}{2}\mathrm{tr}\big\{[B(k_2) - B(k_1)] F^T_{k_1,k_2} B^\dagger(k_1)\big\}. \end{aligned} \tag{A18}$$

As $B(k_2) - B(k_1)$ is considered small for $k_2 - k_1 \ll 2\pi$, we take

$$[B(k_2) - B(k_1)] F^T_{k_1,k_2} \approx B(k_2) - B(k_1), \tag{A19}$$

where we took $F^T_{k_1,k_2} \approx I$ if multiplied by another small number. Hence

$$A^{\mathrm{U}(1)}_{-k_2,-k_1}\Delta k = A^{\mathrm{U}(1)}_{k_1,k_2}\Delta k + \frac{i}{2}\mathrm{tr}\{[B(k_2) - B(k_1)] B^\dagger(k_1)\}. \tag{A20}$$

We then find

$$\begin{aligned} &2(A^{\mathrm{U}(1)}_{0,\Delta k} + A^{\mathrm{U}(1)}_{\Delta k,2\Delta k} + \cdots + A^{\mathrm{U}(1)}_{(N_x/2-1)\Delta k,(N_x/2)\Delta k})\Delta k \\ &\quad = (A^{\mathrm{U}(1)}_{-(N_x/2)\Delta k,-(N_x/2-1)\Delta k} + \cdots + A^{\mathrm{U}(1)}_{-\Delta k,0} \\ &\quad\quad + A^{\mathrm{U}(1)}_{0,\Delta k} + \cdots + A^{\mathrm{U}(1)}_{(N_x/2-1)\Delta k,(N_x/2)\Delta k})\Delta k \\ &\quad\quad - \frac{i}{2}\int_0^\pi dk\, \mathrm{tr}[B^\dagger(k)\nabla_k B(k)]. \end{aligned} \tag{A21}$$

The first term is just the U(1) phase in the contour direction: $\int_{-\pi}^{\pi} A^{\mathrm{U}(1)}(k)dk$. The Wilson loop is then

$$\begin{aligned} W &= \exp\left\{\int_{-\pi}^{\pi} -iA^{\mathrm{U}(1)}(k)\,dk \right. \\ &\quad \left. - \frac{1}{2}\int_0^\pi dk\, \mathrm{tr}[B^\dagger(k)\nabla_k B(k)]\right\} B(\pi) B^\dagger(0). \end{aligned} \tag{A22}$$

As $B(k)$ is unitary, we also know that $\mathrm{tr}[B^\dagger(k)\nabla_k B(k)] = \nabla_k \ln\det B(k)$ and hence

$$\begin{aligned} &\exp\left\{-\frac{1}{2}\int_0^\pi dk\, \mathrm{tr}[B^\dagger(k)\nabla_k B(k)]\right\} \\ &\quad = \exp\left\{-\frac{1}{2}\log\left[\frac{\det B(\pi)}{\det B(0)}\right]\right\} = \sqrt{\frac{\det B(0)}{\det B(\pi)}}. \end{aligned} \tag{A23}$$

The Wilson loop becomes

$$D = \exp\left[\int_{-\pi}^{\pi} -iA^{\mathrm{U}(1)}(k)dk\right]\sqrt{\frac{\det B(0)}{\det B(\pi)}} B(\pi) B^\dagger(0). \tag{A24}$$

As we said before, $B(0)$,$B(\pi)$ are unitary, 2×2 matrices, antisymmetric, so

$$B(0) = \mathrm{Pf}\,[B(0)]\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{A25}$$

and similarly for $B(\pi)$, where Pf is the Pfaffian of the matrix. We hence have

$$D = \exp\left[\int_{-\pi}^{\pi} -iA^{\mathrm{U}(1)}(k)dk\right]\sqrt{\frac{\det B(0)}{\det B(\pi)}}\frac{\mathrm{Pf}[B(\pi)]}{\mathrm{Pf}[B(0)]} I, \tag{A26}$$

where I is the 2×2 identity matrix.

We now make several observations. Obviously, the above equality is valid on both time-reversal invariant lines at $k_y = 0$ and $k_y = \pi$, i.e., we can define two Wilson loops:

$$\begin{aligned} D(k_y = 0) &= \exp\left[\int_{-\pi}^{\pi} -iA^{\mathrm{U}(1)}(k_x, k_y = 0)dk_x\right] \\ &\quad \times \sqrt{\frac{\det B(0,0)}{\det B(\pi,0)}}\frac{\mathrm{Pf}[B(\pi,0)]}{\mathrm{Pf}[B(0,0)]} I, \end{aligned} \tag{A27}$$

and similarly for momentum π.

Second, we notice that the U(1) phase factor is not just the usual Abelian Berry phase but only *half* of it. Indeed, as per our definition,

$$A^{m,n}_{k_1,k_2} = A^{\mathrm{U}(1),i}_{k_1,k_2}\delta_{mn} + A^{\mathrm{SU}(2),i}_{k_1,k_2}(\sigma^i)_{mn}. \tag{A28}$$

This implies

$$A^{\mathrm{U}(1),i}_{\vec{k}} = \frac{1}{2}\sum_m \langle m,\vec{k}|m,\vec{k}\rangle, \tag{A29}$$

which has a $1/2$ difference from the usual form. This difference is actually important. Define

$$\Phi(k_y) = \oint_{-\pi}^{\pi} dk_x\, A_x(k_x,k_y) = i \log \det D(k_y). \quad \text{(A30)}$$

We then have

$$\int_0^{\pi} \nabla_{k_y} \Phi(k_y) = \int_0^{\pi} i \nabla_{k_y} \log \det D(k_y) = \Phi(\pi) - \Phi(0) + 2\pi M_n, \quad \text{(A31)}$$

where M_n is the winding number of the phase $\Phi(k_y)$. The phase $\Phi(k_y)$ is the sum of the phases $\phi_1(k_y)$ and $\phi_2(k_y)$ of the two eigenvalues of the Wilson loop *both defined in the interval* $[0,2\pi]$. Each of these eigenvalues has a winding number which adds to M_n, and the system will become nontrivial if the system has an odd M_n. We now take the Wilson loop W from $k_x = -\pi,\pi$ at $k_y = 0$, and then from $k_x = \pi, -\pi$ at $k_y = \pi$:

$$\begin{aligned} W &= D(k_y = 0)[D(k_y = \pi)]^{-1} \\ &= e^{(i/2)[\Phi(0)-\Phi(k)]} \prod_{i=\vec{G}_i/2} \frac{\sqrt{\det B(\vec{G}_i/2)}}{\mathrm{Pf}\,[B(\vec{G}_i/2)]} \\ &= \exp\left\{\frac{1}{2}\left[-\int_0^{\pi} \nabla_{k_y} \log \det D(k_y)\right] - \pi i M_n\right\} \\ &\quad \times \prod_{i=\vec{G}_i/2} \frac{\sqrt{\det B(\frac{\vec{G}_i}{2})}}{\mathrm{Pf}\,[B(\vec{G}_i/2)]}. \end{aligned}$$

$\vec{G}_i/2$ are the time-reversal invariant momenta $(0,0),(0,\pi),(\pi,0),(\pi,\pi)$. As such,

$$\begin{aligned} D(k_y = 0) e\left\{\frac{1}{2}\left[\int_0^{\pi} \nabla_{k_y} \log \det D(k_y)\right]\right\} [D(k_y = \pi)]^{-1} \\ = e^{-\pi i M_n} \prod_{i=\vec{G}_i/2} \frac{\sqrt{\det B(\vec{G}_i/2)}}{\mathrm{Pf}[B(\vec{G}_i/2)]}. \end{aligned} \quad \text{(A32)}$$

We have proved that both $D(k_y = 0)$ and $D(k_y = \pi)$ are proportional to unity matrix, up to a sign. For a smooth gauge, the difference in sign is taken by the contour term $\exp\{\frac{1}{2}[\int_0^{\pi} \nabla_{k_y} \log \det D(k_y)]\}$ to give

$$D(k_y = 0) e\left\{\frac{1}{2}\left[\int_0^{\pi} \nabla_{k_y} \log \det D(k_y)\right]\right\} [D(k_y = \pi)]^{-1} = I \quad \text{(A33)}$$

to give

$$e^{\pi i M_n} = \prod_{i=\vec{G}_i/2} \frac{\sqrt{\det B(\vec{G}_i/2)}}{\mathrm{Pf}[B(\vec{G}_i/2)]}, \quad \text{(A34)}$$

which says that the Pfaffian invariant is just the parity of the band switch number M_n. For M_n odd, it is nontrivial. Note that although our proof above is explicit only for two occupied bands, it can be easily extended to the $2N_{\mathrm{occupied}}$ band case when we realize that that case is just a tensor product (upon removing accidental degeneracies) of N_{occupied} time-reversal invariant multiplets for which the above expression applies.

APPENDIX B: WILSON LOOP AND THE Z_2 INVARIANT EXPRESSED AS AN OBSTRUCTION

An alternative formulation of the $\mathbb{Z}_2$ invariant has been defined by Fu and Kane,[7] where the $\mathbb{Z}_2$ invariant is expressed as an obstruction of the U(1) Berry's phase gauge field in half of the BZ. This approach has some convenience in its similarity with the Chern number formula of the quantum Hall states by Thouless *et al.*[2] The application of this approach to numerical calculation of the $\mathbb{Z}_2$ invariant in finite size systems has been studied by Fukui and Hatsugai.[18] Here we provide an alternative proof of the relation between our Wilson loop approach and the $\mathbb{Z}_2$ invariant through the obstruction formula. We start by reviewing the obstruction formulation of Fu and Kane[7] (Appendix A1). Consider $|n,k\rangle$ the occupied Bloch bands. We make the gauge choice

$$|n,-k\rangle = \mathcal{T}_{nm} T(|m,k\rangle), \quad \text{(B1)}$$

with $\mathcal{T}$ an antisymmetric matrix satisfying $\mathcal{T}^2 = -1$. Comparing to Eq. (A1), the gauge choice here corresponds to the requirement that $B_{nm}(k)$ is independent from k. More explicitly, with $2N$-occupied bands we can label the bands in pairs as $|n,k\rangle$ and $|\bar{n},k\rangle$ with $n = 1,\ldots,N$. Time reversal acts as $|\bar{n},-k\rangle = T|n,k\rangle$, $|n,-k\rangle = -T|\bar{n},k\rangle$, so that the wave functions in the lower half BZ defined by $k_y \in [-\pi,0]$ is determined by those in the upper half BZ, denoted by $\tau_{1/2}$. With this gauge choice, for a topological insulator it is not possible to make a continuous and single-valued choice of the wave functions in the whole BZ. However, it is always possible to define the wave functions continuously in the half BZ $\tau_{1/2}$, so that all obstructions are pushed to the boundary between the two half BZs, i.e., the two lines $k_y = 0$ and $k_y = \pi$. In such a gauge choice, Fu and Kane shows that the $\mathbb{Z}_2$ invariant is given by an obstruction in the half BZ:

$$\Delta = \frac{1}{2\pi}\left(\oint_{\partial\tau_{1/2}} \mathbf{a}\cdot\mathbf{dl} - \int_{\tau_{1/2}} d^2k\, F_{xy}\right) \mathrm{mod}\, 2, \quad \text{(B2)}$$

where $a_i = i\sum_n \langle nk|\partial_i|nk\rangle$ is the U(1) part of the Berry phase connection. By first integrating over k_x and defining the U(1) Wilson loop,

$$\Phi(k_y) = \oint_{-\pi}^{\pi} dk_x\, a_x(k_x,k_y). \quad \text{(B3)}$$

The $\mathbb{Z}_2$ invariant can be expressed as

$$\Delta = \frac{1}{2\pi}\left(\int_0^{\pi} dk_y \partial_{k_y} \Phi(k_y) - [\Phi(\pi) - \Phi(0)]\right) \mathrm{mod}\, 2. \quad \text{(B4)}$$

If we do not have any restriction on the gauge choice (other than requiring that the wave functions be continuous in $\tau_{1/2}$ so that $\Phi(k_y)$ is continuous and well defined), Δ can be 0, or any arbitrary integer. However, the gauge choice Eq. (B1) removes this ambiguity. Consider the states $|nk\rangle$ for $k_y = 0$. A gauge transformation

$$|nk\rangle \to e^{i\varphi_k}|nk\rangle \quad \text{(B5)}$$

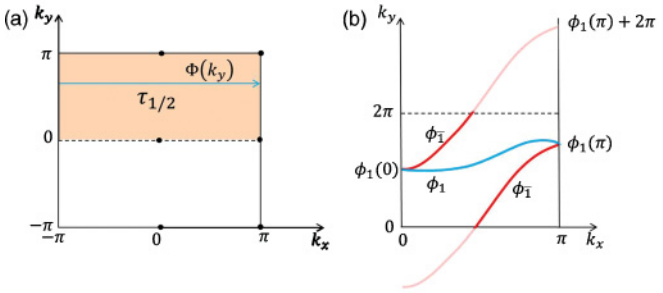

FIG. 8. (Color online) (a) Definition of the half BZ and the paths to define Wilson loops. (b) Schematic picture of eigenvalues $\phi_n(k_y)$ for a two-band model. $\phi_1(k_y)$ has winding number 0 and $\phi_{\bar{1}}(k_y)$ has winding number 1, when $\phi_1(0)$ and $\phi_1(\pi)$ are chosen to be in $[0,2\pi)$.

corresponds to a gauge transformation $\mathbf{A} \to \mathbf{A} + \nabla_{\mathbf{k}}\varphi$, where $A_i^{nm} = i\langle nk|\partial_i|mk\rangle$, which leads to the change in the flux

$$\Phi(k_y = 0) \to \Phi(k_y = 0) + 2N \oint_{-\pi}^{\pi} dk_x\, \partial_x \varphi_{k_x,0}. \quad \text{(B6)}$$

However, to preserve the condition (B1) we have to require

$$\varphi_{\mathbf{k}} = -\varphi_{-\mathbf{k}}, \quad \text{(B7)}$$

so that

$$\oint_{-\pi}^{\pi} dk_x\, \partial_x \varphi_{k_x,0} = 2\int_0^{\pi} \partial_x \varphi_{k_x,0} = 2(\varphi_{\pi,0} - \varphi_{0,0}). \quad \text{(B8)}$$

Since we also have $\varphi_{\pi,0}$ and $\varphi_{0,0} = 0 \,\mathrm{mod}\pi$ from Eq. (B7), the allowed change of $\Phi(k_y = 0)$ in the gauge transformations that preserves the gauge choice (B1) can only be

$$\begin{aligned}\Phi(k_y = 0) &\to \Phi(k_y = 0) + 4N(\varphi_{\pi,0} - \varphi_{0,0}) \\ &= \Phi(k_y = 0) + 4\pi n, n \in \mathbb{Z}. \end{aligned} \quad \text{(B9)}$$

The same is true for $\Phi(k_y = \pi)$. Consequently, $\Phi(k_y = 0)$ and $\Phi(k_y = \pi)$ are well-defined modular 4π, so that the $\mathbb{Z}_2$ quantity Δ defined by Eq. (B4) is well defined.

Now we relate this result to the non-Abelian Wilson loop. When there are $2N$ bands occupied, a U($2N$) Berry phase gauge field $A_i^{nm} = i\langle nk|\partial_i|mk\rangle$ is defined. We can define the U($2N$) Wilson loop along the same equal-k_y loops:

$$W(k_y) = P\exp\left[-i\oint dk_x\, A_x(k_x,k_y)\right] \in \mathrm{U}(2N). \quad \text{(B10)}$$

The U(1) gauge field is related to the U($2N$) gauge field by $a_i = \mathrm{Tr}\, A_i$ so that the U(1) flux $\Phi(k_y)$ is related to $W(k_y)$ as $e^{-i\Phi(k_y)} = \det W(k_y)$. Denote the eigenvalues of $W(k_y)$ as $e^{i\phi_n(k_y)}$ with $n = 1,\ldots,2N$, we have $\Phi(k_y) = \sum_n -\phi_n(k_y) \mathrm{mod}\, 2\pi$. Thus

$$\Delta = \frac{1}{2\pi}\left(\sum_n \int_0^{\pi} dk_y\, \partial_{k_y}\phi_n(k_y) - \sum_n [\phi_n(\pi) - \phi_n(0)]\right) \times \mathrm{mod}\, 2. \quad \text{(B11)}$$

Now we study the effect of the gauge choice (B1) on the U($2N$) gauge field.

$$\begin{aligned}\mathbf{A}_{-\mathbf{k}}^{nm} &= i\langle n,-k|\nabla_{-\mathbf{k}}|m,-k\rangle \\ &= -i\mathcal{T}_{ml}\mathcal{T}_{np}T(\langle pk|\nabla_{\mathbf{k}}|lk\rangle) \\ &= \mathcal{T}_{ml}\mathcal{T}_{np}\left(\mathbf{A}_{\mathbf{k}}^{pl}\right)^* = \left(\mathcal{T}\mathbf{A}_{\mathbf{k}}^T\mathcal{T}^{-1}\right)_{nm} \end{aligned} \quad \text{(B12)}$$

$$\Rightarrow W(-k_y) = \mathcal{T}W^T(k_y)\mathcal{T}^{-1}. \quad \text{(B13)}$$

Thus for $k_y = 0$ or π, $e^{i\phi_n(k_y)}$ is doubly degenerate. If we label a pair of degenerate eigenvalues by n and $\bar{n}$, the gauge choice (B1) corresponds to the choice of $\phi_n(k_y) = \phi_{\bar{n}}(k_y)$. Indeed, we see that if we make this choice, an ambiguity of 2π in ϕ_n leads to an ambiguity of 4π in $\sum_n \phi_n$. Thus Δ defined in Eq. (B11) is well defined. To simplify the formula, we choose $\phi_n(0)$ and $\phi_n(\pi)$ to be in $[0,2\pi)$. Thus

$$\int_0^{\pi} dk_y\, \partial_{k_y}\phi_n(k_y) = \phi_n(\pi) - \phi_n(0) + 2\pi M_n, \quad \text{(B14)}$$

with M_n the winding number of phase ϕ_n, which is equal to the number of times ϕ_n crosses the line $\phi_n = 2\pi$ from below. For example, in Fig. 8(b) ϕ_1 has winding number 0 and $\phi_{\bar{1}}$ has winding number 1. In this way we get

$$\Delta = \sum_n M_n \,\mathrm{mod}\, 2. \quad \text{(B15)}$$

The number $\sum_n M_n$ simply counts how many eigenvalues ϕ_n crosses $\phi = 2\pi$ line (or any other reference line) from below. Thus the $\mathbb{Z}_2$ invariant is simply determined by the parity of the number of eigenvalue curves $\phi_n(k_y)$ which crosses a reference line $\phi =$ constant.

[1] R. B. Laughlin, Phys. Rev. B **23**, 5632 (1981).
[2] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[3] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[4] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[5] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).
[6] R. Roy, Phys. Rev. B **79**, 195322 (2009).
[7] L. Fu and C. L. Kane, Phys. Rev. B **74**, 195312 (2006).
[8] B. A. Bernevig and S. C. Zhang, Phys. Rev. Lett. **96**, 106802 (2006).
[9] C. Wu, B. A. Bernevig, and S. C. Zhang, Phys. Rev. Lett. **96**, 106401 (2006).
[10] C. Xu and J. E. Moore, Phys. Rev. B **73**, 045322 (2006).
[11] A. Roth, C. Brune, H. Buhmann, L. W. Molenkamp, J. Maciejko, X. Qi, and S. Zhang, Science **325**, 294 (2009).
[12] M. Konig, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X. Qi, and S. Zhang, Science **318**, 766 (2007).
[13] B. A. Bernevig, T. L. Hughes, and S. Zhang, Science **314**, 1757 (2006).
[14] X. L. Qi, Y. S. Wu, and S. C. Zhang, Phys. Rev. B **74**, 085308 (2006).
[15] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).

[16]L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[17]J. C. Y. Teo, L. Fu, and C. L. Kane, Phys. Rev. B **78**, 045426 (2008).
[18]T. Fukui and Y. Hatsugai, J. Phys. Soc. Jpn. **76**, 053702 (2007).
[19]Y. Ran, Y. Zhang, and A. Vishwanath, Nat. Phys. **5**, 298 (2009).
[20]H. Zhang, C. Liu, X. Qi, X. Dai, Z. Fang, and S. Zhang, Nat. Phys. **5**, 438 (2009).
[21]Y. L. Chen, J. G. Analytis, J. Chu, Z. K. Liu, S. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fang, S. C. Zhang, I. R. Fisher, Z. Hussain, and Z.-X. Shen, Science **325**, 178 (2009).
[22]D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008).
[23]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).
[24]J. G. Analytis, J. H. Chu, Y. Chen, F. Corredor, R. D. McDonald, Z. X. Shen, and I. R. Fisher, Phys. Rev. B **81**, 205407 (2010).
[25]D. Hsieh, Y. Xia, D. Qian, L. Wray, F. Meier, J. H. Dil, J. Osterwalder, L. Patthey, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Phys. Rev. Lett. **103**, 146401 (2009).
[26]S. R. Park, W. S. Jung, C. Kim, D. J. Song, C. Kim, S. Kimura, K. D. Lee, and N. Hur, Phys. Rev. B **81**, 041405 (2010).
[27]S. Shen, e-print arXiv:0909.4125.
[28]B. Yan, C. Liu, H. Zhang, C. Yam, X. Qi, T. Frauenheim, and S. Zhang, Europhys. Lett. **90**, 37002 (2010).
[29]H. Lin, L. A. Wray, Y. Xia, S. Xu, S. Jia, R. J. Cava, A. Bansil, and M. Z. Hasan, Nat. Mater. **9**, 546 (2010).
[30]M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[31]W. Shan, H. Lu, and S. Shen, New J. Phys. **12**, 043048 (2010).
[32]P. Roushan, J. Seo, C. V. Parker, Y. S. Hor, D. Hsieh, D. Qian, A. Richardella, M. Z. Hasan, R. J. Cava, and A. Yazdani, Nature (London) **460**, 1106 (2009).
[33]Z. Alpichshev, J. G. Analytis, J. H. Chu, I. R. Fisher, and A. Kapitulnik, e-print arXiv:1003.2233.
[34]W. C. Lee, C. Wu, D. P. Arovas, and S. C. Zhang, Phys. Rev. B **80**, 245439 (2009).
[35]Y. Zhang, K. He, C. Chang, C. Song, L. Wang, X. Chen, J. Jia, Z. Fang, X. Dai, W. Shan, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, Nat. Phys. **6**, 584 (2010).
[36]T. Zhang, P. Cheng, X. Chen, J. Jia, X. Ma, K. He, L. Wang, H. Zhang, X. Dai, Z. Fang, Xincheng Xie, and Qi-Kun Xue, Phys. Rev. Lett. **103**, 266803 (2009).
[37]X. Zhou, C. Fang, W. F. Tsai, and J. P. Hu, Phys. Rev. B **80**, 245317 (2009).
[38]H. M. Guo and M. Franz, Phys. Rev. B **81**, 041102 (2010).
[39]D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, J. Osterwalder, L. Patthey, J. G. Checkelsky, N. P. Ong, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **460**, 1101 (2009).
[40]D. Hsieh, Y. Xia, L. Wray, D. Qian, A. Pal, J. H. Dil, J. Osterwalder, F. Meier, G. Bihlmayer, C. L. Kane, G. Bihlmayer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Science **323**, 919 (2009).
[41]T. Fukui, Y. Hatsugai, and H. Suzuki, J. Phys. Soc. Jpn. **74**, 1674 (2005).
[42]D. Xiao, Y. Yao, W. Feng, J. Wen, W. Zhu, X. Q. Chen, G. M. Stocks, and Z. Zhang, Phys. Rev. Lett. **105**, 096404 (2010).
[43]R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
[44]D. J. Thouless, Phys. Rev. B **27**, 6083 (1983).
[45]S. Kivelson, Phys. Rev. B **26**, 4269 (1982).
[46]X. Dai, T. L. Hughes, X. L. Qi, Z. Fang, and S. C. Zhang, Phys. Rev. B **77**, 125319 (2008).
[47]A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B **25**, 6367 (1982).
[48]C. X. Liu, H. J. Zhang, B. Yan, X. L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S. C. Zhang, Phys. Rev. B **81**, 041307 (2010).
[49]W. Zhang, R. Yu, H. Zhang, X. Dai, and Z. Fang, New J. Phys. **12**, 065013 (2010).
[50]S. Murakami, Phys. Rev. Lett. **97**, 236805 (2006).
[51]Y. Liu and R. E. Allen, Phys. Rev. B **52**, 1566 (1995).
[52]A. A. Soluyanov and D. Vanderbilt, Phys. Rev. B **83**, 035108 (2011) .
[53]Z. Ringel and Y. E. Kraus, Phys. Rev. B **83**, 245115 (2011).