



## Topological field theory of time-reversal invariant insulators

Xiao-Liang Qi, Taylor L. Hughes, and Shou-Cheng Zhang

*Department of Physics, Stanford University, Stanford, California 94305, USA*

(Received 27 February 2008; published 24 November 2008)

We show that the fundamental time-reversal invariant (TRI) insulator exists in 4+1 dimensions, where the effective-field theory is described by the (4+1)-dimensional Chern-Simons theory and the topological properties of the electronic structure are classified by the second Chern number. These topological properties are the natural generalizations of the time reversal-breaking quantum Hall insulator in 2+1 dimensions. The TRI quantum spin Hall insulator in 2+1 dimensions and the topological insulator in 3+1 dimensions can be obtained as descendants from the fundamental TRI insulator in 4+1 dimensions through a dimensional reduction procedure. The effective topological field theory and the  $Z_2$  topological classification for the TRI insulators in 2+1 and 3+1 dimensions are naturally obtained from this procedure. All physically measurable topological response functions of the TRI insulators are completely described by the effective topological field theory. Our effective topological field theory predicts a number of measurable phenomena, the most striking of which is the topological magnetoelectric effect, where an electric field generates a topological contribution to the magnetization in the same direction, with a universal constant of proportionality quantized in odd multiples of the fine-structure constant  $\alpha=e^2/\hbar c$ . Finally, we present a general classification of all topological insulators in various dimensions and describe them in terms of a unified topological Chern-Simons field theory in phase space.

DOI: [10.1103/PhysRevB.78.195424](https://doi.org/10.1103/PhysRevB.78.195424)

PACS number(s): 73.43.-f, 78.20.-e, 78.20.Ls, 03.65.Vf

### I. INTRODUCTION

Most states or phases of condensed matter can be described by local order parameters and the associated broken symmetries. However, the quantum Hall (QH) state<sup>1-4</sup> gives the first example of topological states of matter which have topological quantum numbers different from ordinary states of matter and are described in the low energy limit by topological field theories. Soon after the discovery of the integer QH effect, the quantization of Hall conductance in units of  $e^2/h$  was shown to be a general property of two-dimensional (2D) time reversal-breaking (TRB) band insulators.<sup>5</sup> The integral of the curvature of Berry's phase gauge field defined over the magnetic Brillouin zone (BZ) was shown to be a topological invariant called the first Chern number, which is physically measured as the quanta of the Hall conductance. In the presence of many-body interactions and disorder, the Berry curvature and the first Chern number can be defined over the space of twisted boundary conditions.<sup>6</sup> In the long wavelength limit, both the integer and the fractional QH effect can be described by the topological Chern-Simons field theory<sup>7</sup> in 2+1 dimensions. This effective topological field theory captures all physically measurable topological effects, including the quantization of the Hall conductance, the fractional charge, and the statistics of quasiparticles.<sup>8</sup>

Insulators in 1+1 dimensions can also have unique topological effects. Solitons in charge-density wave insulators can have fractional charge or spin-charge separation.<sup>9</sup> The electric polarization of these insulators can be expressed in terms of the integral of Berry's phase gauge field in momentum space.<sup>10,11</sup> During an adiabatic pumping cycle, the change in electric polarization or the net charge pumped across the one-dimensional (1D) insulator is given by the integral of the Berry curvature over the hybrid space of momentum and the adiabatic pumping parameter. This integral

is quantized to be a topological integer.<sup>12</sup> Both the charge of the soliton and the adiabatic pumping current can be obtained from the Goldstone-Wilczek formula.<sup>13</sup>

In this paper we shall show that the topological effects in the (1+1)-dimensional insulator can be obtained from the QH effect of the (2+1)-dimensional TRB insulator by a procedure called dimensional reduction. In this procedure one of the momenta is replaced by an adiabatic parameter or field, and the Goldstone-Wilczek formula, and thus, all topological effects of the (1+1)-dimensional insulators, can be derived from the (2+1)-dimensional QH effect. The procedure of dimensional reduction can be generalized to the higher dimensional time reversal invariant (TRI) insulators and beyond, which is the key result of this paper.

In recent years, the QH effect of the (2+1)-dimensional TRB insulators has been generalized to TRI insulators in various dimensions. The first example of a topologically nontrivial TRI state in condensed matter context was the four-dimensional (4D) generalization of the quantum Hall (4DQH) effect proposed in Ref. 14. The effective theory of this model is given by the Chern-Simons topological field theory in 4+1 dimensions.<sup>15</sup> The quantum spin Hall (QSH) effect has been proposed in (2+1)-dimensional TRI quantum models.<sup>16,17</sup> The QSH insulator state has a gap for all bulk excitations but has topologically protected gapless edge states, where opposite spin states counterpropagate.<sup>16,18,19</sup> Recently the QSH state has been theoretically predicted<sup>20</sup> and experimentally observed in HgTe quantum wells.<sup>21</sup> TRI topological insulators have also been classified in 3+1 dimensions.<sup>22-24</sup> These three-dimensional (3D) states all carry spin Hall current in the insulating state.<sup>25</sup>

The topological properties of the (4+1)-dimensional TRI insulator can be described by the second Chern number defined over four-dimensional momentum space. On the other hand, TRI insulators in 2+1 and 3+1 dimensions are de-

scribed by a  $Z_2$  topological invariant defined over momentum space.<sup>16,22–24,26–30</sup> In the presence of interactions and disorder, the momentum space  $Z_2$  invariant is not well defined: however, one can define a more general  $Z_2$  topological invariant in terms of spin-charge separation associated with a  $\pi$  flux.<sup>31,32</sup> One open question in this field concerns the relationship between the classification of the (4+1)-dimensional TRI insulator by the second Chern number and the classification of the (3+1)- and (2+1)-dimensional TRI insulators by the  $Z_2$  number.

The effective theory of the (4+1)-dimensional TRI insulator is given by the topological Chern-Simons field theory.<sup>15,33</sup> While the (2+1)-dimensional Chern-Simons theory describes a linear topological response to an external  $U(1)$  gauge field,<sup>7,8</sup> the (4+1)-dimensional Chern-Simons theory describes a nonlinear topological response to an external  $U(1)$  gauge field. The key outstanding theoretical problem in this field is the search for the topological field theory describing the TRI insulators in 2+1 and 3+1 dimensions, from which all measurable topological effects can be derived.

In this paper, we solve this outstanding problem by constructing topological field theories for the (2+1)- and (3+1)-dimensional TRI insulators using the procedure of dimensional reduction. We show that the (4+1)-dimensional topological insulator is the fundamental state from which all lower dimensional TRI insulators can be derived. This procedure is analogous to the dimensional reduction from the (2+1)-dimensional TRB topological insulator to the (1+1)-dimensional insulators. There is a deep reason why the fundamental TRB topological insulator exists in 2+1 dimensions, while the fundamental TRI topological insulator exists in 4+1 dimensions. The reason goes back to the von Neumann-Wigner classification<sup>34</sup> of level crossings in TRB unitary quantum systems and the TRI symplectic quantum systems. Generically three parameters need to be tuned to obtain a level crossing in a TRB unitary system, while five parameters need to be tuned to obtain a level crossing in a TRI symplectic system. These level crossing singularities give rise to the nontrivial topological curvatures on the 2D and 4D parameter surfaces which enclose the singularities. Fundamental topological insulators are obtained in space dimensions where all these parameters are momentum variables. Once the fundamental TRI topological insulator is identified in 4+1 dimensions, the lower dimensional versions of TRI topological insulators can be easily obtained by dimensional reduction. In this procedure, one or two momentum variables of the (4+1)-dimensional topological insulator are replaced by adiabatic parameters or fields, and the (4+1)-dimensional Chern-Simons topological field theory is reduced to topological field theories involving both the external  $U(1)$  gauge field and the adiabatic fields. For the 3+1 TRI insulators, the topological field theory is given by that of the “axion Lagrangian,” or the (3+1)-dimensional  $\theta$  vacuum term, familiar in the context of quantum chromodynamics (QCD), where the adiabatic field plays the role of the axion field or the  $\theta$  angle. From these topological field theories, all physically measurable topological effects of the (3+1)- and the (2+1)-dimensional TRI insulators can be derived. We predict a number of topological effects in this pa-

per, the most striking of which is the topological magnetoelectric (TME) effect, where an electric field induces a magnetic field in the same direction, with a universal constant of proportionality quantized in odd multiples of the fine-structure constant  $\alpha=e^2/\hbar c$ . We also present an experimental proposal to measure this effect in terms of Faraday rotation. Our dimensional reduction procedure also naturally produces the  $Z_2$  classification of the (3+1)- and the (2+1)-dimensional TRI topological insulators in terms of the integer second Chern class of the (4+1)-dimensional TRI topological insulators.

The remaining parts of the paper are organized as follows. In Sec. II we review the physical consequences of the first Chern number, namely, the (2+1)-dimensional QH effect and (1+1)-dimensional fractional charge and topological pumping effects. We begin with the (2+1)-dimensional time-reversal breaking insulators and study the topological transport properties. We then present a dimensional reduction procedure that allows us to consider related topological phenomena in (1+1)-dimensions and (0+1)-dimensions. Subsequently, we define a  $Z_2$  classification of these lower dimensional descendants which relies on the presence of a discrete particle-hole symmetry. This will serve as a review and a warm-up exercise for the more complicated phenomena we consider in the later sections. In Secs. III–V we discuss consequences of a nontrivial second Chern number beginning with a parent (4+1)-dimensional topological insulator in Sec. III. In Secs. IV and V we continue studying the consequences of the second Chern number but in the physically realistic (3+1)-dimensional and (2+1)-dimensional models which are the descendants of the initial (4+1)-dimensional system. We present effective actions describing all of the physical systems and their responses to applied electromagnetic fields. This provides the first effective-field theory for the TRI topological insulators in (3+1)-dimensions and (2+1)-dimensions. For these two descendants of the (4+1)-dimensional theory, we show that the  $Z_2$  classification of the descendants is obtained from the second Chern number classification of the parent TRI insulator. Finally, in Sec. VI we unify all of the results into families of topological effective actions defined in a phase space formalism. From this we construct a family tree of all topological insulators, some of which are only defined in higher dimensions, and with topological  $Z_2$  classifications which repeat every eight dimensions.

This paper contains many results on topological insulators, but it can also be read by advanced students as a pedagogical and self-contained introduction of topology applied to condensed matter physics. Physical models are presented in the familiar tight-binding forms, and all topological results can be derived by exact and explicit calculations, using techniques such as response theory already familiar in condensed matter physics. During the course of reading this paper, we suggest the readers to consult Appendix A which covers all of our conventions.

## II. TRB TOPOLOGICAL INSULATORS IN 2+1 DIMENSIONS AND DIMENSIONAL REDUCTION

In this section, we review the physics of the TRB topological insulators in 2+1 dimensions. We shall use the ex-

ample of a translationally invariant tight-binding model<sup>35</sup> which realizes the QH effect without Landau levels. We discuss the procedure of dimensional reduction, from which all topological effects of the (1+1)-dimensional insulators can be obtained. This section serves as a simple pedagogical example for the more complex case of the TRI insulators presented in Sec. III and IV.

### A. First Chern number and topological response function in (2+1)-dimensions

In general, the tight-binding Hamiltonian of a (2+1)-dimensional band insulator can be expressed as

$$H = \sum_{m,n;\alpha,\beta} c_{m\alpha}^\dagger h_{mn}^{\alpha\beta} c_{n\beta} \quad (1)$$

with  $m, n$  as the lattice sites and  $\alpha, \beta = 1, 2, \dots, N$  as the band indices for a  $N$ -band system. With translation symmetry  $h_{mn}^{\alpha\beta} = h^{\alpha\beta}(\vec{r}_m - \vec{r}_n)$ , the Hamiltonian can be diagonalized in a Bloch wave function basis,

$$H = \sum_{\mathbf{k}} c_{\mathbf{k}\alpha}^\dagger h^{\alpha\beta}(\mathbf{k}) c_{\mathbf{k}\beta}. \quad (2)$$

The minimal coupling to an external electromagnetic field is given by  $h_{mn}^{\alpha\beta} \rightarrow h_{mn}^{\alpha\beta} e^{iA_{mn}}$ , where  $A_{mn}$  is a gauge potential defined on a lattice link with sites  $m, n$  at the end. To linear order, the Hamiltonian coupled to the electromagnetic field is obtained as

$$H \approx \sum_{\mathbf{k}} c_{\mathbf{k}}^\dagger h(\mathbf{k}) c_{\mathbf{k}} + \sum_{\mathbf{k}, \mathbf{q}} A^i(-\mathbf{q}) c_{\mathbf{k}+\mathbf{q}/2}^\dagger \frac{\partial h(\mathbf{k})}{\partial k_i} c_{\mathbf{k}-\mathbf{q}/2}$$

with the band indices omitted. The dc response of the system to external field  $A^i(\mathbf{q})$  can be obtained by the standard Kubo formula,

$$\sigma_{ij} = \lim_{\omega \rightarrow 0} \frac{i}{\omega} Q_{ij}(\omega + i\delta),$$

$$Q_{ij}(i\nu_m) = \frac{1}{\Omega \beta_{\mathbf{k},n}} \sum_{\mathbf{k}} \text{tr} \{ J_i(\mathbf{k}) G[\mathbf{k}, i(\omega_n + \nu_m)] \cdot J_j(\mathbf{k}) G(\mathbf{k}, i\omega_n) \}, \quad (3)$$

with the dc  $J_i(\mathbf{k}) = \partial h(\mathbf{k}) / \partial k_i$ ,  $i, j = x, y$ , Green's function  $G(\mathbf{k}, i\omega_n) = [i\omega_n - h(\mathbf{k})]^{-1}$ , and  $\Omega$  as the area of the system. When the system is a band insulator with  $M$  fully occupied bands, the longitudinal conductance vanishes, i.e.,  $\sigma_{xx} = 0$ , as expected, while  $\sigma_{xy}$  has the form shown in Ref. 5,

$$\sigma_{xy} = \frac{e^2}{h} \frac{1}{2\pi} \int dk_x \int dk_y f_{xy}(\mathbf{k}), \quad (4)$$

with

$$f_{xy}(\mathbf{k}) = \frac{\partial a_y(\mathbf{k})}{\partial k_x} - \frac{\partial a_x(\mathbf{k})}{\partial k_y},$$

$$a_i(\mathbf{k}) = -i \sum_{\alpha \in \text{occ}} \langle \alpha \mathbf{k} | \frac{\partial}{\partial k_i} | \alpha \mathbf{k} \rangle, \quad i = x, y.$$

Physically,  $a_i(\mathbf{k})$  is the  $U(1)$  component of Berry's phase gauge field (adiabatic connection) in momentum space. The quantization of the first Chern number,

$$C_1 = \frac{1}{2\pi} \int dk_x \int dk_y f_{xy}(\mathbf{k}) \in \mathbb{Z}, \quad (5)$$

is satisfied for any continuous states  $|\alpha \mathbf{k}\rangle$  defined on the BZ.

Due to charge conservation, the QH response  $j_i = \sigma_H \epsilon^{ij} E_j$  also induces another response equation,

$$j_i = \sigma_H \epsilon^{ij} E_j \Rightarrow \frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{j} = -\sigma_H \nabla \times \mathbf{E} = \sigma_H \frac{\partial B}{\partial t} \quad (6)$$

$$\Rightarrow \rho(B) - \rho_0 = \sigma_H B, \quad (7)$$

where  $\rho_0 = \rho(B=0)$  is the charge density in the ground state. Equations (6) and (7) can be combined together in a covariant way,

$$j^\mu = \frac{C_1}{2\pi} \epsilon^{\mu\nu\tau} \partial_\nu A_\tau, \quad (8)$$

where  $\mu, \nu, \tau = 0, 1, 2$  are temporal and spatial indices. Here and below we will take the units  $e = \hbar = 1$  so that  $e^2/h = 1/2\pi$ .

The response [Eq. (8)] can be described by the topological Chern-Simons field theory of the external field  $A_\mu$ ,

$$S_{\text{eff}} = \frac{C_1}{4\pi} \int d^2x \int dt A_\mu \epsilon^{\mu\nu\tau} \partial_\nu A_\tau, \quad (9)$$

in the sense that  $\delta S_{\text{eff}} / \delta A_\mu = j^\mu$  recovers the response [Eq. (8)]. Such an effective action is topologically invariant, in agreement with the topological nature of the first Chern number. All topological responses of the QH state are contained in the Chern-Simons theory.<sup>8</sup>

### B. Example: Two band models

To make the physical picture clearer, the simplest case of a two-band model can be studied as an example.<sup>35</sup> The Hamiltonian of a two-band model can be generally written as

$$h(\mathbf{k}) = \sum_{a=1}^3 d_a(\mathbf{k}) \sigma^a + \epsilon(\mathbf{k}) \mathbb{I}, \quad (10)$$

where  $\mathbb{I}$  is the  $2 \times 2$  identity matrix and  $\sigma^a$  are the three Pauli matrices. Here we assume that the  $\sigma^a$  represent a spin or pseudospin degree of freedom. If it is a real spin then the  $\sigma^a$  are thus odd under time reversal. If the  $d_a(\mathbf{k})$  are odd in  $\mathbf{k}$  then the Hamiltonian is time-reversal invariant. However, if any of the  $d_a$  contain a constant term then the model has explicit time-reversal symmetry breaking. If the  $\sigma^a$  is a pseudospin then one has to be more careful. Since, in this case,  $\mathcal{T}^2 = 1$  then only  $\sigma^y$  is odd under time reversal (because it is imaginary) while  $\sigma^x, \sigma^z$  are even. The identity matrix is even under time reversal and  $\epsilon(\mathbf{k})$  must be even in  $\mathbf{k}$  to preserve

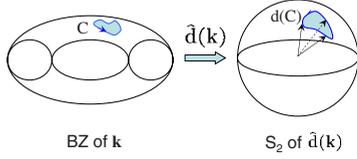


FIG. 1. (Color online) Illustration of Berry's phase curvature in a two-band model. Berry's phase  $\oint_C \mathbf{A} \cdot d\mathbf{r}$  around a path  $C$  in the BZ is half of the solid angle subtended by the image path  $d(C)$  on the sphere  $S_2$ .

time reversal. The energy spectrum is easily obtained:  $E_{\pm}(\mathbf{k}) = \varepsilon(\mathbf{k}) \pm \sqrt{\sum_a d_a^2(\mathbf{k})}$ . When  $\sum_a d_a^2(\mathbf{k}) > 0$  for all  $\mathbf{k}$  in the BZ, the two bands never touch each other. If we also require that  $\max_{\mathbf{k}}[E_-(\mathbf{k})] < \min_{\mathbf{k}}[E_+(\mathbf{k})]$ , so that the gap is not closed indirectly, then a gap always exists between the two bands of the system. In the single particle Hamiltonian  $h(\mathbf{k})$ , the vector  $\mathbf{d}(\mathbf{k})$  acts as a "Zeeman field" applied to a "pseudospin"  $\sigma_i$  of a two level system. The occupied band satisfies  $[\mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}]|-\mathbf{k}\rangle = -|\mathbf{d}(\mathbf{k})\rangle|-\mathbf{k}\rangle$ , which thus corresponds to the spinor with spin polarization in the  $-\mathbf{d}(\mathbf{k})$  direction. Thus Berry's phase gained by  $|-\mathbf{k}\rangle$  during an adiabatic evolution along some path  $C$  in  $\mathbf{k}$  space is equal to Berry's phase a spin-1/2 particle gains during the adiabatic rotation of the magnetic field along the path  $\mathbf{d}(C)$ . This is known to be half of the solid angle subtended by  $\mathbf{d}(C)$ , as shown in Fig. 1. Consequently, the first Chern number  $C_1$  is determined by the winding number of  $\mathbf{d}(\mathbf{k})$  around the origin,<sup>35,36</sup>

$$C_1 = \frac{1}{4\pi} \int dk_x \int dk_y \hat{\mathbf{d}} \cdot \frac{\partial \hat{\mathbf{d}}}{\partial k_x} \times \frac{\partial \hat{\mathbf{d}}}{\partial k_y}. \quad (11)$$

From the response equations we know that a nonzero  $C_1$  implies a quantized Hall response. The Hall effect can only occur in a system with time-reversal symmetry breaking so if  $C_1 \neq 0$  then time-reversal symmetry is broken. Historically, the first example of such a two-band model with a nonzero Chern number was a honeycomb lattice model with imaginary next-nearest-neighbor hopping proposed by Haldane.<sup>37</sup>

To be concrete, we shall study a particular two band model introduced in Ref. 35, which is given by

$$h(\mathbf{k}) = (\sin k_x)\sigma_x + (\sin k_y)\sigma_y + (m + \cos k_x + \cos k_y)\sigma_z. \quad (12)$$

This Hamiltonian corresponds to form (10) with  $\varepsilon(\mathbf{k}) \equiv 0$  and  $d(\mathbf{k}) = (\sin k_x, \sin k_y, m + \cos k_x + \cos k_y)$ . The Chern number of this system is<sup>35</sup>

$$C_1 = \begin{cases} 1 & \text{for } 0 < m < 2 \\ -1 & \text{for } -2 < m < 0 \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

In the continuum limit for  $m$  close to  $-2$ , this model reduces to the (2+1)-dimensional massive Dirac Hamiltonian,

$$h(\mathbf{k}) = k_x\sigma_x + k_y\sigma_y + (m+2)\sigma_z = \begin{pmatrix} m+2 & k_x - ik_y \\ k_x + ik_y & -m-2 \end{pmatrix}.$$

In a real space, this model can be expressed in tight-binding form as

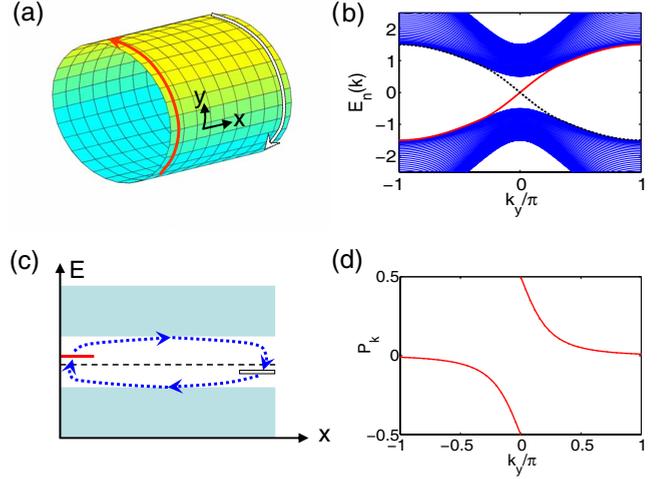


FIG. 2. (Color online) (a) Illustration of a square lattice with cylindrical geometry and the chiral edge states on the boundary. The definitions of  $x$  and  $y$  axes are also shown. (b) One-dimensional energy spectrum of the model in Eq. (12) with  $m = -1.5$ . The red solid and black dashed lines stand for the left and right moving edge states, respectively, while all other blue lines are bulk energy levels. (c) Illustration of the edge states evolution for  $k_y = 0 \rightarrow 2\pi$ . The arrow shows the motion of end states in the space of center-of-mass position versus energy. (d) Polarization of the one-dimensional system versus  $k_y$  (see text).

$$H = \sum_n \left[ c_n^\dagger \frac{\sigma_z - i\sigma_x}{2} c_{n+\hat{x}} + c_n^\dagger \frac{\sigma_z - i\sigma_y}{2} c_{n+\hat{y}} + \text{H.c.} \right] + m \sum_n c_n^\dagger \sigma_z c_n. \quad (14)$$

Physically, such a model describes the quantum anomalous Hall effect realized with both strong spin-orbit coupling ( $\sigma_x$  and  $\sigma_y$  terms) and ferromagnetic polarization ( $\sigma_z$  term). Initially this model was introduced for its simplicity in Ref. 35; however, recently, it was shown that it can be physically realized in  $\text{Hg}_{1-x}\text{Mn}_x\text{Te}/\text{Cd}_{1-x}\text{Mn}_x\text{Te}$  quantum wells with a proper amount of Mn spin polarization.<sup>38</sup>

### C. Dimensional reduction

To see how topological effects of (1+1)-dimensional insulators can be derived from the first Chern number and the QH effect through the procedure of dimensional reduction, we start by studying the QH system on a cylinder. An essential consequence of the nontrivial topology in the QH system is the existence of chiral edge states. For the simplest case with the first Chern number  $C_1 = 1$ , there is one branch of chiral fermions on each boundary.<sup>39</sup> These edge states can be solved for explicitly by diagonalizing Hamiltonian (14) in a cylindrical geometry, that is, with periodic boundary conditions in the  $y$  direction and open boundary conditions in the  $x$  direction, as shown in Fig. 2(a). Note that with this choice  $k_y$  is still a good quantum number. By defining the partial Fourier transformation,

$$c_{k_y\alpha}(x) = \frac{1}{\sqrt{L_y}} \sum_y c_\alpha(x,y) e^{ik_y y},$$

with  $(x,y)$  as the coordinates of square lattice sites, the Hamiltonian can be rewritten as

$$H = \sum_{k_y,x} \left[ c_{k_y}^\dagger(x) \frac{\sigma_z - i\sigma_x}{2} c_{k_y}(x+1) + \text{H.c.} \right] + \sum_{k_y,x} c_{k_y}^\dagger(x) [\sin k_y \sigma_y + (m + \cos k_y) \sigma_z] c_{k_y}(x) \equiv \sum_{k_y} H_{1D}(k_y). \quad (15)$$

In this way, the 2D system can be treated as  $L_y$  independent 1D tight-binding chains, where  $L_y$  is the period of the lattice in the  $y$  direction. The eigenvalues of the 1D Hamiltonian  $H_{1D}(k_y)$  can be obtained numerically for each  $k_y$ , as shown in Fig. 2(b). An important property of the spectrum is the presence of edge states, which lie in the bulk energy gap and are spatially localized at the two boundaries:  $x=0, L_x$ . The chiral nature of the edge states can be seen from their energy spectrum. From Fig. 2(b) we can see that the velocity  $v = \partial E / \partial k$  is always positive for the left edge state and negative for the right one. The QH effect can be easily understood in this edge state picture by Laughlin's gauge argument.<sup>3</sup> Consider a constant electric field  $E_y$  in the  $y$  direction, which can be chosen as

$$A_y = -E_y t, \quad A_x = 0.$$

The Hamiltonian is written as  $H = \sum_{k_y} H_{1D}(k_y + A_y)$  and the current along the  $x$  direction is given by

$$J_x = \sum_{k_y} J_x(k_y) \quad (16)$$

with  $J_x(k_y)$  as the current of the 1D system. In this way, the Hall response of the 2D system is determined by the current response of the parametrized 1D systems  $H_{1D}[q(t)]$  to the temporal change in the parameter  $q(t) = k_y + A_y(t)$ . The gauge vector  $A_y$  corresponds to a flux  $\Phi = A_y L_y$  threading the cylinder. During a time period  $0 \leq t \leq 2\pi / L_y E_y$ , the flux changes from 0 to  $2\pi$ . The charge that flows through the system during this time is given by

$$\Delta Q = \int_0^{\Delta t} dt \sum_{k_y} J_x(k_y) \equiv \sum_{k_y} \Delta P_x(k_y) \Big|_0^{\Delta t} \quad (17)$$

with  $\Delta t = 2\pi / L_y E_y$ . In the second equality we use the relation between the current and charge polarization  $P_x(k_y)$  of the 1D systems,  $J_x(k_y) = dP_x(k_y) / dt$ . In the adiabatic limit, the 1D system stays in the ground state of  $H_{1D}[q(t)]$ , so that the change in polarization  $\Delta P_x(k_y)$  is given by  $\Delta P_x(k_y) = P_x(k_y - 2\pi / L_y) - P_x(k_y)$ . Thus in the  $L_y \rightarrow \infty$  limit  $\Delta Q$  can be written as

$$\Delta Q = - \oint_0^{2\pi} dk_y \frac{\partial P_x(k_y)}{\partial k_y}. \quad (18)$$

Therefore, the charge flow due to the Hall current generated by the flux through the cylinder equals the charge flow through the one-dimensional system  $H_{1D}(k_y)$  when  $k_y$  is cycled adiabatically from 0 to  $2\pi$ . From the QH response we

know  $\Delta Q = \sigma_H \Delta t E_y L_y = 2\pi \sigma_H$  is quantized as an integer, which is easy to understand in the 1D picture. During the adiabatic change in  $k_y$  from 0 to  $2\pi$ , the energy and position of the edge states will change, as shown in Fig. 2(c). Since the edge state energy is always increasing (decreasing) with  $k_y$  for a state on the left (right) boundary, the charge is always "pumped" to the left (right) for the half-filled system, which leads to  $\Delta Q = -1$  for each cycle. This quantization can also be explicitly shown by calculating the polarization  $P_x(k_y)$ , as shown in Fig. 2(d), where the jump of  $P_x$  by one leads to  $\Delta Q = -1$ . In summary, we have shown that the QH effect in the tight-binding model of Eq. (12) can be mapped to an adiabatic pumping effect<sup>12</sup> by diagonalizing the system in one direction and mapping the momentum  $k$  to a parameter.

Such a dimensional reduction procedure is not restricted to specific models and can be generalized to any 2D insulators. For any insulator with Hamiltonian (2), we can define the corresponding 1D systems,

$$H_{1D}(\theta) = \sum_{k_x} c_{k_x}^\dagger \theta h(k_x, \theta) c_{k_x, \theta}, \quad (19)$$

in which  $\theta$  replaces the  $y$ -direction momentum  $k_y$  and effectively takes the place of  $q(t)$ . When  $\theta$  is time dependent, the current response can be obtained by a similar Kubo formula to Eq. (3), except that the summation over all  $(k_x, k_y)$  is replaced by that over only  $k_x$ . More explicitly, such a linear response is defined as

$$J_x(\theta) = G(\theta) \frac{d\theta}{dt},$$

$$G(\theta) = \lim_{\omega \rightarrow 0} \frac{i}{\omega} Q(\omega + i\delta; \theta),$$

$$Q(i\omega_n; \theta) = - \sum_{k_x, i\nu_m} \text{tr} \left\{ J_x(k_x; \theta) G_{1D}[k_x, i(\nu_m + \omega_n); \theta] \times \frac{\partial h(k_x; \theta)}{\partial \theta} G_{1D}(k_x, i\omega_n; \theta) \right\} \frac{1}{L_x \beta}. \quad (20)$$

Similar to Eq. (4) of the 2D case, the response coefficient  $G(k)$  can be expressed in terms of Berry's phase gauge field as

$$G(\theta) = - \oint \frac{dk_x}{2\pi} f_{x\theta}(k_x, \theta) = \oint \frac{dk_x}{2\pi} \left( \frac{\partial a_x}{\partial \theta} - \frac{\partial a_\theta}{\partial k_x} \right) \quad (21)$$

with the sum rule

$$\int G(\theta) d\theta = C_1 \in \mathbb{Z}. \quad (22)$$

If we choose a proper gauge so that  $a_\theta$  is always single valued, the expression of  $G(\theta)$  can be further simplified to

$$G(\theta) = \frac{\partial}{\partial \theta} \left( \oint \frac{dk_x}{2\pi} a_x(k_x, \theta) \right) \equiv \frac{\partial P(\theta)}{\partial \theta}. \quad (23)$$

Physically, the loop integral

$$P(\theta) = \oint dk_x a_x / 2\pi \quad (24)$$

is nothing but the *charge polarization* of the 1D system,<sup>10,11</sup> and the response equation [Eq. (20)] simply becomes  $J_x = \partial P / \partial t$ . Since the polarization  $P$  is defined as the shift of the electron center-of-mass position away from the lattice sites, it is only well-defined mod. 1. Consequently, the change  $\Delta P = P(\theta=2\pi) - P(\theta=0)$  through a period of adiabatic evolution is an integer equal to  $-C_1$  and corresponds to the charge pumped through the system. Such a relation between quantized pumping and the first Chern number was shown by Thouless.<sup>12</sup>

Similar to the QH case, the current response can lead to a charge-density response, which can be determined by the charge conservation condition. When the parameter  $\theta$  has a smooth spatial dependence  $\theta = \theta(x, t)$ , response equation [Eq. (20)] still holds. From the continuity equation we obtain

$$\frac{\partial \rho}{\partial t} = -\frac{\partial J_x}{\partial x} = -\frac{\partial^2 P(\theta)}{\partial x \partial t} \Rightarrow \rho = -\frac{\partial P(\theta)}{\partial x}, \quad (25)$$

in which  $\rho$  is defined with respect to the background charge. Similar to Eq. (8), the density and current response can be written together as

$$j_\mu = -\epsilon_{\mu\nu} \frac{\partial P[\theta(x, t)]}{\partial x_\nu}, \quad (26)$$

where  $\mu, \nu=0, 1$  are time and space. It should be noted that only differentiation with respect to  $x, t$  appears in Eq. (26). This means, as expected, the current and density response of the system do not depend on the parametrization. In general, when the Hamiltonian has smooth space and time dependence, the single particle Hamiltonian  $h(k)$  becomes  $h(k, x, t) \equiv h[k, \theta(x, t)]$ , which has the eigenstates  $|\alpha; k, x, t\rangle$  with  $\alpha$  as the band index. Then relabeling  $t, x, k$  as  $q_A$ ,  $A=0, 1, 2$ , we can define the *phase space* Berry's phase gauge field,

$$\mathcal{A}_A = -i \sum_\alpha \langle \alpha; q_A | \frac{\partial}{\partial q_A} | \alpha; q_A \rangle, \quad (27)$$

$$\mathcal{F}_{AB} = \partial_A \mathcal{A}_B - \partial_B \mathcal{A}_A,$$

and the *phase space current*,

$$j_A^P = -\frac{1}{4\pi} \epsilon_{ABC} \mathcal{F}_{BC}. \quad (28)$$

The physical current is obtained by integration over the wave-vector manifold,

$$j_\mu = \int dk j_\mu^P = -\int \frac{dk}{2\pi} \epsilon^{\mu 2\nu} \mathcal{F}_{2\nu}, \quad (29)$$

where  $\mu, \nu=0, 1$ . This recovers Eq. (26). Note that we could have also looked at the component  $j_k = \int dk j_k^P$  but this current does not have a physical interpretation.

Before moving to the next topic, we would like to apply this formalism to the case of the Dirac model, which reproduces the well-known result of fractional charge in the Su-Schrieffer-Heeger (SSH) model<sup>9</sup> or equivalently the Jackiw-Rebbi model.<sup>40</sup> To see this, consider the following slightly different version of tight-binding model (12):

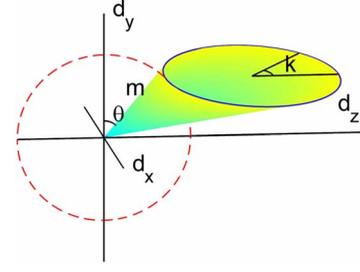


FIG. 3. (Color online) Illustration of the  $\mathbf{d}(k, \theta)$  vector for the 1D Dirac model (30). The horizontal blue solid circle shows the orbit of  $\mathbf{d}(k)$  vector in the 3D space for  $k \in [0, 2\pi)$  with  $\theta$  fixed. The red dashed circle shows the track of the blue solid circle under the variation in  $\theta$ . The cone shows the solid angle  $\Omega(\theta)$  surrounded by the  $\mathbf{d}(k)$  curve, which is  $4\pi$  times the polarization  $P(\theta)$ .

$$h(k, \theta) = \sin k \sigma_x + (\cos k - 1) \sigma_z + m(\sin \theta \sigma_y + \cos \theta \sigma_z) \quad (30)$$

with  $m > 0$ . In the limit  $m \ll 1$ , the Hamiltonian has the continuum limit  $h(k, \theta) \approx k \sigma_x + m(\sin \theta \sigma_y + \cos \theta \sigma_z)$ , which is the continuum Dirac model in (1+1) dimensions, with a real mass  $m \cos \theta$  and an imaginary mass  $m \sin \theta$ . As discussed in Sec. II B, the polarization  $\oint dk_x a_x / 2\pi$  is determined by the solid angle subtended by the curve  $\mathbf{d}(k) = (\sin k, m \sin \theta, m \cos \theta + \cos k - 1)$ , as shown in Fig. 3. In the limit  $m \ll 1$  one can show that the solid angle  $\Omega(\theta) = 2\theta$  so that  $P(\theta) \approx \theta / 2\pi$ , in which case Eq. (26) reproduces the Goldstone-Wilczek formula,<sup>13</sup>

$$j_\mu = -\epsilon_{\mu\nu} \partial_\nu \theta. \quad (31)$$

Specifically, a charge  $Q = -\int_{-\infty}^{\infty} (d\theta/dx)(dx/2\pi) = -[\theta(+\infty) - \theta(-\infty)]/2\pi$  is carried by a domain wall of the  $\theta$  field. In particular, for an antiphase domain wall,  $\theta(+\infty) - \theta(-\infty) = \pi$ , we obtain fractional charge  $q = 1/2$ . Our phase space formula (28) is a new result, and it provides a generalization of the Goldstone-Wilczek formula to the most general one-dimensional insulator.

#### D. $Z_2$ classification of particle-hole symmetric insulators in (1+1) dimensions

In the last section, we have shown how the first Chern number of Berry's phase gauge field appears in an adiabatic pumping effect and the domain wall charge of one-dimensional insulators. In these cases, an adiabatic spatial or temporal variation in the single-particle Hamiltonian, through its parametric dependence on  $\theta(x, t)$ , is required to define the Chern number. In other words, the first Chern number is defined for a *parametrized family* of Hamiltonians  $h(k, x, t)$  rather than for a single 1D Hamiltonian  $h(k)$ . In this section, we will show a different application of the first Chern number, in which a  $Z_2$  topological classification is obtained for particle-hole symmetric insulators in one dimensions. Such a relation between Chern number and  $Z_2$  topology can be easily generalized to the more interesting case of second Chern number, where a similar  $Z_2$  characterization is obtained for TRI insulators, as will be shown in Secs. IV C and V B.

For a one-dimensional tight-binding Hamiltonian  $H = \sum_{mn} c_{m\alpha}^\dagger h_{mn}^{\alpha\beta} c_{n\beta}$ , the particle-hole transformation is defined by  $c_{m\alpha} \rightarrow C^{\alpha\beta} c_{m\beta}^\dagger$ , where the *charge-conjugation matrix*  $C$  satisfies  $C^\dagger C = \mathbb{I}$  and  $C^* C = \mathbb{I}$ . Under periodic boundary conditions the symmetry requirement is

$$H = \sum_k c_k^\dagger h(k) c_k = \sum_k c_{-k} C^\dagger h(k) C c_{-k}^\dagger \Rightarrow C^\dagger h(-k) C = -h^T(k). \quad (32)$$

From Eq. (32) it is straightforward to see the symmetry of the energy spectrum: if  $E$  is an eigenvalue of  $h(0)$ , so is  $-E$ . Consequently, if the dimension of  $h(k)$  is odd, there must be at least one zero mode with  $E=0$ . Since the chemical potential is constrained to vanish by the traceless condition of  $h$ , such a particle-hole symmetric system cannot be gapped unless the dimension of  $h(k)$  is even. Since we are only interested in the classification of insulators, we will focus on the case with  $2N$  bands per lattice site.

Now consider two particle-hole symmetric insulators with Hamiltonians  $h_1(k)$  and  $h_2(k)$ . In general, a continuous interpolation  $h(k, \theta)$ ,  $\theta \in [0, \pi]$  between them can be defined so that

$$h(k, 0) = h_1(k), \quad h(k, \pi) = h_2(k). \quad (33)$$

Moreover, it is always possible to find a proper parametrization so that  $h(k, \theta)$  is gapped for all  $\theta \in [0, \pi]$ . In other words, the topological space of all 1D insulating Hamiltonians  $h(k, \theta)$  is connected, which is a consequence of the von Neumann–Wigner theorem.<sup>34</sup>

Suppose  $h(k, \theta)$  is such a ‘‘gapped interpolation’’ between  $h_1(k)$  and  $h_2(k)$ . In general,  $h(k, \theta)$  for  $\theta \in (0, \pi)$  does not necessarily satisfy the particle-hole symmetry. For  $\theta \in [\pi, 2\pi]$ , define

$$h(k, \theta) = -[C^{-1}h(-k, 2\pi - \theta)C]^T. \quad (34)$$

We choose this parametrization so that if we replaced  $\theta$  by a momentum wave vector then the corresponding higher dimensional Hamiltonian would be particle-hole symmetric. Due to the particle-hole symmetry of  $h(k, \theta=0)$  and  $h(k, \theta=\pi)$ ,  $h(k, \theta)$  is continuous for  $\theta \in [0, 2\pi]$  and  $h(k, 2\pi) = h(k, 0)$ . Consequently, the adiabatic evolution of  $\theta$  from 0 to  $2\pi$  defines a cycle of adiabatic pumping in  $h(k, \theta)$ , and a first Chern number can be defined in the  $(k, \theta)$  space. As discussed in Sec. II C, the Chern number  $C[h(k, \theta)]$  can be expressed as a winding number of the polarization,

$$C[h(k, \theta)] = \oint d\theta \frac{\partial P(\theta)}{\partial \theta},$$

$$P(\theta) = \oint \frac{dk}{2\pi} \sum_{E_\alpha(k) < 0} (-i) \langle k, \theta; \alpha | \partial_k | k, \theta; \alpha \rangle,$$

where the summation is carried out over the occupied bands. In general, two different parametrizations  $h(k, \theta)$  and  $h'(k, \theta)$  can lead to different Chern numbers  $C[h(k, \theta)] \neq C[h'(k, \theta)]$ . However, the symmetry constraint in Eq. (34) guarantees that the two Chern numbers always differ by an even integer:  $C[h(k, \theta)] - C[h'(k, \theta)] = 2n$ ,  $n \in \mathbb{Z}$ .

To prove this conclusion, we first study the behavior of  $P(\theta)$  under a particle-hole transformation. For an eigenstate  $|k, \theta; \alpha\rangle$  of the Hamiltonian  $h(k, \theta)$  with eigenvalue  $E_\alpha(k, \theta)$ , Eq. (34) leads to

$$h(-k, 2\pi - \theta) C |k, \theta; \alpha\rangle^* = -E_\alpha(k) C |k, \theta; \alpha\rangle^*, \quad (35)$$

in which  $|k, \theta; \alpha\rangle^*$  is the complex conjugate state:  $|k, \theta; \alpha\rangle^* = \sum_\beta \langle m, \beta | k, \theta; \alpha \rangle^* |m, \beta\rangle$ , where  $m, \beta$  are the position space lattice and orbital index, respectively. Thus  $C |k, \theta; \alpha\rangle^* = |-k, 2\pi - \theta; \bar{\alpha}\rangle$  is an eigenstate of  $h(-k, 2\pi - \theta)$  with energy  $E_{\bar{\alpha}}(-k, 2\pi - \theta) = -E_\alpha(k, \theta)$  and momentum  $-k$ . Such a mapping between eigenstates of  $h(k, \theta)$  and  $h(-k, 2\pi - \theta)$  is one to one. Thus

$$\begin{aligned} P(\theta) &= \oint \frac{dk}{2\pi} \sum_{E_\alpha(k) < 0} (-i) \langle k, \theta; \alpha | \partial_k | k, \theta; \alpha \rangle \\ &= \oint \frac{dk}{2\pi} \sum_{E_{\bar{\alpha}}(-k) > 0} (-i) \langle -k, 2\pi - \theta; \bar{\alpha} | \partial_k | -k, 2\pi - \theta; \bar{\alpha} \rangle^* \\ &\quad \cdot \partial_k | -k, 2\pi - \theta; \bar{\alpha} \rangle^* = -P(2\pi - \theta). \end{aligned} \quad (36)$$

Since  $P(\theta)$  is only well-defined mod. 1, equality (36) actually means  $P(\theta) + P(2\pi - \theta) = 0 \pmod{1}$  (we give a detailed proof in Appendix B). Consequently, for  $\theta=0$  or  $\pi$  we have  $2\pi - \theta = \theta \pmod{2\pi}$ , so that  $P(\theta) = 0$  or  $1/2$ . In other words, the polarization  $P$  is either 0 or  $1/2$  for any particle-hole symmetric insulator, which thus defines a classification of particle-hole symmetric insulators. If two systems have different  $P$  value, they cannot be adiabatically connected without breaking the particle-hole symmetry because  $P \pmod{1}$  is a continuous function during adiabatic deformation and a  $P$  value other than 0 and  $1/2$  breaks particle-hole symmetry. Though such an argument explains physically why a  $Z_2$  classification is defined for particle-hole symmetric system, it is not so rigorous. As discussed in the derivation from Eq. (21) to Eq. (23), the definition  $P(\theta) = \oint dka_k / 2\pi$  relies on a proper gauge choice. To avoid any gauge dependence, a more rigorous definition of the  $Z_2$  classification is shown below, which only involves the gauge-invariant variable  $\partial P(\theta) / \partial \theta$  and Chern number  $C_1$ .

To begin with, symmetry (36) leads to

$$\int_0^\pi dP(\theta) = \int_\pi^{2\pi} dP(\theta), \quad (37)$$

which is independent of gauge choice since only the change in  $P(\theta)$  is involved. This equation shows that the changes in polarization during the first half and the second half of the closed path  $\theta \in [0, 2\pi]$  are always the same.

Now consider two different parametrizations  $h(k, \theta)$  and  $h'(k, \theta)$ , satisfying  $h(k, 0) = h'(k, 0) = h_1(k)$ ,  $h(k, \pi) = h'(k, \pi) = h_2(k)$ . Denoting the polarizations  $P(\theta)$  and  $P'(\theta)$  corresponding to  $h(k, \theta)$  and  $h'(k, \theta)$ , respectively, the Chern number difference between  $h$  and  $h'$  is given by

$$C[h] - C[h'] = \int_0^{2\pi} d\theta \left( \frac{\partial P(\theta)}{\partial \theta} - \frac{\partial P'(\theta)}{\partial \theta} \right). \quad (38)$$

Define the new interpolations  $g_1(k, \theta)$  and  $g_2(k, \theta)$  as

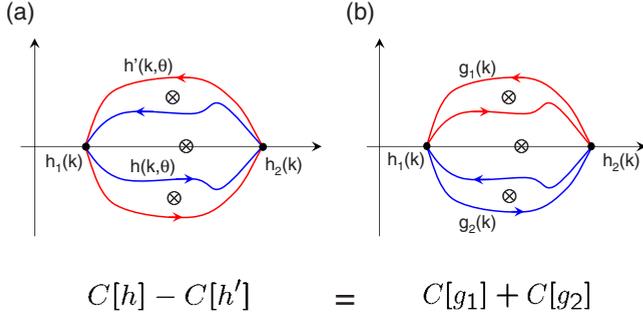


FIG. 4. (Color online) Illustration of the interpolation between two particle-hole symmetric Hamiltonians  $h_1(k)$  and  $h_2(k)$ .

$$g_1(k, \theta) = \begin{cases} h(k, \theta), & \theta \in [0, \pi] \\ h'(k, 2\pi - \theta), & \theta \in [\pi, 2\pi], \end{cases}$$

$$g_2(k, \theta) = \begin{cases} h'(k, 2\pi - \theta), & \theta \in [0, \pi] \\ h(k, \theta), & \theta \in [\pi, 2\pi], \end{cases} \quad (39)$$

where  $g_1(k, \theta)$  and  $g_2(k, \theta)$  are obtained by recombination of the two paths  $h(k, \theta)$  and  $h'(k, \theta)$ , as shown in Fig. 4. From the construction of  $g_1$  and  $g_2$ , it is straightforward to see that

$$C[g_1] = \int_0^\pi d\theta \left( \frac{\partial P(\theta)}{\partial \theta} - \frac{\partial P'(\theta)}{\partial \theta} \right),$$

$$C[g_2] = \int_\pi^{2\pi} d\theta \left( \frac{\partial P(\theta)}{\partial \theta} - \frac{\partial P'(\theta)}{\partial \theta} \right). \quad (40)$$

Thus  $C[h] - C[h'] = C[g_1] + C[g_2]$ . On the other hand, from Eq. (37) we know  $C[g_1] = C[g_2]$ , so that  $C[h] - C[h'] = 2C[g_1]$ . Since  $C[g_1] \in \mathbb{Z}$ , we obtain that  $C[h] - C[h']$  is even for any two interpolations  $h(k, \theta)$  and  $h'(k, \theta)$  between  $h_1(k)$  and  $h_2(k)$ . Intuitively, such a conclusion simply comes from the fact that the Chern numbers  $C[h]$  and  $C[h']$  can be different only if there are singularities between these two paths while the positions of the singularities in the parameter space are always symmetric under particle-hole symmetry, as shown in Fig. 4.

Based on the discussions above, we can define the *relative Chern parity* as

$$N_1[h_1(k), h_2(k)] = (-1)^{C[h(k, \theta)]}, \quad (41)$$

which is independent of the choice of interpolation  $h(k, \theta)$  but only determined by the Hamiltonians  $h_1(k)$ ,  $h_2(k)$ . Moreover, for any three particle-hole symmetric Hamiltonians  $h_1(k)$ ,  $h_2(k)$ ,  $h_3(k)$ , it is easy to prove that the Chern parity satisfies the following associative law:

$$N_1[h_1(k), h_2(k)] N_1[h_2(k), h_3(k)] = N_1[h_1(k), h_3(k)]. \quad (42)$$

Consequently,  $N_1[h_1(k), h_2(k)] = 1$  defines an *equivalence relation* between any two particle-hole symmetric Hamiltonians, which thus classifies all the particle-hole symmetric insulators into two classes. To define these two classes more explicitly, one can define a “vacuum” Hamiltonian as  $h_0(k) \equiv h_0$ , where  $h_0$  is an arbitrary matrix which does not depend on  $k$  and which satisfies the particle-hole symmetry con-

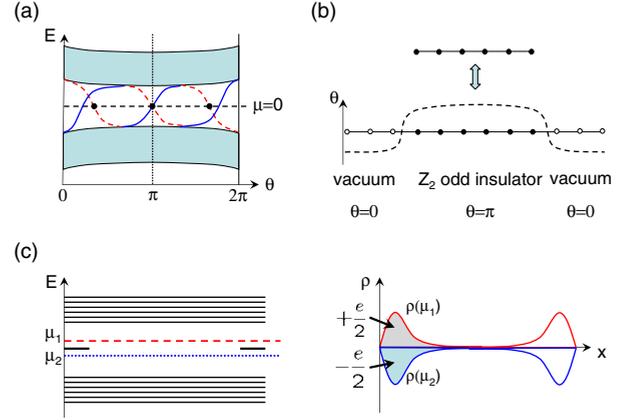


FIG. 5. (Color online) (a) Schematic energy spectrum of a parametrized Hamiltonian  $h_{mn}(\theta)$  with open boundary conditions. The red dashed (blue solid) lines indicate the left (right) end states. The  $\theta$  values with zero-energy *left* edge states are marked by the solid circles. (b) Illustration to show that the open boundary of a  $Z_2$  nontrivial insulator is equivalent to a domain wall between  $\theta = \pi$  (nontrivial) and  $\theta = 0$  (trivial vacuum). (c) Illustration of the charge-density distribution corresponding to two different chemical potentials  $\mu_1$  (red dashed line) and  $\mu_2$  (blue dotted line). The areas below the curves  $\rho(\mu_1)$  and  $\rho(\mu_2)$  are  $+e/2$  and  $-e/2$ , respectively, which show the half charge confined on the boundary.

straint  $C^\dagger h_0 C = -h_0^T$ . Thus  $h_0$  describes a totally local system, in which there is no hopping between different sites. Taking such a trivial system as a reference Hamiltonian, we can define  $N_1[h_0(k), h(k)] \equiv N_1[h(k)]$  as a  $Z_2$  topological quantum number of the Hamiltonian  $h(k)$ . All the Hamiltonians  $h(k)$  with  $N_1[h_0(k), h(k)] = 1$  are classified as  $Z_2$  trivial, while those with  $N_1[h_0(k), h(k)] = -1$  are considered as  $Z_2$  nontrivial. (Again, this classification does not depend on the choice of vacuum  $h_0$  since any two vacua are equivalent.)

Despite its abstract form, such a topological characterization has a direct physical consequence. For a  $Z_2$  nontrivial Hamiltonian  $h_1(k)$ , an interpolation  $h(k, \theta)$  can be defined so that  $h(k, 0) = h_0$ ,  $h(k, \pi) = h_1(k)$ , and the Chern number  $C[h(k, \theta)]$  is an odd integer. If we study the one-dimensional system  $h(k, \theta)$  with open boundary conditions, the tight-binding Hamiltonian can be rewritten in real space as

$$h_{mn}(\theta) = \frac{1}{\sqrt{L}} \sum_k e^{ik(x_m - x_n)} h(k, \theta), \quad \forall 1 \leq m, n \leq L.$$

As discussed in Sec. II C, there are midgap end states in the energy spectrum of  $h_{mn}(\theta)$  as a consequence of the nonzero Chern number. When the Chern number  $C[h(k, \theta)] = 2n - 1$ ,  $n \in \mathbb{Z}$ , there are values  $\theta_s^L \in [0, 2\pi)$ ,  $s = 1, 2, \dots, 2n - 1$ , for which the Hamiltonian  $h_{mn}(\theta_s)$  has zero energy localized states on the left end of the 1D system and the same number of  $\theta_s^R$  values where zero energy states are localized on the right end, as shown in Fig. 5. Due to the particle-hole symmetry between  $h_{mn}(\theta)$  and  $h_{mn}(2\pi - \theta)$ , zero levels always appear in pairs at  $\theta$  and  $2\pi - \theta$ . Consequently, when the Chern number is odd, there must be a zero level at  $\theta = 0$  or  $\theta = \pi$ . Since  $\theta = 0$  corresponds to a trivial insulator with flat bands and no end states, the localized zero mode has to ap-

pear at  $\theta = \pi$ . In other words, one zero energy localized state (or an odd number of such states) is confined at each open boundary of a  $Z_2$  nontrivial particle-hole symmetric insulator.

The existence of a zero level leads to an important physical consequence—a half charge on the boundary of the nontrivial insulator. In a periodic system when the chemical potential vanishes, the average electron density on each site is  $\bar{n}_m = \langle \sum_{\alpha} c_{m\alpha}^{\dagger} c_{m\alpha} \rangle = N$  when there are  $N$  bands filled. In an open boundary system, define  $\rho_m(\mu) = \langle \sum_{\alpha} c_{m\alpha}^{\dagger} c_{m\alpha} \rangle_{\mu} - N$  to be the density deviation with respect to  $N$  on each site. Then particle-hole symmetry leads to  $\rho_m(\mu) = -\rho_m(-\mu)$ . On the other hand, when  $\mu$  is in the bulk gap, the only difference between  $\mu$  and  $-\mu$  is the filling of the zero levels localized on each end  $|0L\rangle$  and  $|0R\rangle$ , so that

$$\lim_{\mu \rightarrow 0^+} [\rho_m(\mu) - \rho_m(-\mu)] = \sum_{\alpha} |\langle m\alpha | 0L \rangle|^2$$

for the sites  $m$  that are far away enough from the right boundary. Thus we have  $\sum_m \rho_m(\mu \rightarrow 0^+) = 1/2$  where the summation is done around the left boundary so that we do not pick up a contribution from the other end. In summary, a charge  $e/2$  ( $-e/2$ ) is localized on the boundary if the zero level is vacant (occupied), as shown in Fig. 5.

The existence of such a half charge can also be understood by viewing the open boundary of a topologically nontrivial insulator as a domain wall between the nontrivial insulator and the vacuum. By defining the interpolation  $h_{mn}(\theta)$ , such a domain wall is described by a spatial dependence of  $\theta$  with  $\theta(x \rightarrow +\infty) = \pi$ ,  $\theta(x \rightarrow -\infty) = 0$ . According to response formula (25), the charge carried by the domain wall is given by

$$Q_d = e \int_{-\infty}^{+\infty} dx \frac{\partial P[\theta(x)]}{\partial x} = e \int_0^{\pi} dP(\theta). \quad (43)$$

By using Eq. (37) we obtain

$$Q_d = \frac{e}{2} \int_0^{2\pi} dP(\theta) = \frac{e}{2} C[h(k, \theta)]. \quad (44)$$

It should be noted that an integer charge can always be added by changing the filling of local states, which means  $Q_d$  is only fixed mod  $e$ . Consequently,  $a \pm e/2$  charge is carried by the domain wall if and only if the Chern number is odd, i.e., when the insulator is nontrivial.

### E. $Z_2$ classification of (0+1)-dimensional particle-hole symmetric insulators

In the last section, we have shown how a  $Z_2$  classification of (1+1)-dimensional particle-hole symmetric insulators is defined by dimensional reduction from (2+1)-dimensional systems. Such a dimensional reduction can be repeated once more to study (0+1)-dimensional systems, that is, a single-site problem. In this section we will show that a  $Z_2$  classification of particle-hole symmetric Hamiltonians in (0+1)-dimensions is also obtained by dimensional reduction. Although such a classification by itself is not as interesting as the higher dimensional counterparts, it does provide a simplest example of the “dimensional reduction chain” (2

+1) dimensions  $\rightarrow$  (1+1) dimensions  $\rightarrow$  (0+1) dimensions, which can be later generalized to its higher-dimensional counterpart (4+1) dimensions  $\rightarrow$  (3+1) dimensions  $\rightarrow$  (2+1) dimensions. In other words, the  $Z_2$  classification of the (0+1)-dimensional particle-hole symmetric insulators can help us to understand the classification of (2+1)-dimensional TRI insulators as it is dimensionally reduced from the (4+1)-dimensional TRI insulator.

For a free single-site fermion system with Hamiltonian matrix  $h$ , the particle-hole symmetry is defined as

$$C^{\dagger} h C = -h^T. \quad (45)$$

Given any two particle-hole symmetric Hamiltonians  $h_1$  and  $h_2$ , we follow the same procedure as the last section and define a continuous interpolation  $h(\theta)$ ,  $\theta \in [0, 2\pi]$  satisfying

$$h(0) = h_1, \quad h(\pi) = h_2, \quad C^{\dagger} h(\theta) C = -h(2\pi - \theta)^T, \quad (46)$$

where  $h(\theta)$  is gapped for all  $\theta$ . The Hamiltonian  $h(\theta)$  is the dimensional reduction in a (1+1)-dimensional Hamiltonian  $h(k)$ , with the wave vector  $k$  replaced by the parameter  $\theta$ . Constraint (46) is identical to the particle-hole symmetry condition (32), so that  $h(\theta)$  corresponds to a particle-hole symmetric (1+1)-dimensional insulator. As shown in last section,  $h(\theta)$  is classified by the value of the “Chern parity”  $N_1[h(\theta)]$ . If  $N_1[h(\theta)] = -1$ , no continuous interpolation preserving particle-hole symmetry can be defined between  $h(\theta)$  and the vacuum Hamiltonian  $h(\theta) = h_0$ ,  $\forall \theta \in [0, 2\pi]$ . To obtain the classification of (0+1)-dimensional Hamiltonians, consider two different interpolations  $h(\theta)$  and  $h'(\theta)$  between  $h_1$  and  $h_2$ . According to associative law (42), we know  $N_1[h(\theta)]N_1[h'(\theta)] = N_1[h(\theta), h'(\theta)]$ , where  $N_1[h(\theta), h'(\theta)]$  is the relative Chern parity between two interpolations. In the following we will prove  $N_1[h(\theta), h'(\theta)] = 1$  for any two interpolations  $h$  and  $h'$  satisfying condition (46). As a result,  $N_1[h(\theta)]$  is independent of the choice of interpolation between  $h_1$  and  $h_2$ , so that  $N_0[h_1, h_2] \equiv N_1[h(\theta)]$  can be defined as a function of  $h_1$  and  $h_2$ . The  $Z_2$  quantity  $N_0$  defined for (0+1)-dimensional Hamiltonians plays exactly the same role as  $N_1[h(k), h'(k)]$  in the (1+1)-dimensional case, from which a  $Z_2$  classification can be defined.

To prove  $N_1[h(\theta), h'(\theta)] = 1$  for any two interpolations, first define a continuous deformation  $g(\theta, \varphi)$  between  $h(\theta)$  and  $h'(\theta)$ , which satisfies the conditions below:

$$g(\theta, \varphi = 0) = h(\theta), \quad g(\theta, \varphi = \pi) = h'(\theta),$$

$$g(0, \varphi) = h_1, \quad g(\pi, \varphi) = h_2,$$

$$C^{\dagger} g(\theta, \varphi) C = -g(2\pi - \theta, 2\pi - \varphi)^T. \quad (47)$$

From the discussions in last section it is easy to confirm that such a continuous interpolation is always possible, in which  $g(\theta, \varphi)$  is gapped for all  $\theta$  and  $\varphi$ . In the two-dimensional parameter space  $\theta, \varphi$  one can define the Berry phase gauge field and the first Chern number  $C_1[g(\theta, \varphi)]$ . By the definition of the Chern parity, we have  $N_1[h(\theta), h'(\theta)] = (-1)^{C_1[g(\theta, \varphi)]}$ . However, the parametrized Hamiltonian  $g(\theta, \varphi)$  can be viewed in two different ways: it not only defines an interpolation between  $h(\theta)$  and  $h'(\theta)$  but also defines an interpolation between  $g(0, \varphi) = h_1$  and  $g(\pi, \varphi) = h_2$ .

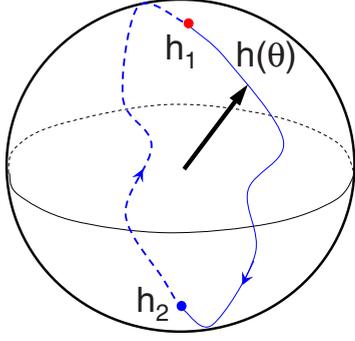


FIG. 6. (Color online) Illustration of the  $2 \times 2$  single-site Hamiltonians. Each point on the sphere represents a unit vector  $\hat{\mathbf{d}} = \vec{d}/|d|$ , and the north and south poles correspond to the particle-hole symmetric Hamiltonians  $h_{1,2} = \pm \sigma^3$ , respectively. The blue path shows an interpolation between  $h_1$  and  $h_2$  satisfying constraint (46), which always encloses a solid angle  $\Omega = 2\pi$ .

Since  $g(0, \varphi)$  and  $g(\pi, \varphi)$  are “vacuum Hamiltonians” without any  $\varphi$  dependence, they have trivial relative Chern parity, which means  $N_1[h(\theta), h'(\theta)] = N_1[g(0, \varphi), g(\pi, \varphi)] = N_1[h_1, h_2] = 1$ .

In conclusion, from the discussion above we have proved that any two interpolations  $h(\theta)$  and  $h'(\theta)$  belong to the same  $Z_2$  class, so that the Chern parity  $N_1[h(\theta)]$  only depends on the end points  $h_1$  and  $h_2$ . Consequently, the quantity  $N_0[h_1, h_2] \equiv N_1[h(\theta)]$  defines a relation between each pair of particle-hole symmetric Hamiltonians  $h_1$  and  $h_2$ . After picking a reference Hamiltonian  $h_0$ , one can *define* all the Hamiltonians with  $N_0[h_0, h] = 1$  as “trivial” and  $N_0[h_0, h] = -1$  as nontrivial. The main difference between this classification and the one for (1+1)-dimensional systems is that there is no natural choice of the reference Hamiltonian  $h_0$ . In other words, the names trivial and “nontrivial” only have relative meaning in the (0+1)-dimensional case. However, the classification is still meaningful in the sense that any two Hamiltonians with  $N_0[h_1, h_2] = -1$  cannot be adiabatically connected without breaking particle-hole symmetry. In other words, the manifold of single-site particle-hole symmetric Hamiltonians is disconnected, with at least two connected pieces.

As a simple example, we study  $2 \times 2$  Hamiltonians. A general  $2 \times 2$  single-site Hamiltonian can be decomposed as

$$h = d_0 \sigma^0 + \sum_{a=1}^3 d_a \sigma^a, \quad (48)$$

where  $\sigma^0 = \mathbb{I}$  and  $\sigma^{1,2,3}$  are the Pauli matrices. When  $C = \sigma^1$ , particle-hole symmetry requires  $C^\dagger h C = -h^T$ , from which we obtain  $d_0 = d_1 = d_2 = 0$ . Thus  $h = d_3 \sigma^3$ , in which  $d_3 \neq 0$  so as to make  $h$  gapped. Consequently, we can see that the two  $Z_2$  classes are simply  $d_3 > 0$  and  $d_3 < 0$ . When an adiabatic interpolation  $h(\theta) = d_0(\theta) \sigma^0 + \sum_a d_a(\theta) \sigma^a$  is defined from  $d_3 > 0$  to  $d_3 < 0$ , the spin vector  $\vec{d}(\theta)$  has to rotate from the north pole to the south pole and then return along the image path determined by particle-hole symmetry (46), as shown in Fig. 6. The topological quantum number  $N_0[h_1, h_2]$  is simply determined by Berry’s phase enclosed by the path  $d_a(\theta)$ , which

is  $\pi$  when  $h_1$  and  $h_2$  are on different poles and 0 otherwise. From this example we can understand the  $Z_2$  classification intuitively. In Sec. V B we show that the  $Z_2$  classification of (2+1)-dimensional TRI insulators—the class that corresponds to the QSH effect—is obtained as a direct analog of the (0+1)-dimensional case discussed above.

### III. SECOND CHERN NUMBER AND ITS PHYSICAL CONSEQUENCES

In this section, we shall generalize the classification of the (2+1)-dimensional TRB topological insulator in terms of the first Chern number and the (2+1)-dimensional Chern-Simons theory to the classification of the (4+1)-dimensional TRI topological insulator in terms of the second Chern number and the (4+1)-dimensional Chern-Simons theory. We then generalize the dimensional reduction chain (2+1) dimensions  $\rightarrow$  (1+1) dimensions  $\rightarrow$  (0+1) dimensions to the case of (4+1) dimensions  $\rightarrow$  (3+1) dimensions  $\rightarrow$  (2+1) dimensions for TRI insulators. Many novel topological effects are predicted for the TRI topological insulators in (3+1) dimensions and (2+1) dimensions.

#### A. Second Chern number in (4+1)-dimensional nonlinear response

In this section, we will show how the second Chern number appears as a nonlinear-response coefficient of (4+1)-dimensional band insulators in an external  $U(1)$  gauge field, which is in exact analogy with the first Chern number as the Hall conductance of a (2+1)-dimensional system. To describe such a nonlinear response, it is convenient to use the path-integral formalism. The Hamiltonian of a (4+1)-dimensional insulator coupled to a  $U(1)$  gauge field is written as

$$H[A] = \sum_{m,n} (c_{m\alpha}^\dagger h_{mn}^{\alpha\beta} e^{iA_{mn}} c_{n\beta} + \text{H.c.}) + \sum_m A_{0m} c_{m\alpha}^\dagger c_{m\alpha}. \quad (49)$$

The effective action of gauge field  $A^\mu$  is obtained by the following path integral:

$$e^{iS_{\text{eff}}[A]} = \int D[c] D[c^\dagger] \exp \left\{ i \int dt \left[ \sum_m c_{m\alpha}^\dagger (i\partial_t) c_{m\alpha} - H[A] \right] \right\} \\ = \det[(i\partial_t - A_{0m}) \delta_{mn}^{\alpha\beta} - h_{mn}^{\alpha\beta} e^{iA_{mn}}], \quad (50)$$

which determines the response of the fermionic system through the equation

$$j_\mu(\mathbf{x}) = \frac{\delta S_{\text{eff}}[A]}{\delta A_\mu(\mathbf{x})}. \quad (51)$$

In the case of the (2+1)-dimensional insulators, the effective action  $S_{\text{eff}}$  contains a Chern-Simons term  $(C_1/4\pi) A_\mu \epsilon^{\mu\nu\tau} \partial_\nu A_\tau$ , as shown in Eq. (9) of Sec. II A, in which the first Chern number  $C_1$  appears as the coefficient. For the (4+1)-dimensional system, a similar topological term is in general present in the effective action, which is the second Chern-Simons term,

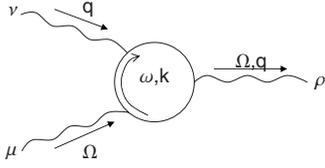


FIG. 7. The Feynman diagram that contributes to topological term (52). The loop is a fermion propagator and the wavy lines are external legs corresponding to the gauge field.

$$S_{\text{eff}} = \frac{C_2}{24\pi^2} \int d^4x dt \epsilon^{\mu\nu\rho\sigma\tau} A_\mu \partial_\nu A_\rho \partial_\sigma A_\tau, \quad (52)$$

where  $\mu, \nu, \rho, \sigma, \tau = 0, 1, 2, 3, 4$ . As shown in Refs. 33, 41, and 42, the coefficient  $C_2$  can be obtained by the one-loop Feynman diagram in Fig. 7, which can be expressed in the following symmetric form:

$$C_2 = -\frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4k d\omega}{(2\pi)^5} \text{Tr} \left[ \left( G \frac{\partial G^{-1}}{\partial q^\mu} \right) \left( G \frac{\partial G^{-1}}{\partial q^\nu} \right) \left( G \frac{\partial G^{-1}}{\partial q^\rho} \right) \right. \\ \left. \times \left( G \frac{\partial G^{-1}}{\partial q^\sigma} \right) \left( G \frac{\partial G^{-1}}{\partial q^\tau} \right) \right], \quad (53)$$

in which  $q^\mu = (\omega, k_1, k_2, k_3, k_4)$  is the frequency-momentum vector and  $G(q^\mu) = [\omega + i\delta - h(\mathbf{k})]^{-1}$  is the single-particle Green's function.

Now we are going to show the relation between  $C_2$  defined in Eq. (53) and the non-Abelian Berry's phase gauge field in momentum space. To make the statement clear, we first write down the conclusion.

For any (4+1)-dimensional band insulator with single particle Hamiltonian  $h(\mathbf{k})$ , the nonlinear-response coefficient  $C_2$  defined in Eq. (53) is equal to the second Chern number of the non-Abelian Berry's phase gauge field in the BZ, i.e.,

$$C_2 = \frac{1}{32\pi^2} \int d^4k \epsilon^{ijkl} \text{tr}[f_{ij} f_{kl}], \quad (54)$$

with

$$f_{ij}^{\alpha\beta} = \partial_i a_j^{\alpha\beta} - \partial_j a_i^{\alpha\beta} + i[a_i, a_j]^{\alpha\beta},$$

$$a_i^{\alpha\beta}(\mathbf{k}) = -i \langle \alpha, \mathbf{k} | \frac{\partial}{\partial k_i} | \beta, \mathbf{k} \rangle,$$

where  $i, j, k, \ell = 1, 2, 3, 4$ .

The index  $\alpha$  in  $a_i^{\alpha\beta}$  refers to the occupied bands; therefore, for a general multiband model,  $a_i^{\alpha\beta}$  is a non-Abelian gauge field and  $f_{ij}^{\alpha\beta}$  is the associated non-Abelian field strength. Here we sketch the basic idea of Eq. (54) and leave the explicit derivation to Appendix C. The key point to simplify Eq. (53) is noticing its *topological invariance*, i.e., under any continuous deformation of the Hamiltonian  $h(\mathbf{k})$ , as long as no level crossing occurs at the Fermi level,  $C_2$  remains invariant. Denote the eigenvalues of the single particle Hamiltonian  $h(\mathbf{k})$  as  $\epsilon_\alpha(\mathbf{k})$ ,  $\alpha = 1, 2, \dots, N$  with  $\epsilon_\alpha(\mathbf{k}) \leq \epsilon_{\alpha+1}(\mathbf{k})$ . When  $M$  bands are filled, one can always define a continuous deformation of the energy spectrum so that  $\epsilon_\alpha(\mathbf{k}) \rightarrow \epsilon_G$  for  $\alpha \leq M$  and  $\epsilon_\alpha(\mathbf{k}) \rightarrow \epsilon_E$  for  $\alpha > M$  (with  $\epsilon_E > \epsilon_G$ ), while all the corresponding eigenstates  $|\alpha, \mathbf{k}\rangle$  remain

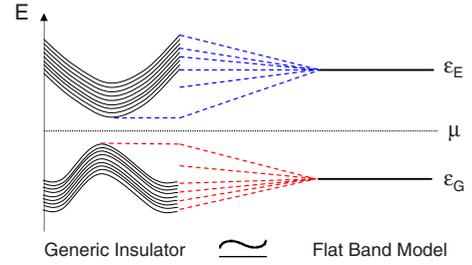


FIG. 8. (Color online) Illustration showing that a band insulator with arbitrary band structure  $\epsilon_i(k)$  can be continuously deformed to a flat band model with the same eigenstates. Since no level crossing occurs at the Fermi level, the two Hamiltonians are topologically equivalent.

invariant. In other words, each Hamiltonian  $h(\mathbf{k})$  can be continuously deformed to some ‘‘flat band’’ model, as shown in Fig. 8. Since both Eq. (53) and the second Chern number in Eq. (54) are topologically invariant, we only need to demonstrate Eq. (54) for the flat band models, of which the Hamiltonians have the form

$$h_0(\mathbf{k}) = \epsilon_G \sum_{1 \leq \alpha \leq M} |\alpha, \mathbf{k}\rangle \langle \alpha, \mathbf{k}| + \epsilon_E \sum_{\beta > M} |\beta, \mathbf{k}\rangle \langle \beta, \mathbf{k}| \\ \equiv \epsilon_G P_G(\mathbf{k}) + \epsilon_E P_E(\mathbf{k}), \quad (55)$$

Here  $P_G(\mathbf{k}) [P_E(\mathbf{k})]$  is the projection operator to the occupied (unoccupied) bands. Non-Abelian gauge connections can be defined in terms of these projection operators in a way similar to Ref. 43. Correspondingly, the single particle Green's function can also be expressed by the projection operators  $P_G, P_E$ , and Eq. (54) can be proved by straightforward algebraic calculations, as shown in Appendix C.

In summary, we have shown that for any (4+1)-dimensional band insulator, there is a (4+1)-dimensional Chern-Simons term (52) in the effective action of the external  $U(1)$  gauge field, of which the coefficient is the second Chern number of the non-Abelian Berry phase gauge field. Such a relation between Chern number and Chern-Simons term in the effective action is an exact analogy of the Thouless-Kohmoto-Nightingale-den Nijs (TKNN) formula<sup>5</sup> in (2+1)-dimensional QH effect. By applying the equation of motion [Eq. (51)], we obtain

$$j^\mu = \frac{C_2}{8\pi^2} \epsilon^{\mu\nu\rho\sigma\tau} \partial_\nu A_\rho \partial_\sigma A_\tau, \quad (56)$$

which is the nonlinear response to the external field  $A_\mu$ . For example, consider a field configuration,

$$A_x = 0, \quad A_y = B_z x, \quad A_z = -E_z t, \quad A_w = A_t = 0, \quad (57)$$

where  $x, y, z, w$  are the spatial coordinates and  $t$  is time. The only nonvanishing components of the field curvature are  $F_{xy} = B_z$  and  $F_{zt} = -E_z$ , which according to Eq. (56) generates the current,

$$j_w = \frac{C_2}{4\pi^2} B_z E_z.$$

If we integrate the equation above over the  $x, y$  dimensions [with periodic boundary conditions and assuming  $E_z$  does not depend on  $(x, y)$ ], we obtain

$$\int dxdy j_w = \frac{C_2}{4\pi^2} \left( \int dxdy B_z \right) E_z \equiv \frac{C_2 N_{xy}}{2\pi} E_z, \quad (58)$$

where  $N_{xy} = \int dxdy B_z / 2\pi$  is the number of flux quanta through the  $xy$  plane, which is always quantized to be an integer. This is exactly the 4DQH effect proposed in Ref. 14. Thus, from this example we can understand a physical consequence of the second Chern number: in a  $(4+1)$ -dimensional insulator with second Chern number  $C_2$ , a quantized Hall conductance  $C_2 N_{xy} / 2\pi$  in the  $zw$  plane is induced by magnetic field with flux  $2\pi N_{xy}$  in the perpendicular  $(xy)$  plane.

Similar to the  $(2+1)$ -dimensional case, the physical consequences of the second Chern number can also be understood better by studying the surface states of an open-boundary system, which for the  $(4+1)$ -dimensional case is described by a  $(3+1)$ -dimensional theory. In the next section we will study an explicit example of a  $(4+1)$ -dimensional topological insulator, which helps us to improve our understanding of the physical picture of the  $(4+1)$ -dimensional topology, especially, after dimensional reduction to the lower-dimensional physical systems.

### B. TRI topological insulators based on lattice Dirac models

In Sec. II B, we have shown that the model introduced in Ref. 35 realizes the fundamental TRB topological insulator in  $(2+1)$  dimensions and it reduces to the Dirac model in the continuum limit. Generalizing this construction, we propose the lattice Dirac model to be the realization of the fundamental TRI topological insulator in  $(4+1)$  dimensions. Such a model has also been studied in the field theory literature.<sup>41,44</sup> The continuum Dirac model in  $(4+1)$  dimensions is expressed as

$$H = \int d^4x [\psi^\dagger(x) \Gamma^i (-i\partial_i) \psi(x) + m \psi^\dagger \Gamma^0 \psi] \quad (59)$$

with  $i=1,2,3,4$  as the spatial dimensions and  $\Gamma^\mu, \mu=0,1,\dots,4$  as the five Dirac matrices satisfying the Clifford algebra,

$$\{\Gamma^\mu, \Gamma^\nu\} = 2\delta_{\mu\nu} \mathbb{I} \quad (60)$$

with  $\mathbb{I}$  as the identity matrix.<sup>45</sup>

The lattice (tight-binding) version of this model is written as

$$H = \sum_{n,i} \left[ \psi_n^\dagger \left( \frac{c\Gamma^0 - i\Gamma^i}{2} \right) \psi_{n+i} + \text{H.c.} \right] + m \sum_n \psi_n^\dagger \Gamma^0 \psi_n \quad (61)$$

or, in momentum space,

$$H = \sum_{\mathbf{k}} \psi_{\mathbf{k}}^\dagger \left[ \sum_i \sin k_i \Gamma^i + \left( m + c \sum_i \cos k_i \right) \Gamma^0 \right] \psi_{\mathbf{k}}. \quad (62)$$

Such a Hamiltonian can be written in the compact form,

$$H = \sum_{\mathbf{k}} \psi_{\mathbf{k}}^\dagger d_a(\mathbf{k}) \Gamma^a \psi_{\mathbf{k}} \quad (63)$$

with

$$d_a(\mathbf{k}) = \left[ \left( m + c \sum_i \cos k_i \right), \sin k_x, \sin k_y, \sin k_z, \sin k_w \right]$$

as a five-dimensional vector. Similar to the  $(2+1)$ -dimensional two-band models we studied in Sec. II B, a single particle Hamiltonian with the form  $h(\mathbf{k}) = d_a(\mathbf{k}) \Gamma^a$  has two eigenvalues  $E_{\pm}(\mathbf{k}) = \pm \sqrt{\sum_a d_a^2(\mathbf{k})}$ , but with the key difference that here both eigenvalues are doubly degenerate. When  $\sum_a d_a^2(\mathbf{k}) \equiv d^2(\mathbf{k})$  is nonvanishing in the whole BZ, the system is gapped at half-filling, with the two bands with  $E = E_{\pm}(\mathbf{k})$  filled. Since there are two occupied bands, an  $SU(2) \times U(1)$  adiabatic connection can be defined.<sup>43,46,47</sup> Starting from Hamiltonian (62), one can determine the single-particle Green's function and substitute it into the expression for the second Chern number in Eq. (53). We obtain

$$C_2 = \frac{3}{8\pi^2} \int d^4k \epsilon^{abcde} \hat{d}_a \partial_x \hat{d}_b \partial_y \hat{d}_c \partial_z \hat{d}_d \partial_w \hat{d}_e, \quad (64)$$

which is the winding number of the mapping  $\hat{d}_a(\mathbf{k}) \equiv d_a(\mathbf{k})/|d(\mathbf{k})|$  from the BZ  $T^4$  to the sphere  $S^4$  and  $a, b, c, d, e = 0, 1, 2, 3, 4$ . More details of this calculation are presented in Appendix D.

Since winding number (64) is equal to the second Chern number of Berry's phase gauge field, it is topologically invariant. It is easy to calculate  $C_2$  in lattice Dirac model (61). Considering the lattice Dirac model with a fixed positive parameter  $c$  and tunable mass term  $m$ ,  $C_2(m)$  as a function of  $m$  can change only if the Hamiltonian is gapless, i.e., if  $\sum_a d_a^2(\mathbf{k}, m) = 0$  for some  $\mathbf{k}$ . It is easy to determine that  $C_2(m) = 0$  in the limit  $m \rightarrow +\infty$  since the unit vector  $\hat{d}_a(\mathbf{k}) \rightarrow (1, 0, 0, 0, 0)$  in that limit. Thus we only need to study the change in  $C_2(m)$  at each quantum critical points, namely, at critical values of  $m$  where the system becomes gapless.

The solutions of equation  $\sum_a d_a^2(\mathbf{k}, m) = 0$  lead to five critical values of  $m$  and corresponding  $\mathbf{k}$  points as listed below:

$$m = \begin{cases} -4c, & \mathbf{k} = (0, 0, 0, 0) \\ -2c, & \mathbf{k} \in P[(\pi, 0, 0, 0)] \\ 0, & \mathbf{k} \in P[(\pi, 0, \pi, 0)] \\ 2c, & \mathbf{k} \in P[(\pi, \pi, \pi, 0)] \\ 4c, & \mathbf{k} = (\pi, \pi, \pi, \pi), \end{cases} \quad (65)$$

in which  $P[\mathbf{k}]$  stands for the set of all the wave vectors obtained from index permutations of wave vector  $\mathbf{k}$ . For example,  $P[(\pi, 0, 0, 0)]$  consists of  $(\pi, 0, 0, 0)$ ,  $(0, \pi, 0, 0)$ ,  $(0, 0, \pi, 0)$ , and  $(0, 0, 0, \pi)$ . As an example, we can study the change in  $C_2(m)$  around the critical value  $m = -4c$ . In the limit  $m + 4c \ll 2c$ , the system has its minimal gap at  $\mathbf{k} = \mathbf{0}$ , around which the  $d_a(\mathbf{k})$  vector has the approximate form  $d_a(\mathbf{k}) \approx (\delta m, k_x, k_y, k_z, k_w) + o(|k|)$ , with  $\delta m \equiv m + 4c$ . Taking a

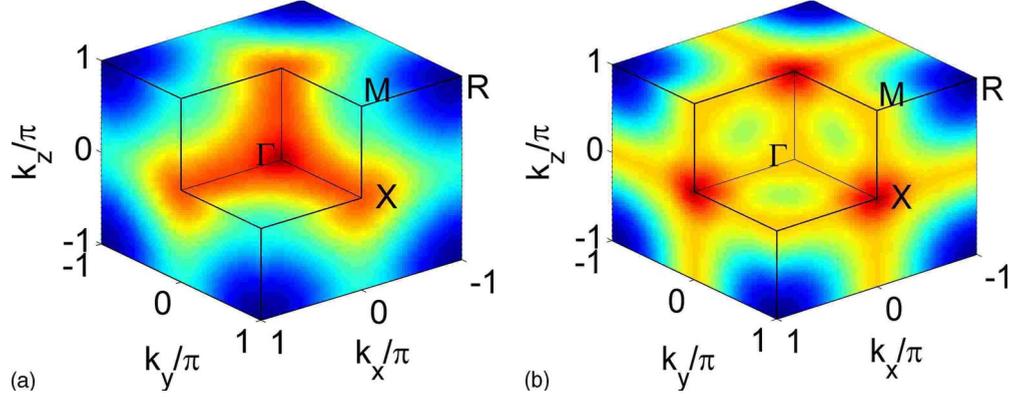


FIG. 9. (Color online) Three-dimensional energy spectrum of the surface states for the parameters (a)  $c=1, m=-3$  and (b)  $c=1, m=-1$ . The Dirac points where energy gap vanishes are marked with deepest red color. For  $m=-3$  there is one Dirac point at  $\Gamma$  point, while for  $m=-1$  there are three of them at  $X$  points.

cutoff  $\Lambda \ll 2\pi$  in momentum space, one can divide expression (64) of  $C_2$  into low-energy and high-energy parts,

$$C_2 = \frac{3}{8\pi^2} \left( \int_{|\mathbf{k}| \leq \Lambda} d^4k + \int_{|\mathbf{k}| > \Lambda} d^4k \right) \epsilon^{abcde} \hat{d}_a \partial_x \hat{d}_b \partial_y \hat{d}_c \partial_z \hat{d}_d \partial_w \hat{d}_e \\ \equiv C_2^{(1)}(\delta m, \Lambda) + C_2^{(2)}(\delta m, \Lambda).$$

Since there is no level crossing in the region  $|\mathbf{k}| > \Lambda$ , the jump of  $C_2$  at  $\delta m=0$  can only come from  $C_2^{(1)}$ . In the limit  $|\delta m| < \Lambda \ll 2\pi$ , the continuum approximation of  $d_a(\mathbf{k})$  can be applied to obtain

$$C_2^{(1)}(\delta m, \Lambda) \approx \frac{3}{8\pi^2} \int_{|\mathbf{k}| \leq \Lambda} d^4k \frac{\delta m}{(\delta m^2 + \mathbf{k}^2)^{5/2}},$$

which can be integrated and leads to

$$\Delta C_{2, \delta m=0^-}^{\delta m=0^+} = \Delta C_2^{(1)} \frac{\delta m=0^+}{\delta m=0^-} = 1. \quad (66)$$

From the analysis above we see that the change in the second Chern number is determined only by the effective continuum model around the level crossing wave vector(s). In this case the continuum model is just the Dirac model. Similar analysis can be carried out at the other critical  $m$ 's, which leads to the following values of the second Chern number:

$$C_2(m) = \begin{cases} 0, & m < -4c \text{ or } m > 4c \\ 1, & -4c < m < -2c \\ -3, & -2c < m < 0 \\ 3, & 0 < m < 2c \\ -1, & 2c < m < 4c. \end{cases} \quad (67)$$

A more general formula is given in Ref. 41.

After obtaining the second Chern number, we can study the surface states of the topologically nontrivial phases of this model. In the same way as in Sec. II C, we can take open boundary conditions for one dimension, say,  $w$ , and periodic boundary conditions for all other dimensions, so that  $k_x, k_y, k_z$  are still good quantum numbers. The Hamiltonian is transformed to a sum of 1D tight-binding models,

$$H = \sum_{\vec{k}, w} \left[ \psi_{\vec{k}}^\dagger(w) \left( \frac{c\Gamma^0 - i\Gamma^4}{2} \right) \psi_{\vec{k}}(w+1) + \text{H.c.} \right] + \sum_{\vec{k}, w} \psi_{\vec{k}}^\dagger(w) \\ \times \left[ \sin k_i \Gamma^i + \left( m + c \sum_i \cos k_i \right) \Gamma^0 \right] \psi_{\vec{k}}(w), \quad (68)$$

in which  $\vec{k} = (k_x, k_y, k_z)$ ,  $i=1, 2, 3$ , and  $w=1, 2, \dots, L$  are the  $w$  coordinates of lattice sites. The single-particle energy spectrum can be obtained as  $E_\alpha(\vec{k})$ ,  $\alpha=1, 2, \dots, 4L$ , among which the midgap surface states are found when  $C_2 \neq 0$ , as shown in Fig. 9. When the Chern number is  $C_2$ , there are  $|C_2|$  branches of gapless surface states with linear dispersion, so that the low energy effective theory is described by  $|C_2|$  flavors of chiral fermions,<sup>44</sup>

$$H = \text{sgn}(C_2) \int \frac{d^3p}{(2\pi)^3} \sum_{i=1}^{|C_2|} v_i \psi_i^\dagger(\vec{p}) \vec{\sigma} \cdot \vec{p} \psi_i(\vec{p}). \quad (69)$$

The factor  $\text{sgn}(C_2)$  ensures that the chirality of the surface states is determined by the sign of the Chern number. From such a surface theory we can obtain a more physical understanding of the nonlinear response equation [Eq. (56)] to an external  $U(1)$  gauge field. Taking the same gauge-field configuration as in Eq. (57), the nonvanishing components of the field curvature are  $F_{xy} = B_z$  and  $F_{zt} = -E_z$ . Consequently, the (3+1)-dimensional surface states are coupled to a magnetic field  $\mathbf{B} = B_z \hat{z}$  and an electric field  $\mathbf{E} = E_z \hat{z}$ . For simplicity, consider the system with  $-4c < m < -2c$  and  $C_2 = 1$ , in which the surface theory is a single chiral fermion with the single particle Hamiltonian,

$$h = v \vec{\sigma} \cdot (\vec{p} + \vec{A}) = v \sigma_x p_x + v \sigma_y (p_y + B_z x) + v \sigma_z (p_z - E_z t).$$

If  $E_z$  is small enough so that the time dependence of  $A_z(t) = -E_z t$  can be treated adiabatically, the single particle energy spectrum can be solved for a fixed  $A_z$  as

$$E_{n\pm}(p_z) = \pm v \sqrt{(p_z + A_z)^2 + 2n|B_z|}, \quad n = 1, 2, \dots,$$

$$E_0(p_z) = v(p_z + A_z) \text{sgn}(B_z). \quad (70)$$

When the size of the surface is taken as  $L_x \times L_y \times L_z$  with periodic boundary conditions, each Landau level has the de-

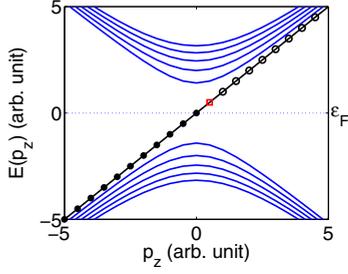


FIG. 10. (Color online) Illustration of the surface Landau-level spectrum given by Eq. (70). Each level in the figure is  $N_{xy}$ -fold degenerate. The solid circles are the occupied states of the zeroth Landau level and the red open square shows the states that are filled when the gauge vector potential  $A_z$  is shifted adiabatically from 0 to  $2\pi/L_z$ .

generacy  $N_{xy} = L_x L_y B_z / 2\pi$ . Similar to Laughlin's gauge argument for QH edge states,<sup>4</sup> the effect of an infinitesimal electric field  $E_z$  can be obtained by adiabatically shifting the momentum  $p_z \rightarrow p_z + E_z t$ . As shown in Fig. 10, from the time  $t=0$  to  $t=T \equiv 2\pi/L_z E_z$ , the momentum is shifted as  $p_z \rightarrow p_z + 2\pi/L_z$ , so that the net electron number of the surface 3D system increases by  $N_{xy}$ . In other words, a “generalized Hall current”  $I_w$  must be flowing toward the  $w$  direction,

$$I_w = \frac{N_{xy}}{T} = \frac{L_x L_y L_z B_z E_z}{4\pi^2}.$$

This generalized Hall current is the key property of the 4DQH effect studied in Ref. 14.

In terms of current density we obtain  $j_w = B_z E_z / 4\pi^2$ , which is consistent with the result of Eq. (56) discussed in the last section. More generally the current density  $j_w$  can be written as

$$j_w = C_2 \frac{\mathbf{E} \cdot \mathbf{B}}{4\pi^2} = \frac{C_2}{32\pi^2} \epsilon^{\mu\nu\sigma\tau} F_{\mu\nu} F_{\sigma\tau}, \quad (71)$$

which is the *chiral anomaly* equation of massless (3+1)-dimensional Dirac fermions.<sup>48,49</sup> Since the gapless states on the 3D edge of the 4D lattice Dirac model are chiral fermions, the current  $I_w$  carries away chiral charge, leading to the nonconservation of chirality on the 3D edge.

#### IV. DIMENSIONAL REDUCTION TO (3+1)-DIMENSIONAL TRI INSULATORS

As shown in Sec. II C, one can start from a (2+1)-dimensional TRB topological insulator described by a Hamiltonian  $h(k_x, k_y)$  and perform the procedure of dimensional reduction by replacing  $k_y$  by a parameter  $\theta$ . The same dimensional reduction procedure can be carried out for the (4+1)-dimensional TRI insulator with a nonvanishing second Chern number. From this procedure, one obtains the topological effective theory of insulators in (3+1) dimensions and (2+1) dimensions. Specifically, for TRI insulators a general  $Z_2$  topological classification is defined. Compared to the earlier proposals of the  $Z_2$  topological invariant,<sup>16,22–24,27–30</sup> our approach provides a direct relation-

ship between the topological quantum number and the physically measurable topological response of the corresponding system. We discuss a number of theoretical predictions, including the TME effect, and propose experimental settings where these topological effects can be measured in realistic materials.

#### A. Effective action of (3+1)-dimensional insulators

To perform the dimensional reduction explicitly, in the following we show the derivation for the (4+1)-dimensional Dirac model (61). However, each step of the derivation is applicable to any other insulator model, so the conclusion is completely generic.

The Hamiltonian of Dirac model (61) coupled to an external  $U(1)$  gauge field is given by

$$H[A] = \sum_{n,i} \left[ \psi_n^\dagger \left( \frac{c\Gamma^0 - i\Gamma^i}{2} \right) e^{iA_{n,n+i}} \psi_{n+i} + \text{H.c.} \right] + m \sum_n \psi_n^\dagger \Gamma^0 \psi_n. \quad (72)$$

Now consider a special “Landau-gauge” configuration satisfying  $A_{n,n+i} = A_{n+\hat{w},n+\hat{w}+\hat{i}}$ ,  $\forall n$ , which is translationally invariant in the  $w$  direction. Thus, under periodic boundary conditions the  $w$ -direction momentum  $k_w$  is a good quantum number, and the Hamiltonian can be rewritten as

$$H[A] = \sum_{k_w, \vec{x}, s} \left[ \psi_{\vec{x}, k_w}^\dagger \left( \frac{c\Gamma^0 - i\Gamma^s}{2} \right) e^{iA_{\vec{x}, \vec{x}+\hat{s}}} \psi_{\vec{x}+\hat{s}, k_w} + \text{H.c.} \right] + \sum_{k_w, \vec{x}, s} \psi_{\vec{x}, k_w}^\dagger \{ \sin(k_w + A_{\vec{x}4}) \Gamma^4 + [m + c \cos(k_w + A_{\vec{x}4})] \Gamma^0 \} \psi_{\vec{x}, k_w},$$

where  $\vec{x}$  stands for the three-dimensional coordinates,  $A_{\vec{x}4} \equiv A_{\vec{x}, \vec{x}+\hat{w}}$ , and  $s=1, 2, 3$  stands for the  $x, y, z$  directions. In this expression, the states with different  $k_w$  decouple from each other, and the (4+1)-dimensional Hamiltonian  $H[A]$  reduces to a series of (3+1)-dimensional Hamiltonians. Pick one of these (3+1)-dimensional Hamiltonians with fixed  $k_w$  and rename  $k_w + A_{\vec{x}4} = \theta_{\vec{x}}$ , we obtain the (3+1)-dimensional model,

$$H_{3D}[A, \theta] = \sum_{\vec{x}, s} \left[ \psi_{\vec{x}}^\dagger \left( \frac{c\Gamma^0 - i\Gamma^s}{2} \right) e^{iA_{\vec{x}, \vec{x}+\hat{s}}} \psi_{\vec{x}+\hat{s}} + \text{H.c.} \right] + \sum_{\vec{x}, s} \psi_{\vec{x}}^\dagger [ \sin \theta_{\vec{x}} \Gamma^4 + (m + c \cos \theta_{\vec{x}}) \Gamma^0 ] \psi_{\vec{x}}, \quad (73)$$

which describes a band insulator coupled to an electromagnetic field  $A_{\vec{x}, \vec{x}+\hat{s}}$  and an adiabatic parameter field  $\theta_{\vec{x}}$ .

Due to its construction, the response of model (73) to  $A_{\vec{x}, \vec{x}+\hat{s}}$  and  $\theta_{\vec{x}}$  fields is closely related to the response of the (4+1)-dimensional Dirac model (61) to the  $U(1)$  gauge field. To study the response properties of the (3+1)-dimensional system, the effective action  $S_{3D}[A, \theta]$  can be defined as

$$\exp^{iS_{3D}[A,\theta]} = \int D[\psi]D[\bar{\psi}] \exp \left\{ i \int dt \left[ \sum_{\vec{x}} \bar{\psi} \dot{\alpha} (i\partial_{\tau} - A_{\vec{x}0}) \psi_{\vec{x}} - H[A, \theta] \right] \right\}.$$

A Taylor expansion of  $S_{3D}$  can be carried out around the field configuration  $A_s(\vec{x}, t) \equiv 0$ ,  $\theta(\vec{x}, t) \equiv \theta_0$ , which contains a nonlinear-response term directly derived from the (4+1)-dimensional Chern-Simons action (52),

$$S_{3D} = \frac{G_3(\theta_0)}{4\pi} \int d^3x dt \epsilon^{\mu\nu\sigma\tau} \delta\theta \partial_{\mu} A_{\nu} \partial_{\sigma} A_{\tau}. \quad (74)$$

Compared to Eq. (52), the field  $\delta\theta(\vec{x}, t) = \theta(\vec{x}, t) - \theta_0$  plays the role of  $A_4$ , and the coefficient  $G_3(\theta_0)$  is determined by the same Feynman diagram (7) but evaluated for the three-dimensional Hamiltonian (73). Consequently,  $G_3(\theta_0)$  can be calculated and is equal to Eq. (53), but without the integration over  $k_w$ ,

$$G_3(\theta_0) = -\frac{\pi}{6} \int \frac{d^3k d\omega}{(2\pi)^4} \text{Tr} \epsilon^{\mu\nu\sigma\tau} \left[ \left( G \frac{\partial G^{-1}}{\partial q^{\mu}} \right) \left( G \frac{\partial G^{-1}}{\partial q^{\nu}} \right) \times \left( G \frac{\partial G^{-1}}{\partial q^{\sigma}} \right) \left( G \frac{\partial G^{-1}}{\partial q^{\tau}} \right) \left( G \frac{\partial G^{-1}}{\partial \theta_0} \right) \right], \quad (75)$$

where  $q^{\mu} = (\omega, k_x, k_y, k_z)$ . Due to the same calculation as in Sec. III A and Appendix C,  $G_3(\theta_0)$  is determined from the Berry phase curvature as

$$G_3(\theta_0) = \frac{1}{8\pi^2} \int d^3k \epsilon^{ijk} \text{tr}[f_{\theta i} f_{jk}], \quad (76)$$

in which the Berry phase gauge field is defined in the four-dimensional space  $(k_x, k_y, k_z, \theta_0)$ , i.e.,

$$a_i^{\alpha\beta} = -i \langle \vec{k}, \theta_0; \alpha | (\partial / \partial k_i) | \vec{k}, \theta_0; \beta \rangle$$

and

$$a_{\theta}^{\alpha\beta} = -i \langle \vec{k}, \theta_0; \alpha | (\partial / \partial \theta_0) | \vec{k}, \theta_0; \beta \rangle.$$

Compared to second Chern number (54), we know that  $G_3(\theta_0)$  satisfies the sum rule,

$$\int G_3(\theta_0) d\theta_0 = C_2 \in \mathbb{Z}, \quad (77)$$

which is in exact analogy with the sum rule of the pumping coefficient  $G_1(\theta)$  in Eq. (22) of the (1+1)-dimensional system. Recall that  $G_1(\theta)$  can be expressed as  $\partial P_1(\theta) / \partial \theta$ , where  $P_1(\theta)$  is simply the charge polarization. In comparison, a generalized polarization  $P_3(\theta_0)$  can also be defined in (3+1) dimensions so that  $G_3(\theta_0) = \partial P_3(\theta_0) / \partial \theta_0$ . (Recently, a similar quantity has also been considered in Ref. 50 from the point of view of semiclassical particle dynamics.) The conventional electric polarization  $\mathbf{P}$  couples linearly to the external electric field  $\mathbf{E}$  and the magnetic polarization  $\mathbf{M}$  couples linearly to the magnetic field  $\mathbf{B}$ ; however, as we shall show,  $P_3$  is a pseudoscalar which couples nonlinearly to the external electromagnetic field combination  $\mathbf{E} \cdot \mathbf{B}$ . For this reason, we coin the term ‘‘magnetoelectric polarization’’ for  $P_3$ .

To obtain  $P_3(\theta_0)$ , one needs to introduce the non-Abelian Chern-Simons term,

$$\mathcal{K}^A = \frac{1}{16\pi^2} \epsilon^{ABCD} \text{Tr} \left[ \left( f_{BC} - \frac{1}{3} [a_B, a_C] \right) \cdot a_D \right], \quad (78)$$

which is a vector in the four-dimensional parameter space  $q = (k_x, k_y, k_z, \theta_0)$  and  $A, B, C, D = x, y, z, \theta$ ,  $\mathcal{K}^A$  satisfies

$$\partial_A \mathcal{K}^A = \frac{1}{32\pi^2} \epsilon^{ABCD} \text{tr}[f_{AB} f_{CD}] \Rightarrow G_3(\theta_0) = \int d^3k \partial_A \mathcal{K}^A.$$

When the second Chern number is nonzero, there is an obstruction to the definition of  $a_A$ , which implies that  $\mathcal{K}_A$  cannot be a single-valued continuous function in the whole parameter space. However, in an appropriate gauge choice,  $\mathcal{K}_A^i$ ,  $i = x, y, z$ , can be single valued, so that  $G_3(\theta_0) = \int d^3k \partial_{\theta} \mathcal{K}^{\theta} \equiv \partial P_3(\theta_0) / \partial \theta_0$ , with

$$P_3(\theta_0) = \int d^3k \mathcal{K}^{\theta} = \frac{1}{16\pi^2} \int d^3k \epsilon^{\theta ijk} \text{Tr} \left[ \left( f_{ij} - \frac{1}{3} [a_i, a_j] \right) \cdot a_k \right]. \quad (79)$$

Thus,  $P_3(\theta_0)$  is given by the integral of the non-Abelian Chern-Simons three form over momentum space. This is analogous to the charge polarization defined as the integral of the adiabatic connection one form over a path in momentum space.

As is well known, the three-dimensional integration of the Chern-Simons term is only gauge-invariant modulo of an integer. Under a gauge transformation  $a_i \rightarrow u^{-1} a_i u - i u^{-1} \partial_i u$  [ $u \in U(M)$  when  $M$  bands are occupied], the change in  $P_3$  is

$$\Delta P_3 = \frac{i}{24\pi^2} \int d^3k \epsilon^{\theta ijk} \text{Tr}[(u^{-1} \partial_i u)(u^{-1} \partial_j u)(u^{-1} \partial_k u)],$$

which is an integer. Thus  $P_3(\theta_0)$ , just like  $P_1(\theta)$ , is only defined mod 1, and its change during a variation in  $\theta_0$  from 0 to  $2\pi$  is well defined and given by  $C_2$ .

Effective action (74) can be further simplified by introducing  $G_3 = \partial P_3 / \partial \theta$ . Integration by parts of  $S_{3D}$  leads to

$$S_{3D} = \frac{1}{4\pi} \int d^3x dt \epsilon^{\mu\nu\sigma\tau} A_{\mu} (\partial P_3 / \partial \theta) \partial_{\nu} \delta\theta \partial_{\sigma} A_{\tau}.$$

$(\partial P_3 / \partial \theta) \partial_{\nu} \delta\theta$  can be written as  $\partial_{\nu} P_3$ , where  $P_3(\vec{x}, t) = P_3[\theta(\vec{x}, t)]$  has space-time dependence determined by the  $\theta$  field. Such an expression is only meaningful when the space-time dependence of  $\theta$  field is smooth and adiabatic, so that locally  $\theta$  can still be considered as a parameter. In summary, the effective action is finally written as

$$S_{3D} = \frac{1}{4\pi} \int d^3x dt \epsilon^{\mu\nu\sigma\tau} P_3(x, t) \partial_{\mu} A_{\nu} \partial_{\sigma} A_{\tau}. \quad (80)$$

This effective topological action for the (3+1)-dimensional insulator is one of the central results of this paper. As we shall see later, many physical consequences can be directly derived from it. It should be emphasized that this effective

action is well defined for an arbitrary (3+1)-dimensional insulator Hamiltonian  $h(\vec{k}, \vec{x}, t)$  in which the dependence on  $\vec{x}, t$  is adiabatic. We obtained this effective theory by the dimensional reduction from a (4+1)-dimensional system, and we presented it this way since we believe that this derivation is both elegant and unifying. However, for readers who are not interested in the relationship to higher dimensional physics, a self-contained derivation can also be carried out directly in (3+1) dimensions, as we explained earlier, by integrating out the fermions in the presence of the  $A_\mu(x, t)$  and the  $\theta(x, t)$  external fields.

This effective action is known in the field theory literature as axion electrodynamics,<sup>51–53</sup> where the adiabatic field  $P_3$  plays the role of the axion field.<sup>54,55</sup> When the  $P_3$  field becomes a constant parameter independent of space and time, this effective action is referred to as the topological term for the  $\theta$  vacuum.<sup>56,57</sup> The axion field has not yet been experimentally identified, and it remains as a deep mystery in particle physics. Our work shows that the same physics can occur in a condensed matter system, where the adiabatic “axion” field  $P_3(x, t)$  has a direct physical interpretation and can be accessed and controlled experimentally.

From the discussion above it is clear that 3D TRI topological insulators realize a nontrivial solitonic background  $\theta$  field. In Ref. 28 Fu and Kane suggested several candidate materials which could be 3D topological insulators. These 3D materials are topologically nontrivial because of band inversion mechanism similar to that of the HgTe quantum wells.<sup>20</sup> Dai *et al.*<sup>58</sup> closely studied the strained bulk HgTe. We will keep this system in mind since it has a simple physical interpretation, and its essential physics can be described by the Dirac model presented earlier. We can consider the trivial vacuum outside the material to have a constant axion field with the value  $\theta=0$  and the interior of a 3D topological insulator to have a  $\theta=\pi$  background field. The value  $\theta=\pi$  does not violate time reversal (or the combined charge conjugation and parity symmetry (CP) in high-energy language). HgTe is a zero-gap semiconductor and has no topologically protected features. However, when strained, the system develops a bulk insulating gap between the  $p$ -wave light-hole “conduction band” and the  $p$ -wave heavy-hole “valence band” around the  $\Gamma$  point. To study the topological features we must also include the  $s$ -wave band which in a conventional material such as GaAs would be a conduction band. Because of the strong spin-orbit coupling in HgTe, the band structure is actually inverted, *i.e.*, the  $s$ -wave band becomes a valence band. For a moment we will ignore the heavy-hole band and only consider the light-hole and  $s$ -wave band.<sup>58</sup> The effective Hamiltonian of these two bands is a massive Dirac Hamiltonian but with a negative mass. The negative mass indicates a phase shift of  $\pi$  in the vacuum angle  $\theta$  from its original unshifted value in the trivial vacuum. The axion domain wall structure at the surface of the topological insulator traps fermion zero modes which are simply the topologically protected surface states. If we include the effects of the heavy-hole band the dispersions of the bulk bands and surface states are quantitatively modified. However, as long as the crystal is strained enough to maintain the bulk gap the topological phenomena will be unaffected and the boundary of the 3D topological insulator can still be described as an

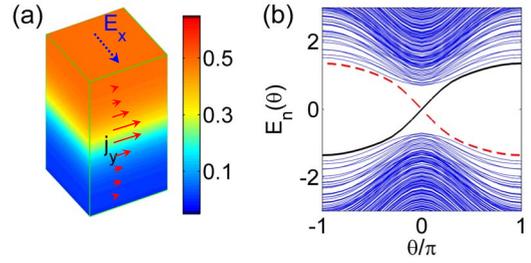


FIG. 11. (Color online) (a) Illustration of the Hall effect induced by a spatial gradient of  $P_3$ . The colors represent different values of  $P_3$ , which decrease from 0 at the bottom to  $1/2$  on top. The blue dashed arrow shows the direction of a uniform electric field  $E_x$  and the red solid arrows show the Hall current density induced, given by the formula  $j_y = (\partial_z P_3 / 2\pi) E_x$ . (b) Energy spectrum of the (3+1)-dimensional lattice Dirac model (73) in a magnetic field  $B_z$  toward the  $z$  direction. The boundary conditions are periodic in the  $x$  and  $y$  directions and open on the  $z$  direction. The red dashed and black solid curves show the surface states on the top and bottom surfaces, respectively, each of which is a  $B_z L_x L_y / 2\pi$ -fold degenerate Landau Level. The parameters of model (73) are chosen to be  $m=-3$ ,  $c=1$ .

axion domain wall. Thus, this material in condensed matter physics provides a direct realization of axion electrodynamics.

## B. Physical consequences of the effective action $S_{3D}$

In this section we present the general physical consequences of the effective topological action (80) for (3+1)-dimensional insulators coupled to a  $P_3$  polarization, and in Sec. IV C we focus on its consequences for TRI insulators. Since the effective action is quadratic in  $A_\mu$ , it describes a linear response to the external electromagnetic fields which depends on the spatial and temporal gradients of  $P_3$ . Taking a variation in  $S_{3D}[A, \theta]$  we obtain the response equation,

$$j^\mu = \frac{1}{2\pi} \epsilon^{\mu\nu\sigma\tau} \partial_\nu P_3 \partial_\sigma A_\tau. \quad (81)$$

The physical consequences [Eq. (81)] can be understood by studying the following two cases.

### 1. Hall effect induced by spatial gradient of $P_3$

Consider a system in which  $P_3 = P_3(z)$  only depends on  $z$ . For example, this can be realized by the lattice Dirac model (73) with  $\theta = \theta(z)$ . (This type of domain wall has also been considered in Ref. 59.) In this case Eq. (81) becomes

$$j^\mu = \frac{\partial_z P_3}{2\pi} \epsilon^{\mu\nu\rho} \partial_\nu A_\rho, \quad \mu, \nu, \rho = t, x, y,$$

which describes a QH effect in the  $xy$  plane with the Hall conductivity  $\sigma_{xy} = \partial_z P_3 / 2\pi$ , as shown in Fig. 11(a). For a uniform electric field  $E_x$  in the  $x$  direction, the Hall current density is  $j_y = (\partial_z P_3 / 2\pi) E_x$ . Thus the integration over  $z$  in a finite range gives the 2D current density in the  $xy$  plane,

$$J_y^{2D} = \int_{z_1}^{z_2} dz j_y = \frac{1}{2\pi} \left( \int_{z_1}^{z_2} dP_3 \right) E_x.$$

In other words, the net Hall conductance of the region  $z_1 \leq z \leq z_2$  is

$$\sigma_{xy}^{2D} = \int_{z_1}^{z_2} dP_3 / 2\pi, \quad (82)$$

which only depends on the change in  $P_3$  in this region and is not sensitive to any details of the function  $P_3(z)$ . Analogously in the (1+1)-dimensional case, if we perform the spatial integration of Eq. (25), we obtain the total charge induced by the charge polarization  $P$ ,

$$Q = - \int_{z_1}^{z_2} dP / 2\pi. \quad (83)$$

By comparing these two equations, we see that the relation between  $P_3$  and Hall conductance in (3+1)-dimensional insulators is the same as the relation between charge polarization  $P$  and the total charge in the (1+1)-dimensional case. As a specific case, a domain wall between two homogeneous materials with different  $P_3$  will carry Hall conductance  $\sigma_H = \Delta P_3 / 2\pi$ , while the fractional charge carried by a domain wall in (1+1) dimensions is given by  $Q = -\Delta P / 2\pi$ .

## 2. TME induced by temporal gradient of $P_3$

When  $P_3 = P_3(t)$  is spatially uniform, but time dependent, Eq. (81) becomes

$$j^i = - \frac{\partial_t P_3}{2\pi} \epsilon^{ijk} \partial_j A_k, \quad i, j, k = x, y, z.$$

In other words, we have

$$\vec{j} = - \frac{\partial_t P_3}{2\pi} \vec{B}. \quad (84)$$

Since the charge polarization  $\vec{P}$  satisfies  $\vec{j} = \partial_t \vec{P}$ , in a static uniform magnetic field  $B$  we have  $\partial_t \vec{P} = -\partial_t (P_3 \vec{B} / 2\pi)$ , so that

$$\vec{P} = - \frac{\vec{B}}{2\pi} (P_3 + \text{const}). \quad (85)$$

Such an equation describes the charge polarization induced by a magnetic field, which is a magnetoelectric effect. Compared to similar effects in multiferroic materials,<sup>60,61</sup> the magnetoelectric effect obtained here is of topological origin and only determined by the magnetoelectric polarization  $P_3$ .

Similar to the (1+1)-dimensional adiabatic pumping effect, the response [Eq. (84)] can be understood in a surface state picture. For example, consider the lattice Dirac model (73) with periodic boundary conditions in the  $x, y$  directions and open boundary conditions in the  $z$  direction. In the presence of a static magnetic field  $B_z$  in the  $z$  direction, the single particle energy spectrum  $E_n(\theta)$  can be solved for at a fixed  $\theta$  value. As shown in Fig. 11(b), midgap states appear for generic  $\theta$ , which are localized on the (2+1)-dimensional boundary. It should be noticed that each state is  $N$ -fold de-

generate where  $N = B_z L_x L_y / 2\pi$  is the Landau-level degeneracy. In the lattice Dirac model, when  $-4c < m < -2c$  so that  $C_2 = \int_{\theta=0}^{\theta=2\pi} dP_3 = 1$ , we find that during a period  $\theta=0 \rightarrow 2\pi$ ,  $N$  degenerate surface states on the bottom boundary sink below Fermi level and  $N$  states on the top float up. Consequently, when  $\theta$  is adiabatically tuned from 0 to  $2\pi$ , there are  $N$  electrons pumped from the top surface to the bottom one, which is in consistent with the result of Eq. (84),

$$\Delta Q = \int dt \int dx dy j_z = - \frac{\int_0^{2\pi} dP_3}{2\pi} B_z L_x L_y = -NC_2.$$

Just like the relation between (2+1)-dimensional QH edge states and the midgap end states in the (1+1)-dimensional pumping effect, there is a direct relationship between the (3+1)-dimensional surface states of the (4+1)-dimensional lattice Dirac model and the adiabatic pumping discussed above. As discussed in Sec. III B, the surface theory of a (4+1)-dimensional lattice Dirac model with nontrivial  $C_2$  is a (3+1)-dimensional chiral fermion. As shown in Fig. 10, the energy spectrum in a magnetic field  $B_z$  has a chiral dependence on the wave vector  $p_z$ . During the dimensional reduction procedure,  $p_z$  (in the notation of Sec. III B) is replaced by the parameter  $\theta$ , so that the chiral energy spectrum  $E(p_z)$  changes to the ‘‘chiral’’  $\theta$  dependence of  $E_n(\theta)$  in Fig. 11(b). In other words, the adiabatic pumping in the magnetic field in the (3+1)-dimensional system is a dimensionally reduced version of the chiral anomaly on the surface of a (4+1)-dimensional topological insulator.

The TME leads to a striking consequence if magnetic monopoles are present. For a uniform  $P_3$ , Eq. (84) leads to

$$\nabla \cdot \vec{j} = - \frac{\partial_t P_3}{2\pi} \nabla \cdot \vec{B}.$$

Suppose we consider a compact  $U(1)$  electromagnetic field on a lattice, where the monopole density  $\rho_m = \nabla \cdot \vec{B} / 2\pi$  can be nonvanishing, then we obtain

$$\partial_t \rho_e = (\partial_t P_3) \rho_m. \quad (86)$$

Therefore, when  $P_3$  is adiabatically changed from zero to  $\Theta / 2\pi$ , the magnetic monopole will acquire a charge of

$$Q_e = \frac{\Theta}{2\pi} Q_m. \quad (87)$$

Such a relation was first derived by Witten in the context of the topological term obtained from QCD.<sup>62</sup>

## C. $Z_2$ topological classification of time-reversal invariant insulators

In Sec. II D we have seen how a  $Z_2$  topological classification is obtained for (1+1)-dimensional particle-hole symmetric insulators. The key point for that case is to show that any interpolation between two particle-hole symmetric insulators  $h_1(k)$  and  $h_2(k)$  carries the same parity of Chern number, so that the relative Chern parity is well defined for each two Hamiltonians with particle-hole symmetry. In this sec-

tion, we will show that the same approach can be applied to (3+1)-dimensional insulators, where the time-reversal symmetry plays the same role as particle-hole symmetry does in (1+1) dimensions.

For a Hamiltonian  $H = \sum_{m,n} c_{m\alpha}^\dagger h_{mn}^{\alpha\beta} c_{n\beta}$ , the time-reversal transformation is an antiunitary operation defined by  $c_{m\alpha} \rightarrow T^{\alpha\beta} c_{m\beta}$ , where the *time-reversal* matrix  $T$  satisfies  $T^\dagger T = \mathbb{1}$  and  $T^* T = -\mathbb{1}$ . In  $\vec{k}$  space time-reversal symmetry requires

$$T^\dagger h(-\vec{k}) T = h^T(\vec{k}). \quad (88)$$

The condition  $T^* T = -\mathbb{1}$  is essential and leads to the Kramers degeneracy. Now we will follow the same approach as in Sec. II D and show how to define a  $Z_2$  invariant for the TRI insulators in (3+1) dimensions. For any two TRI band insulators  $h_1(\vec{k})$  and  $h_2(\vec{k})$ , an interpolation  $h(\vec{k}, \theta)$  can be defined, satisfying

$$h(\vec{k}, 0) = h_1(\vec{k}), \quad h(\vec{k}, \pi) = h_2(\vec{k}),$$

$$T^\dagger h(-\vec{k}, -\theta) T = h^T(\vec{k}, \theta), \quad (89)$$

and  $h(\vec{k}, \theta)$  is gapped for any  $\theta \in [0, 2\pi]$ . Since the interpolation is periodic in  $\theta$ , a second Chern number  $C_2[h(\vec{k}, \theta)]$  of the Berry phase gauge field can be defined in the  $(\vec{k}, \theta)$  space. In the same way as in Sec. II D, we will demonstrate below that  $C_2[h(\vec{k}, \theta)] - C_2[h'(\vec{k}, \theta)] = 0 \pmod{2}$  for any two interpolations  $h$  and  $h'$ . First of all, two new interpolations  $g_{1,2}(\vec{k}, \theta)$  can be defined by Eq. (39), which we repeat here for convenience,

$$g_1(k, \theta) = \begin{cases} h(k, \theta), & \theta \in [0, \pi] \\ h'(k, 2\pi - \theta), & \theta \in [\pi, 2\pi] \end{cases}$$

$$g_2(k, \theta) = \begin{cases} h'(k, 2\pi - \theta), & \theta \in [0, \pi] \\ h(k, \theta), & \theta \in [\pi, 2\pi] \end{cases}$$

By their definition,  $g_1$  and  $g_2$  satisfy  $C_2[h] - C_2[h'] = C_2[g_1] + C_2[g_2]$  and  $T^\dagger g_1(-\vec{k}, -\theta) T = g_2^T(\vec{k}, \theta)$ . To demonstrate  $C_2[g_1] = C_2[g_2]$ , consider an eigenstate  $|\vec{k}, \theta; \alpha\rangle_1$  of  $g_1(\vec{k}, \theta)$  with eigenvalue  $E_\alpha(\vec{k}, \theta)$ . We have

$$\begin{aligned} g_2^T(-\vec{k}, -\theta) T^\dagger |\vec{k}, \theta; \alpha\rangle_1 &= T^\dagger g_1(\vec{k}, \theta) |\vec{k}, \theta; \alpha\rangle_1 \\ &= E_\alpha(\vec{k}, \theta) T^\dagger |\vec{k}, \theta; \alpha\rangle_1 \\ &\Rightarrow g_2(-\vec{k}, -\theta) T^T (|\vec{k}, \theta; \alpha\rangle_1)^* \\ &= E_\alpha(\vec{k}, \theta) T^T (|\vec{k}, \theta; \alpha\rangle_1)^*. \end{aligned}$$

Thus  $T^T (|\vec{k}, \theta; \alpha\rangle_1)^*$  is an eigenstate of  $g_2(-\vec{k}, -\theta)$  with the same eigenvalue  $E_\alpha(\vec{k}, \theta)$ . Expand over the eigenstates  $|\vec{k}, -\theta; \beta\rangle_2$  of  $g_2(-\vec{k}, -\theta)$ , we have

$$T^T (|\vec{k}, \theta; \alpha\rangle_1)^* = \sum_\beta U_{\alpha\beta}(\vec{k}, \theta) |\vec{k}, -\theta; \beta\rangle_2. \quad (90)$$

Consequently the Berry phase gauge vector of the  $g_1$  and  $g_2$  systems satisfies

$$\begin{aligned} a_{1j}^{\alpha\beta}(\vec{k}, \theta) &= -i \langle \vec{k}, \theta; \alpha | \partial_j | \vec{k}, \theta; \beta \rangle_1 \\ &= -i \left[ \sum_{\gamma, \delta} U_{\alpha\gamma}^* \langle -\vec{k}, -\theta; \gamma | \partial_j (U_{\beta\delta} | -\vec{k}, -\theta; \delta \rangle_2) \right]^* \\ &= \sum_{\gamma, \delta} U_{\alpha\gamma} a_{2j}^{\gamma\delta*}(-\vec{k}, -\theta) (U^\dagger)_{\delta\beta} \\ &\quad - i \sum_\gamma U_{\alpha\gamma}(\vec{k}, \theta) \partial_j U_{\beta\gamma}^*(\vec{k}, \theta). \end{aligned} \quad (91)$$

In other words,  $a_{1j}^{\alpha\beta}(\vec{k}, \theta)$  is equal to  $a_{2j}^{\alpha\beta}(-\vec{k}, -\theta)$  up to a gauge transformation. Consequently, the Berry phase curvature satisfies  $f_{1ij}^{\alpha\beta}(\vec{k}, \theta) = U_{\alpha\gamma} f_{2ij}^{\gamma\delta*}(-\vec{k}, -\theta) (U^\dagger)_{\delta\beta}$ , which thus leads to  $C_2[g_1(\vec{k}, \theta)] = C_2[g_2(\vec{k}, \theta)]$ . In summary, we have proved  $C_2[h(\vec{k}, \theta)] - C_2[h'(\vec{k}, \theta)] = 2C_2[g(\vec{k}, \theta)] = 0 \pmod{2}$  for any two symmetric interpolations  $h$  and  $h'$ . Thus the ‘‘relative second Chern parity,’’

$$N_3[h_1(\vec{k}), h_2(\vec{k})] = (-1)^{C_2[h(\vec{k}, \theta)]},$$

is well defined for any two time-reversal invariant (3+1)-dimensional insulators, independent of the choice of interpolation. In the same way as in (1+1) dimensions, a vacuum Hamiltonian  $h_0(\vec{k}) \equiv h_0$ ,  $\forall \vec{k}$  can be defined as a reference. All the Hamiltonians with  $N_3[h_0, h] = -1$  are called  $Z_2$  nontrivial, while those with  $N_3[h_0, h] = 1$  are trivial. Similar to the (1+1)-dimensional case, there is a more intuitive, but less rigorous, way to define the  $Z_2$  invariant  $N_3$ . Through the derivation of Eq. (91) one can see that for a TRI Hamiltonian satisfying Eq. (88), Berry’s phase gauge potential satisfies  $a_i(\vec{k}) = U a_i(-\vec{k}) U^\dagger - i U \partial_i U^\dagger$ , so that the magnetoelectric polarization  $P_3$  satisfies

$$2P_3 = \frac{i}{24\pi^2} \int d^3k \epsilon^{ijk} \text{Tr}[(U \partial_i U^\dagger)(U \partial_j U^\dagger)(U \partial_k U^\dagger)] \in \mathbb{Z}.$$

Consequently, there are only two inequivalent TRI values of  $P_3$ , which are  $P_3 = 0$  and  $P_3 = 1/2$ . For two Hamiltonians  $h_1$  and  $h_2$ , the second Chern number  $C_2[h(\vec{k}, \theta)] = 2(P_3[h_2] - P_3[h_1]) \pmod{2}$ , so the difference of  $P_3$  determines the relative Chern parity  $N_3[h_1, h_2]$  by  $N_3[h_1, h_2] = (-1)^{2(P_3[h_1] - P_3[h_2])}$ . Since the trivial Hamiltonian  $h_0$  obviously has  $P_3 = 0$ , we know that all the Hamiltonians with  $P_3 = 1/2$  are topologically nontrivial, while those with  $P_3 = 0$  are trivial.

Once the  $Z_2$  classification is obtained, the physical consequences of this topological quantum number can be studied by effective theory (80), as has been done in the last section. In the (1+1)-dimensional case, we have shown that a zero-energy localized state exists at each open boundary of a  $Z_2$  nontrivial particle-hole symmetric insulator, which leads to a half charge  $Q_d = e/2 \pmod{e}$  confined on the boundary. Similarly, the nontrivial (3+1)-dimensional insulators also have topologically protected surface states. The easiest way to study the surface physics of the (3+1)-dimensional insulator is again by dimensional reduction. As discussed above, for any three-dimensional Hamiltonian  $h_1(\vec{k})$ , an interpolation  $h(\vec{k}, \theta)$  can be defined between  $h_1$  and the vacuum Hamiltonian  $h_0$ . If we interpret  $\theta$  as the fourth momentum,  $h(\vec{k}, \theta)$  defines a (4+1)-dimensional band insulator. Moreover, the

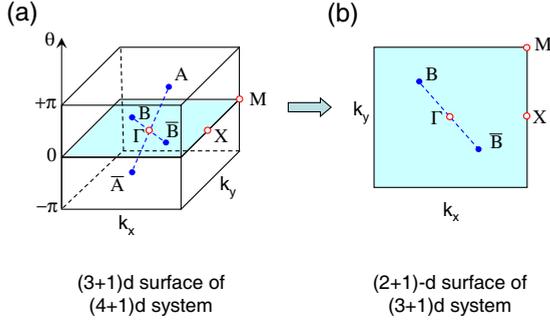


FIG. 12. (Color online) Illustration of the nodal points in the surface state energy spectrum of  $(4+1)$ -dimensional and  $(3+1)$ -dimensional insulators. (a) The nodal points in the  $(k_x, k_y, \theta)$  BZ for a  $z=\text{const.}$  surface of the  $(4+1)$ -dimensional system. The red hollow (blue filled) points stand for nodal points at time-reversal symmetric (asymmetric) wave vectors. The dashed lines are guides for the eyes. There are two pairs of asymmetric nodal points and three symmetric points in this example, which correspond to a bulk second Chern number  $C_2=7$ . (b) The nodal points in the  $(k_x, k_y)$  BZ for a  $z=\text{const.}$  surface of the  $(3+1)$ -dimensional system. According to the dimensional reduction procedure (see text) the 2D surface energy spectrum is given by the  $\theta=0$  slice of the 3D surface spectrum in (a). Since  $\theta=\pi$  corresponds to a vacuum Hamiltonian  $h_0$ , no nodal points exist in that plane. Consequently, the number of nodal points in the 2D BZ (five in this example) has the same parity as  $C_2$  in the  $(4+1)$ -dimensional system.

constraint [Eq. (89)] on  $h(\vec{k}, \theta)$  requires time-reversal symmetry for the corresponding  $(4+1)$ -dimensional system. The Hamiltonian  $h(\vec{k}, \theta)$  can be written in a real-space form and then defined on a four-dimensional lattice with open boundary conditions in the  $z$  direction and periodic boundary conditions for all the other directions. As discussed in Sec. III B, there will be  $|C_2[h]|$  flavors of  $(3+1)$ -dimensional chiral fermions on the surface when the second Chern number  $C_2[h]$  is nonzero. In other words, in the 3D BZ of the surface states there are  $|C_2[h]|$  nodal points  $(k_{xn}, k_{yn}, \theta_n)$ ,  $n=1, \dots, |C_2[h]|$ , where the energy spectrum  $E_n(k_x, k_y, \theta)$  is gapless and disperses linearly as a Dirac cone. From time-reversal symmetry it is easy to prove that the energy spectrum is identical for  $(k_x, k_y, \theta)$  and  $(-k_x, -k_y, -\theta)$ . Consequently, if  $(k_x, k_y, \theta)$  is a nodal point, so is  $(-k_x, -k_y, -\theta)$ . In other words, time-reversal symmetry requires the chiral fermions to appear in pairs, except for the ones at time-reversal symmetric points, as shown in Fig. 12(a). Thus, when the second Chern number  $C_2[h]$  is odd, there must be an odd number of Dirac cones at the eight symmetric points in the 3D BZ. Actually, the  $(4+1)$ -dimensional lattice Dirac model (62) provides an example of TRI insulators with nontrivial second Chern number since one can define  $\Gamma^0$  to be time-reversal even and  $\Gamma^{1,2,3,4}$  to be odd, as in conventional relativistic quantum mechanics<sup>63</sup>. As shown in Fig. 9, all the nodal points of the surface states are located at the symmetric points  $\Gamma, M, R$ , or  $X$ .

Now we return to the surface of the  $(3+1)$ -dimensional insulator. Since  $h_1(\vec{k})=h(\vec{k}, 0)$ ,  $h_0=h(\vec{k}, \pi)$  by definition of the interpolation, the surface energy spectra of  $h_1$  and  $h_0$  are given by the  $\theta=0$  and  $\theta=\pi$  slices of the 3D surface spec-

trum. Since all eight time-reversal symmetric points ( $\Gamma, X$ , and  $M$ ) are at  $\theta=0$  or  $\theta=\pi$ , we know that the net number of Dirac cones on the surface energy spectrum of  $h_1$  and  $h_0$  is odd (even) when  $C_2[h(\vec{k}, \theta)]$  is odd (even). However,  $h_0$  is defined as the vacuum Hamiltonian, which is totally local without any hopping between different sites. Thus, there cannot be any midgap surface states for  $h_0$ . Consequently, the number of 2D Dirac cones in the surface state spectrum of  $h_1$  must be odd (even) when  $C_2[h]$  is odd (even). Since the parity of  $C_2[h]$  determines the  $Z_2$  invariant  $N_3[h_1]$ , we finally reach the conclusion that there must be an odd (even) number of  $(2+1)$ -dimensional gapless Dirac fermions confined on the surface of a  $(3+1)$ -dimensional nontrivial (trivial) topological insulator.

Compared to earlier works on  $Z_2$  invariants and surface states in  $(3+1)$ -dimensions, one can see that the  $Z_2$  nontrivial topological insulator defined here corresponds to the “strong topological insulator” of Ref. 22. The present approach has the advantage of (i) demonstrating the bulk-edge relationship more explicitly, (ii) clarifying the connection between the second Chern number and the  $Z_2$  topological number, and (iii) naturally providing the effective theory that describes the physically measurable topological response properties of the system. The “weak topological insulators” defined in Ref. 22 are not included in the present approach since these  $Z_2$  invariants actually correspond to topological properties of  $(2+1)$ -dimensional insulators (QSH insulators, as will be discussed in next section), just as the QH effect in  $(3+1)$ -dimensional systems<sup>64</sup> still corresponds to a first Chern number, but defined in a 2D projection of the 3D BZ.

#### D. Physical properties of $Z_2$ -nontrivial insulators

In the last section we have defined the  $Z_2$  topological quantum number for the  $(3+1)$ -dimensional TRI insulators and discussed the gapless Dirac fermions on the surface of a nontrivial insulator. Now we will study the physical response properties of the nontrivial insulators. Since the nontrivial insulator has a magnetoelectric polarization  $P_3=1/2 \text{ mod } 1$ , according to Eq. (80) the effective action of the bulk system should be

$$S_{3D} = \frac{2n+1}{8\pi} \int d^3x dt \epsilon^{\mu\nu\sigma\tau} \partial_\mu A_\nu \partial_\sigma A_\tau, \quad (92)$$

in which  $n=P_3-1/2 \in \mathbb{Z}$  is the integer part of  $P_3$ . Under time-reversal symmetry, the term  $\epsilon^{\mu\nu\sigma\tau} \partial_\mu A_\nu \partial_\sigma A_\tau = 2\mathbf{E} \cdot \mathbf{B}$  is odd, so that for general  $P_3$ , effective action (80) breaks time-reversal symmetry. However, when the space-time manifold is closed (i.e., with periodic boundary conditions in the spatial and temporal dimensions), the term  $\int d^3x dt \epsilon^{\mu\nu\sigma\tau} \partial_\mu A_\nu \partial_\sigma A_\tau$  is quantized to be  $8\pi^2 m$ ,  $m \in \mathbb{Z}$ . Consequently,  $S_{3D} = (2n+1)m\pi$  so that the action  $e^{iS_{3D}} = e^{im\pi} = (-1)^m$  is time-reversal invariant and is independent of  $n$ , the integer part of  $P_3$ . This time-reversal property of the effective action is consistent with that of  $P_3$  discussed in the last section. Thus, in a closed space-time manifold, topological action (92) is a consistent effective theory of the  $Z_2$ -nontrivial TRI insulators, and the integer part of  $P_3$  is not a physical quantity. However, the case is different when the system has open boundaries. On a

space-time manifold with boundary, the value of  $S_{3D}$  is not quantized, which thus breaks time-reversal symmetry even for  $P_3=1/2$  or 0. In this case, the integer part  $n$  of  $P_3$  does enter the action  $e^{iS_{3D}}$  and becomes physical; its value depends on the quantitative details of the boundary.

To understand the physics in the open boundary system, we study a semi-infinite nontrivial insulator occupying the space  $z \leq 0$ . Since the vacuum, which fills  $z > 0$ , is effectively a trivial insulator (with an infinitely large gap), effective action (80) can be written in the whole of  $\mathbb{R}^3$  as

$$S_{3D} = \frac{1}{4\pi} \int d^3x dt \epsilon^{\mu\nu\sigma\tau} A_\mu \partial_\nu P_3 \partial_\sigma A_\tau.$$

Since  $P_3=1/2 \bmod 1$  for  $z < 0$  and  $P_3=0 \bmod 1$  for  $z > 0$ , we have

$$\partial_z P_3 = (n + 1/2) \delta(z),$$

where  $n \in \mathbb{Z}$  depends on the nontopological details of the surface, as will be studied later. In this case, the effective action is reduced to a (2+1)-dimensional Chern-Simons term on the surface,

$$S_{\text{surf}} = -\frac{2n+1}{8\pi} \int dx dy dt \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho. \quad (93)$$

This is consistent with the observation in Sec. IV B that a domain wall of  $P_3$  carries a QH effect. The Hall conductance of such a surface of a  $Z_2$  nontrivial insulator is thus  $\sigma_H = (n + 1/2)/2\pi$ , which is quantized as a half odd integer times the quanta  $e^2/h$ . On the other hand, from the discussion in the last section we know that there are always an odd number of (2+1)-dimensional Dirac fermions living on the surface of a nontrivial insulator. Thus the half QH effect on the surface can be easily understood by the parity anomaly of massless Dirac fermions.<sup>65</sup> Here we need to be more careful. The Hall conductance carried by a Dirac fermion is well defined only when the fermion mass is nonvanishing, so that a gap is opened. With the continuum Hamiltonian  $H = k_x \sigma^x + k_y \sigma^y + m \sigma^z$ , the Berry phase curvature can be calculated as in Eq. (11). The  $\mathbf{d}$  vector is given by  $\mathbf{d} = (k_x, k_y, m)$ , which has a meron-type configuration in  $k_x, k_y$  space, and thus carries a Hall conductance,<sup>65</sup>

$$\sigma_H = \frac{1}{4\pi} \text{sgn}(m) \left( = \frac{e^2}{2h} \text{sgn}(m) \right).$$

Now consider the surface of a topological insulator with  $2n+1$  gapless Dirac cones. From the discussion in the last section we know that without breaking time-reversal symmetry, at least one of these Dirac cones cannot be gapped. Now consider a perturbation that breaks time-reversal symmetry, i.e., a term which assigns a mass  $m_i, i=1, 2, \dots, 2n+1$  to each Dirac cone and induces a net Hall conductance  $\sigma_H = \sum_{i=1}^{2n+1} \text{sgn}(m_i)/4\pi$ . Since  $\sum_{i=1}^{2n+1} \text{sgn}(m_i)$  is an odd integer, the Hall conductance we obtain is consistent with surface Chern-Simons theory (93). From this discussion we can understand that effective action (93) describes a surface with time-reversal symmetry breaking, though the bulk system remains time-reversal invariant. The bulk topology requires there to be a  $1/2$  quanta in the Hall conductance, and the

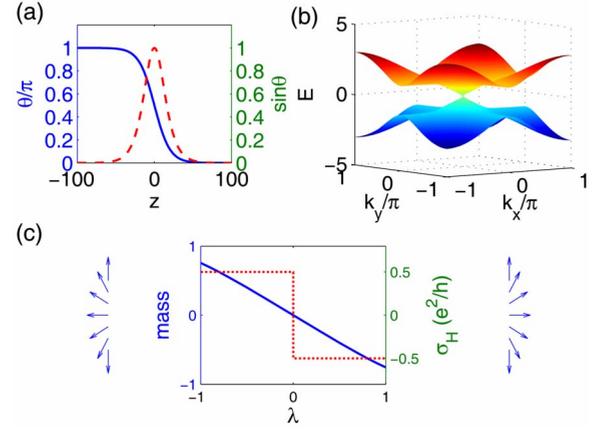


FIG. 13. (Color online) (a) Domain wall configuration  $\theta(z)$  (blue solid line) and corresponding  $\sin \theta(z)$  (red dashed line) as defined by Eq. (94) with  $\xi=20$ . (b) Dispersion relation of the two-dimensional bands trapped on the domain wall for the TRI Hamiltonian  $h_0$ . (c) The masses of surface Dirac fermion (blue solid line) and the half-quantized Hall conductance carried by the domain wall (red dashed line) in the Hamiltonian  $h(\lambda) = h_0 + \lambda h_1$  as a function of  $\lambda$ . The arrows on the left and right sides show schematically the rotation of angle  $\theta$  across the domain wall for  $\lambda < 0$  and  $\lambda > 0$ , respectively.

surface time-reversal symmetry breaking term determines the integer part  $n$ . This is an exact analog of the half charge on an end of the (1+1)-dimensional particle-hole symmetric insulator. As shown in Fig. 5, whether the localized state on the end of a nontrivial insulator is filled or vacant can only be determined by choosing a chemical potential  $\mu_1 > 0$  or  $\mu_2 < 0$  or, equivalently, by breaking the particle-hole symmetry around the boundary. The charge localized at one end of the insulator chain is  $(n+1/2)e$ , in which the integer part  $n$  depends on the symmetry breaking term on the surface, but the  $1/2$  part is guaranteed by the bulk topology.

To show such a surface state picture more explicitly, again the lattice Dirac model can be taken as an example. Consider the (3+1)-dimensional lattice Dirac model (73) with a domain wall configuration of the  $\theta(\vec{x})$  field given by

$$\theta(\vec{x}) = \theta(z) = \frac{\pi}{2} [1 - \tanh(z/4\xi)], \quad (94)$$

which has the asymptotic behavior  $\theta(z \rightarrow -\infty) = \pi$ ,  $\theta(z \rightarrow +\infty) = 0$ . The domain wall width is  $\xi$ , as shown in Fig. 13(a). With periodic boundary conditions in the  $x$  and  $y$  directions, the Hamiltonian can be block diagonalized,

$$\begin{aligned} H = & \sum_{z, k_x, k_y} \left[ \psi_{k_x, k_y}^\dagger(z) \left( \frac{c\Gamma^0 - i\Gamma^3}{2} \right) \psi_{k_x, k_y}(z+1) + \text{H.c.} \right] \\ & + \sum_{z, k_x, k_y} \psi_{k_x, k_y}^\dagger(z) [(m + c \cos \theta(z) + c \cos k_x + c \cos k_y) \Gamma^0 \\ & + \sin k_x \Gamma^1 + \sin k_y \Gamma^2] \psi_{k_x, k_y}(z) \\ & + \sum_{z, k_x, k_y} \psi_{k_x, k_y}^\dagger(z) \sin \theta(z) \Gamma^4 \psi_{k_x, k_y}(z) \equiv H_0 + H_1. \end{aligned} \quad (95)$$

Under a time-reversal transformation,  $\Gamma^0$  is even and  $\Gamma^{1,2,3,4}$  are odd. Thus the only time-reversal odd term in Hamiltonian (95) is the last term, which is localized around the boundary due to the factor  $\sin \theta(z)$  [see Fig. 13(a)]. Hereby we denote  $H=H_0+H_1$  with  $H_1$  as the last term and  $H_0$  as all the other TRI terms and define  $h_0, h_1$  as the single particle Hamiltonian corresponding to  $H_0, H_1$ , respectively. Then the Hamiltonian  $H_0$  describes a *time-reversal invariant* interface between two insulators  $\theta=0$  and  $\theta=\pi$ . For  $-4c < m < -2c$ , the parametrized Hamiltonian  $H(\theta)$ ,  $\theta \in [0, 2\pi]$  has a Chern number  $C_2=1$ , so that the  $\theta=0$  and  $\theta=\pi$  systems have relative Chern parity of  $-1$ . It is easy to show that Hamiltonian (73) for  $\theta=0$ ,  $-4c < m < -2c$  is adiabatically connected to the  $m \rightarrow -\infty$  limit. Thus we know that  $\theta=0$  and  $\theta=\pi$  correspond to  $Z_2$  trivial and nontrivial insulators, respectively. Consequently, on the domain wall at  $z=0$  there should be an odd number of gapless Dirac cones for  $H_0$ . As shown in Fig. 13(b), numerical diagonalization of  $h_0$  shows one single Dirac cone at  $(k_x, k_y)=(0,0)$  on the surface. To understand the effect of  $h_1$  term, notice that  $\{\Gamma^4, h_0\}=0$ , with  $\{\cdot\}$  being the anticommutator. The effective Hamiltonian of the Dirac cone can always be written as  $h_{\text{surf}}=k_x\sigma_x+k_y\sigma_y$  in a proper basis, and it should also anticommute with  $\Gamma^4$  since the bulk Hamiltonian does. Since the only term that anticommutes with  $h_{\text{surf}}$  in the  $2 \times 2$  Hilbert space is  $\sigma_z$ , we know that the effect of  $\Gamma^4$  term is to induce a mass term  $m\sigma_z$  in the lattice Dirac model. More accurately, the amplitude and the sign of  $m$  can be determined by standard perturbation theory. Given the two zero-energy surface states  $|k=0, \alpha\rangle$ ,  $\alpha=1, 2$ , the representation of the matrices  $\sigma_x$  and  $\sigma_y$  in the effective theory  $h_{\text{surf}}$  can be determined by

$$\sigma_{i\alpha\beta}^j = \langle k=0, \alpha | \left. \frac{\partial h_0}{\partial k_i} \right|_{k=0} | k=0, \beta \rangle, \quad i=x, y.$$

Then the  $\sigma^z$  is given by  $\sigma^z = -i\sigma^x\sigma^y$ , so that the mass  $m$  is determined by

$$m = \frac{1}{2} \sum_{\alpha\beta} \sigma_{\alpha\beta}^z \langle k=0, \beta | h_1 | k=0, \alpha \rangle.$$

If we consider the parametrized Hamiltonian  $h=h_0+\lambda h_1$ , then the mass of the surface Dirac fermion is proportional to  $\lambda$  for  $\lambda \rightarrow 0$ . As shown in Fig. 13(c), the mass is positive for  $\lambda < 0$ , which leads to a surface Hall conductance  $\sigma_H = -\text{sgn}(\lambda)/4\pi$ . On the other hand, the surface Hall conductance can also be calculated by the effective theory through Eq. (82). For  $\lambda=1$ , the phase field  $\theta$  winds from  $\pi$  to 0, which leads to  $\sigma_H = \int_{-\infty}^{+\infty} dP_3(z)/2\pi = \int_{\pi}^0 dP_3(\theta)/2\pi = -1/4\pi$  (since  $\int_0^\pi dP_3 = C_2/2 = 1/2$ ). The Hamiltonian for  $\lambda=-1$  can be considered to be the same lattice Dirac Hamiltonian  $H(\theta)$  with  $\theta(z)$  replaced by  $-\theta(z)$ . This keeps  $h_0$  invariant but reverses the sign of  $h_1$ . Consequently, the winding of the  $\theta$  field is from  $-\pi$  to 0, which leads to a Hall conductance  $\sigma_H = \int_{-\pi}^0 dP_3(\theta)/2\pi = 1/4\pi$ . The winding of  $\theta$  in the two cases is shown schematically in Fig. 13(c).

In summary, from this example we learn that the effect of a time-reversal symmetry breaking term on the surface is to assign a mass to the Dirac fermions which determines the winding direction of  $P_3$  through the domain wall. Once each

Dirac cone on the surface gains a mass, the whole system is gapped and the Berry phase curvature is well defined, so that the winding number of  $P_3$  through the domain wall is determined. Physically, the time-reversal symmetry breaking term on the surface can come from magnetic fields or magnetic moments localized on the surface; it could also arise from the spontaneous breaking of time reversal symmetry on the surface due to interactions. Once such a ‘‘T-breaking surface field’’ (denoted by  $M$ ) is applied, effective action (80) is well defined for open boundaries and describes the electromagnetic response of the  $Z_2$  nontrivial insulator. Actually, the T-breaking field should be considered to be an external field applied to the TRI system, such that topological action (80) describes a nonlinear response of the system to the combination of  $M$  and electromagnetic field  $A_\mu$ . For a  $Z_2$  nontrivial insulator occupying a spatial region  $\mathcal{V}$  with boundary  $\partial\mathcal{V}$ , the spatial gradient of  $P_3$  is given by

$$\nabla P_3(\vec{x}) = \left( g[M(\vec{x})] + \frac{1}{2} \right) \int_{\partial\mathcal{V}} d\hat{\mathbf{n}}(\vec{y}) \delta^3(\vec{x} - \vec{y}), \quad (96)$$

where  $g[M(\vec{x})] \in \mathbb{Z}$  is the integer part of the winding number determined by the T-breaking field  $M(\vec{x})$  and  $\hat{\mathbf{n}}$  is the normal vector of the surface. Under such a configuration of  $\nabla P_3$ , effective action (80) is reduced to the surface Chern-Simons action,

$$S_{\text{surf}} = \frac{1}{4\pi} \int_{\partial\mathcal{V}} d\hat{\mathbf{n}}_\mu \left( g[M(\vec{x})] + \frac{1}{2} \right) \epsilon^{\mu\nu\sigma\tau} A_\nu \partial_\sigma A_\tau. \quad (97)$$

Since  $g[M(\vec{x})]$  can only take discrete values, in general, the surface of a nontrivial insulator consists of several domains with different Hall conductances. To obtain more realistic predictions of effective theory (97), in the rest of this section we will study a specific case—the interface between a ferromagnetic insulator and a  $Z_2$  nontrivial insulator, where the surface time-reversal symmetry breaking is generated by the magnetization of the ferromagnet (FM) material. Several specific experimental proposals will be discussed.

### 1. Magnetization-induced QH effect

Consider the ferromagnet-topological insulator heterostructure shown in Fig. 14. The magnetization of the two FM layers can be parallel or antiparallel, and the standard six-terminal measurement can be performed to measure the in-plane Hall conductance. The net Hall conductance is given by the summation of the contributions of the top and bottom surfaces. When the topological insulator is uniform, an outward pointing magnetization vector (i.e., toward the direction of the surface normal vector  $\hat{\mathbf{n}}$ ) will have the same effect, no matter to which surface it is applied. Suppose the Hall current on the top surface induced by an electric field  $\mathbf{E} = E_x \hat{\mathbf{x}}$  is  $\mathbf{j}_t = \hat{\mathbf{n}}_t \times \mathbf{E}/4\pi$ , then on the bottom surface the same formula applies, such that  $\mathbf{j}_b = \hat{\mathbf{n}}_b \times \mathbf{E}/4\pi$ . Since  $\hat{\mathbf{n}}_t = -\hat{\mathbf{n}}_b = \hat{\mathbf{z}}$ , the current  $\mathbf{j}_t = -\mathbf{j}_b$ , as shown in Fig. 14. Consequently, the antiparallel magnetization leads to a vanishing net Hall conductance, while the parallel magnetization leads to  $\sigma_H = e^2/h$ .

Just like the usual integer QH effect, the quantized Hall conductance here is carried by chiral edge states. To under-

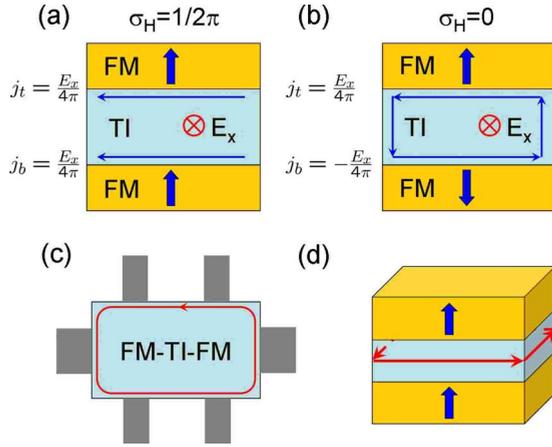


FIG. 14. (Color online) Illustration of the QH effect in ferromagnet-topological insulator heterostructure. [(a) and (b)] The electric field  $E_x$  (with direction into the paper) and the induced Hall currents  $j_t$  and  $j_b$  for parallel and antiparallel magnetizations, respectively. In (b) the Hall currents on the two surfaces are opposite and form a circulating current. (c) A top-down view of the device for Hall measurement. The gray regions are leads that contact to the surface of the topological insulator. (d) For the case of parallel magnetization, the chiral edge states are trapped on the side surfaces of the topological insulator. These carry the quantized Hall current.

stand the edge state picture, notice that for parallel magnetizations in Fig. 14(a), the magnetization vector is outward pointing at the top surface and inward pointing at the bottom surface. Although the Hall conductances of the two surfaces are the same in the global  $x, y, z$  basis, they are opposite in the *local* basis defined with respect to the normal vector  $\hat{n}$ . In other words, the integer  $g[M(\vec{x})]$  in the surface Chern-Simons theory (97) is 0 for the top surface and  $-1$  for the bottom surface. Consequently, the side surface is a domain wall between two different QH regions with Hall conductances that differ by one quantum. Just like a domain wall between  $\nu=0$  and  $\nu=1$  regions in the usual QH system, such a domain wall will trap a chiral Fermi liquid, which in this experimental proposal is responsible for the net Hall effect. It should be noticed that the side surface is two dimensions, so that generically there are also other nonchiral propagating modes on the side surface besides the one branch of chiral edge states. However, the existence of these nonchiral states does not change the stability of the chiral edge state since there is always one more right mover than left mover. The stability of the QH effect is still protected by the “bulk” gap, which is the magnetization-induced gap  $E_M$  in this case. Thus we will expect to observe this QH effect under the following two requirements: (i) temperature  $k_B T \ll E_M$  and (ii) the chemical potential on the top and bottom surfaces remains in the gap induced by the applied magnetization.

We would like to point out that this experimental proposal provides a direct demonstration of the half QH effect on the surface of a topological insulator. If the  $\sigma_H = e^2/h$  measured for parallel magnetization were contributed by one surface, then the magnetization flip of the other surface would have no effect on the net Hall conductance. Thus if an  $e^2/h$  Hall conductance is observed for parallel magnetization and the

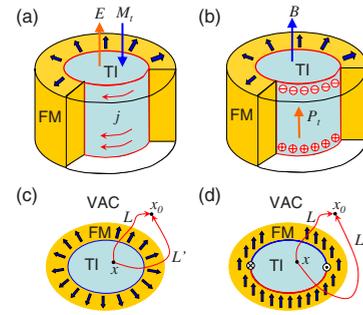


FIG. 15. (Color online) (a) Illustration of the magnetization induced by an electric field in a cylindrical geometry. The magnetization of the FM layer points outward from the side surface of the TI, and a circulating current is induced by the electric field. (b) Illustration of the charge polarization induced by a perpendicular magnetic field in the same geometry. “ $\oplus$ ” and “ $\ominus$ ” indicate the positive and negative charges induced by the magnetic field on the top and bottom surfaces, respectively. (c) Illustration of a topological insulator with fully gapped surface states induced by a hedgehog magnetization configuration.  $P_3$  of the topological insulator is well defined since the integral along any two paths  $L$  and  $L'$  gives the same  $\Delta P_3 = P_3(\vec{x}) - P_3(\vec{x}_0)$ . (d) Illustration of a topological insulator with parallel magnetization on the surface. A 1D domain wall between 2D surface regions with different Hall conductances (the upper and lower hemisphere as shown by red and blue, respectively) is induced, which carries chiral edge states as shown by  $\otimes$  and  $\odot$ . In this case the  $P_3$  of the topological insulator cannot be determined since two different paths  $L$  and  $L'$  lead to different  $\Delta P_3$ .

magnetization flip of either magnet leads to vanishing Hall conductance, one can conclude that the Hall conductance is contributed equally by the two surfaces.

## 2. TME

As has been discussed in Sec. IV B, a TME effect is induced by  $P_3$ , which is described by Eqs. (84) and (85). Now we consider the realization of this effect in a nontrivial topological insulator. Similar to the surface QH effect, a magnetization (or any other time-reversal symmetry breaking term) is necessary to determine the integer part of  $P_3$ . Consider the FM-TI-FM structure in Fig. 14(b). With antiparallel magnetizations, the current induced by an electric field  $E_x$  on the top and bottom surfaces flows in opposite directions. If we consider an isolated system rather than the Hall bar with leads as discussed above, a circulating current is formed, which induces a magnetic field parallel or antiparallel to the electric field. However, in the geometry shown in Fig. 14(b), dissipation occurs when the circulating current flows on the gapless side surface and the adiabatic condition of the TME effect is violated. To obtain the TME effect, a T-breaking gap for the side surface is necessary. This is satisfied in the cylindrical geometry shown in Fig. 15(a). With a magnetization pointing out of the cylinder’s surface, the surface is gapped and has a fixed Hall conductance  $\sigma_H = (n + \frac{1}{2})e^2/h$ . In an electric field parallel to the cylinder as shown in Fig. 15(a), a tangential circulating current is induced on the side surface, with the strength  $j_t = \sigma_H E$ . Such a circulating current appears identical to the surface bound current generated by a constant

magnetization  $M = j_t / c = (n + 1/2) \frac{e^2}{\hbar c} E$ , the direction of which is antiparallel to the electric field. In other words, the surface Hall current acts as a *topological magnetization current* such that the surface half QH effect is equivalent to a topological contribution to the bulk magnetization,

$$\mathbf{M}_t = - \left( n + \frac{1}{2} \right) \frac{e^2}{\hbar c} \mathbf{E}. \quad (98)$$

It should be noticed that the response coefficient is quantized as  $1/4\pi$  times odd multiples of the fine-structure constant  $\alpha = e^2 / \hbar c$ . Although the discussion above is carried only for an infinitely long cylindrical geometry, it can be easily verified that the response equation [Eq. (98)] applies for a sample with arbitrary shape. Generically, if an electric field  $\mathbf{E}$  is applied to a surface of a topological insulator with unit normal vector  $\hat{\mathbf{n}}$ , the Hall current is given by  $\mathbf{j}_H = (n + 1/2) \frac{e^2}{\hbar} \hat{\mathbf{n}} \times \mathbf{E}$ , which agrees with the magnetization current generated by the magnetization  $\mathbf{M}_t$  in Eq. (98). Combined with the conventional nontopological magnetic response  $\mathbf{M}_c$ , we obtain the total magnetization  $\mathbf{M} = \mathbf{M}_c + \mathbf{M}_t$  or, equivalently, the modified constituent equation,

$$\mathbf{H} = \mathbf{B} - 4\pi\mathbf{M}_c + (2n + 1) \frac{e^2}{\hbar c} \mathbf{E}. \quad (99)$$

Similar induction between electric and magnetic fields also occurs when a magnetic field is applied. Consider the process of applying a magnetic field  $B$  parallel to the cylinder as shown in Fig. 15(b). When the magnetic field is turned on from zero, a circulating electric field parallel to the side surface is generated, which then induces a Hall current  $j \propto dB/dt$  parallel or antiparallel to the magnetic field. Consequently, a charge density proportional to  $B$  is accumulated on the top and bottom surfaces, so that a magnetic field induces a topological contribution to the charge polarization, given by

$$\mathbf{P}_t = \left( n + \frac{1}{2} \right) \frac{e^2}{\hbar c} \mathbf{B}. \quad (100)$$

Combined with the conventional nontopological response  $\mathbf{P}_c$ , we obtain

$$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}_c - (2n + 1) \frac{e^2}{\hbar c} \mathbf{B}. \quad (101)$$

The conventional Maxwell's equations, supplemented by constituent relations an (99) and (101), give the complete description of the electrodynamics of the 3D topological insulators.

The alternative descriptions are to use the conventional constituent relations  $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}_c$  and  $\mathbf{H} = \mathbf{B} - 4\pi\mathbf{M}_c$  and work with a set of Maxwell's equations modified by the topological term. The total action of the electromagnetic field including the topological term is given by Eq. (80),

$$\begin{aligned} S_{\text{tot}} &= S_{\text{Maxwell}} + S_{\text{topo}} \\ &= \int d^3x dt \left[ \frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} F_{\mu\nu} \mathcal{P}^{\mu\nu} - \frac{1}{c} j^\mu A_\mu \right] \\ &\quad + \frac{\alpha}{16\pi} \int d^3x dt P_3 \epsilon^{\mu\nu\sigma\tau} F_{\mu\nu} F_{\sigma\tau}, \end{aligned} \quad (102)$$

in which  $\alpha \equiv e^2 / \hbar c$  is the fine-structure constant and  $\mathcal{P}^{0i} = P^i$  and  $\mathcal{P}^{ij} = \epsilon^{ijk} M_k$  are the electric and magnetic polarization vectors, respectively. The equations of motion are obtained by variation in the action over  $A^\mu$  as

$$\frac{1}{4\pi} \partial_\nu F^{\mu\nu} + \partial_\nu \mathcal{P}^{\mu\nu} + \frac{\alpha}{4\pi} \epsilon^{\mu\nu\sigma\tau} \partial_\nu (P_3 F_{\sigma\tau}) = \frac{1}{c} j^\mu. \quad (103)$$

These equations can also be written in the more familiar component form as

$$\begin{aligned} \nabla \cdot \mathbf{D} &= 4\pi\rho + 2\alpha(\nabla P_3 \cdot \mathbf{B}), \\ \nabla \times \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} &= \frac{4\pi}{c} \mathbf{j} - 2\alpha \left( (\nabla P_3 \times \mathbf{E}) + \frac{1}{c} (\partial_t P_3) \mathbf{B} \right), \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} &= 0, \\ \nabla \cdot \mathbf{B} &= 0, \end{aligned} \quad (104)$$

where  $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}_c$  and  $\mathbf{H} = \mathbf{B} - 4\pi\mathbf{M}_c$  include only the nontopological contributions. These are the equations of motion of axion electrodynamics.<sup>51-53</sup> By shifting the topological terms to the left-hand side, redefining  $\mathbf{D}$  and  $\mathbf{H}$  according to Eqs. (99) and (101), and taking  $P_3 = n + \frac{1}{2}$ , we recover the conventional Maxwell's equations, but with modified constituent relations, thus demonstrating the equivalence to the formulation given above. The quantization of the magnetoelectric coefficient  $\partial \mathbf{M} / \partial \mathbf{E}$  or  $\partial \mathbf{P} / \partial \mathbf{B}$  in odd units of the fine-structure constant is a profound topological phenomenon in condensed matter physics. Due to time reversal symmetry in the bulk of the topological insulator, it is not possible to have any contribution to the magnetoelectric coefficient from other nontopological mechanisms, which guarantees the robustness of the TME effect.

It should be emphasized that the value of  $P_3$  in the topological insulator can only be determined when a magnetization is applied to open a gap on the surface. As shown in Fig. 15(c), by defining a path  $L$  from a reference point  $\vec{x}_0$  deep in the vacuum,  $P_3$  can be determined by  $P_3(\vec{x}) = \int_{\vec{x}_0}^{\vec{x}} dl \cdot \nabla P_3$ . However, this definition only applies when the result does not depend on the choice of path. If a magnetic "shell" covered the surface of the topological insulator, with the magnetization outgoing everywhere on the interface, then the change in  $P_3$  across the interface is the same for different points on the surface, such that the bulk  $P_3$  is well defined without dependence on the choice of path. In this case, Eqs. (99) and (101) are well defined, and the integer part of  $P_3$  can change if the magnetization direction is reversed. On the other hand, when there are domain walls on the surface, the integer part of  $P_3$  is not well defined in the bulk of the

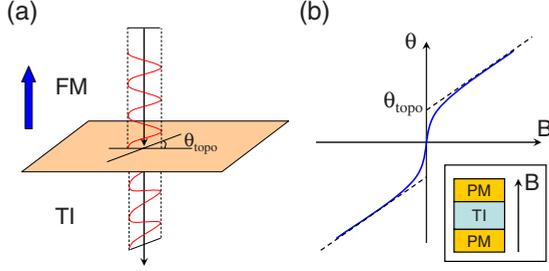


FIG. 16. (Color online) (a) Illustration of the Faraday rotation  $\theta_{\text{topo}}$  on the interface between a ferromagnet (or equivalently, a paramagnet in magnetic field) and a topological insulator. (b) Illustration of the total Faraday rotation angle  $\theta$  as a function of magnetic field  $B$  in the sandwich structure as shown in the inset. The zero-field extrapolation of  $\theta(B)$  determines the topological term  $\theta_{\text{topo}}$ .

topological insulator and Eqs. (99) and (101) do not apply, as shown in Fig. 15(d). Physically, the failure of Eqs. (99) and (101) is simply due to the existence of a QH edge current on the domain wall, which requires the more general Maxwell equations [Eq. (104)] including the contribution of the current. This analysis also provides a new picture of the surface QH effect, that is, the QH effect on the surface is carried by the chiral edge states living on vortex rings of the  $P_3$  field. It is only when there are no vortex rings of  $P_3$  on the surface, which the surface is fully gapped and the electromagnetic response is simply given by Eqs. (99) and (101).

### 3. Low-frequency Faraday rotation

The TME effect can be experimentally observed in the settings discussed above by applying an electric field through a capacitor and measuring the magnetic field by a superconducting quantum interference device (SQUID) device. Alternatively, we consider the experiment of Faraday or Kerr rotation. The modified Maxwell equations [Eq. (104)] can be applied to another phenomenon—photon propagation in the system.<sup>51</sup> It should be noted that effective theory (102) only applies in the low-energy limit  $E \ll E_g$ , where  $E_g$  is the gap of the surface state. Thus, to detect the topological phenomena we should study the low frequency photons with  $\omega \ll E_g/\hbar$ . Consider a FM-TI interface at  $z=0$ , as shown in Fig. 16. Normally incident linearly polarized light can be written as

$$\mathbf{A}(z,t) = \begin{cases} \mathbf{a}e^{i(-kz-\omega t)} + \mathbf{b}e^{i(kz-\omega t)}, & z > 0 \\ \mathbf{c}e^{i(-k'z-\omega t)}, & z < 0, \end{cases} \quad (105)$$

in which  $k=\omega/v$  and  $k'=\omega/v'$  are the wave vectors of the photon in the  $z>0$  and  $z<0$  regions, respectively. The  $\nabla P_3$  terms in Eq. (104) contribute a nonconventional boundary condition at  $z=0$ . Defining  $\nabla P_3 = \Delta \hat{\mathbf{z}} \delta(z)$  (with  $\Delta - 1/2 \in \mathbb{Z}$ ), the boundary conditions are given by

$$\mathbf{a} + \mathbf{b} = \mathbf{c},$$

$$\hat{\mathbf{z}} \times [k(-\mathbf{a} + \mathbf{b})/\mu + k'\mathbf{c}/\mu'] = -\frac{2\alpha\Delta\omega}{c}\mathbf{c},$$

in which the dimensionless constants  $\varepsilon, \varepsilon'$  and  $\mu, \mu'$  are the permittivity and permeability of the  $z>0$  and  $z<0$  materials, respectively. Denote  $a_{\pm} = a_x \pm ia_y$  and the same for  $b_{\pm}, c_{\pm}$ , the equations above lead to

$$a_+ = \frac{1}{2} \left[ 1 + \frac{k'/\mu' - 2i\alpha\Delta\omega/c}{k/\mu} \right] c_+.$$

Consequently, when the incident wave  $\mathbf{a}$  is linearly polarized, the transmission wave  $\mathbf{c}$  is also linearly polarized, with the polarization plane rotated by an angle

$$\theta_{\text{topo}} = \arctan \frac{2\alpha\Delta}{\sqrt{\varepsilon/\mu} + \sqrt{\varepsilon'/\mu'}}. \quad (106)$$

Here we always assume that the magnetization of the FM material is perpendicular to the  $xy$  plane, so that  $\mathbf{H} = \mu\mathbf{B}$  holds for in-plane magnetic fields. In the simplest case  $\mu, \mu' \simeq 1$ ,  $\varepsilon, \varepsilon' \simeq 1$ , and  $\Delta = 1/2$ , we get  $\theta \simeq \alpha \simeq 7.3 \times 10^{-3}$  rad.

Since the ferromagnetic material itself also induces a Faraday rotation, it's necessary to distinguish these two contributions in order to measure topological contribution (106). Replace the FM layers by paramagnetic materials with large susceptibilities and apply an external magnetic field to polarize them. In this case the magnetization is proportional to magnetic field, such that the Faraday rotation contributed by the bulk is also proportional to magnetic field. The net Faraday rotation is given by  $\theta = \theta_{\text{topo}}^{(i)} + \theta_{\text{topo}}^{(b)} + \theta_{\text{bulk}}$ , which has the following dependence on the magnetic field:

$$\theta(B) = uB + 2 \operatorname{sgn}(B) \arctan \frac{\alpha}{\sqrt{\varepsilon/\mu} + \sqrt{\varepsilon'/\mu'}}. \quad (107)$$

Consequently, the topological contribution can be obtained by measuring  $\theta(B)$  at different applied magnetic fields and extracting the linear extrapolation of  $\theta(B)$  as  $B \rightarrow 0^+$ .

Experimentally, the main difficulty of measuring such an effect comes from the low-frequency constraint  $\omega \ll E_g/\hbar$ . For a typical value  $E_g = 10$  meV we get  $f \equiv \omega/2\pi \ll 2.4$  THz, which is in the far infrared or microwave region. In principle, it is possible to find a topological insulator with a larger gap which can support an accurate measurement of Faraday rotation. Similar proposals as above can also be worked out for the rotation of reflected wave (Kerr effect).

## V. DIMENSIONAL REDUCTION TO (2+1) DIMENSIONS

By carrying out the same dimensional reduction procedure once more, we can obtain the topological effective theory for TRI (2+1)-dimensional insulators. Additionally a  $Z_2$  classification can be defined for (2+1)-dimensional TRI insulators, which is in exact analogy to the  $Z_2$  classification of (0+1)-dimensional particle-hole symmetric insulators. We will show that the (2+1)-dimensional  $Z_2$  nontrivial phase

corresponds to the QSH insulator proposed recently<sup>16,17,20,21</sup> and study the physical consequences of the effective theory.

### A. Effective action of (2+1)-dimensional insulators

In Sec. IV A we have seen how a (3+1)-dimensional insulator with a parameter field  $\theta(\vec{x}, t)$  is related to a (4+1)-dimensional insulator through dimensional reduction. In the same way, two parameter fields can be defined to obtain the dimensional reduction from (4+1) dimensions to (2+1) dimensions. In the following we will still take the lattice Dirac model as a canonical example to show the dimensional reduction procedure and derive the effective theory.

Starting from the lattice Dirac model (72) and choosing a special gauge vector configuration satisfying  $A_{n,n+i} = A_{n+\hat{w}, n+\hat{w}+i} = A_{n+\hat{z}, n+\hat{z}+i}$  (so that the gauge vector is homogeneous along  $z$  and  $w$ ) we obtain the Hamiltonian,

$$\begin{aligned}
 H[A] = & \sum_{k_z, k_w, \mathbf{x}} \sum_{s=1,2} \left[ \psi_{\mathbf{x}; k_z, k_w}^\dagger \left( \frac{c\Gamma^0 - i\Gamma^s}{2} \right) e^{iA_{\mathbf{x}, \mathbf{x}+\hat{s}}} \psi_{\mathbf{x}+\hat{s}; k_z, k_w} + \text{H.c.} \right] \\
 & + \sum_{k_z, k_w, \mathbf{x}} \sum_{s=1,2} \psi_{\mathbf{x}; k_z, k_w}^\dagger \cdot \{ \sin(k_z + A_{\mathbf{x}3}) \Gamma^3 \\
 & + \sin(k_w + A_{\mathbf{x}4}) \Gamma^4 + [m + c \cos(k_z + A_{\mathbf{x}3}) \\
 & + c \cos(k_w + A_{\mathbf{x}4})] \Gamma^0 \} \psi_{\mathbf{x}; k_z, k_w},
 \end{aligned}$$

in which  $\mathbf{x}=(x, y)$  is the two-dimensional coordinate. As in the (3+1)-dimensional case, the gauge fields in the  $z$  and  $w$  directions can be replaced by parameter fields  $(k_z + A_{\mathbf{x}3}) \rightarrow \theta_{\mathbf{x}}$  and  $(k_w + A_{\mathbf{x}4}) \rightarrow \varphi_{\mathbf{x}}$ , resulting in the parametrized family of (2+1)-dimensional Hamiltonians,

$$\begin{aligned}
 H_{2D}[A, \theta, \varphi] = & \sum_{\mathbf{x}, s} \left[ \psi_{\mathbf{x}}^\dagger \left( \frac{c\Gamma^0 - i\Gamma^s}{2} \right) e^{iA_{\mathbf{x}, \mathbf{x}+\hat{s}}} \psi_{\mathbf{x}+\hat{s}} + \text{H.c.} \right] \\
 & + \sum_{\mathbf{x}, s} \psi_{\mathbf{x}}^\dagger [ \sin \theta_{\mathbf{x}} \Gamma^3 + \sin \varphi_{\mathbf{x}} \Gamma^4 + (m + c \cos \theta_{\mathbf{x}} \\
 & + c \cos \varphi_{\mathbf{x}}) \Gamma^0 ] \psi_{\mathbf{x}}. \quad (108)
 \end{aligned}$$

By integrating out the fermion fields and expanding the resulting effective action around  $A_s=0$ ,  $\theta=\theta_0$ , and  $\varphi=\varphi_0$ , the same nonlinear term shown in the Feynman diagram in Fig. 7 leads to the topological term,

$$S_{2D} = \frac{G_2(\theta_0, \varphi_0)}{2\pi} \int d^2x dt \epsilon^{\mu\nu\rho} A_\mu \partial_\nu \delta\theta \partial_\rho \delta\varphi, \quad (109)$$

in which the coefficient  $G_2(\theta_0, \varphi_0)$  is determined by the same correlation function as Eq. (53) but without the integrations over  $k_z, k_w$ ,

$$\begin{aligned}
 G_2(\theta_0, \varphi_0) = & \frac{2\pi}{3} \int \frac{d^2k d\omega}{(2\pi)^3} \text{Tr} \epsilon^{\mu\nu\rho} \left[ \left( G \frac{\partial G^{-1}}{\partial q^\mu} \right) \left( G \frac{\partial G^{-1}}{\partial q^\nu} \right) \right. \\
 & \times \left. \left( G \frac{\partial G^{-1}}{\partial q^\rho} \right) \left( G \frac{\partial G^{-1}}{\partial \theta_0} \right) \left( G \frac{\partial G^{-1}}{\partial \varphi_0} \right) \right] \\
 = & \frac{1}{4\pi} \int d^2k e^{ij} \text{Tr} [2f_i \theta f_{j\varphi} - f_{ij} \theta_\varphi], \quad (110)
 \end{aligned}$$

in which  $\mu, \nu, \rho=0, 1, 2$ ,  $i, j=1, 2$ ,  $q^\mu=(\omega, k_x, k_y)$ , and the

Berry curvature is defined in the four-dimensional parameter space  $(k_x, k_y, \theta, \varphi)$ . The coefficient  $G_2(\theta_0, \varphi_0)$  satisfies the sum rule,

$$\int G_2(\theta_0, \varphi_0) d\theta_0 d\varphi_0 = 2\pi C_2. \quad (111)$$

To simplify the expression further, the Chern-Simons form  $\mathcal{K}^A$  in Eq. (78) can be introduced again. Here  $A$  runs over  $k_x, k_y, \theta, \varphi$ , and  $G_2(\theta_0, \varphi_0)$  can be written in terms of  $\mathcal{K}^A$  as

$$G_2(\theta_0, \varphi_0) = -2\pi \int d^2k (\partial_x \mathcal{K}^x + \partial_y \mathcal{K}^y + \partial_\theta \mathcal{K}^\theta + \partial_\varphi \mathcal{K}^\varphi).$$

Similar to the (3+1)-dimensional case, the momentum derivative terms  $\partial_{(x,y)} \mathcal{K}^{(x,y)}$  lead to vanishing contributions if  $\mathcal{K}^{(x,y)}$  is single valued, in which case  $G_2$  can be expressed as

$$G_2(\theta_0, \varphi_0) = \partial_\theta \Omega_\varphi - \partial_\varphi \Omega_\theta,$$

with

$$\Omega_\varphi = -2\pi \int d^2k \mathcal{K}^\theta, \quad \Omega_\theta = 2\pi \int d^2k \mathcal{K}^\varphi. \quad (112)$$

Notice that

$$\begin{aligned}
 -\mathcal{K}^\theta = & -\frac{1}{16\pi^2} \epsilon^{ij} \text{Tr} \left[ \left( f_{ij} - \frac{1}{3} [a_i, a_j] \right) \cdot a_\varphi \right. \\
 & \left. - 2 \left( f_{i\varphi} - \frac{1}{3} [a_i, a_\varphi] \right) \cdot a_j \right]
 \end{aligned}$$

and similarly for  $\mathcal{K}^\varphi$ . We know that the vector  $\Omega=(\Omega_\theta, \Omega_\varphi)$  has the correct transformation properties of a gauge vector potential under the coordinate transformations of the parameter space  $(\theta, \varphi)$  and also under gauge transformations of the wave functions. Consequently, when the parameters  $\theta = \theta(x^\mu)$  and  $\varphi = \varphi(x^\mu)$  have smooth dependence on space-time coordinates, an effective gauge vector potential  $\Omega_\mu$  can be defined in (2+1)-dimensional *space time* as

$$\Omega_\mu \equiv \Omega_\theta \partial_\mu \delta\theta + \Omega_\varphi \partial_\mu \delta\varphi, \quad (113)$$

the gauge curvature of which is related to  $G_2$  as

$$\begin{aligned}
 \partial_\mu \Omega_\nu - \partial_\nu \Omega_\mu = & (\partial_\theta \Omega_\varphi - \partial_\varphi \Omega_\theta) (\partial_\mu \delta\theta \partial_\nu \delta\varphi - \partial_\nu \delta\theta \partial_\mu \delta\varphi) \\
 = & G_2 (\partial_\mu \delta\theta \partial_\nu \delta\varphi - \partial_\nu \delta\theta \partial_\mu \delta\varphi). \quad (114)
 \end{aligned}$$

Mathematically,  $G_2$  is a density of second Chern form in the 2D parameter space  $(\theta, \varphi)$ , and  $(\theta, \varphi)=[\theta(x^\mu), \varphi(x^\mu)]$  defines a smooth map from the (2+1)-dimensional space-time manifold to the 2D parameter space. The curvature of the gauge potential  $\Omega_\mu$  is the pullback of  $G_2$  to (2+1)-dimensional space time. By making use of Eq. (113) effective action (109) can be rewritten in a parameter-independent form,

$$S_{2D} = \frac{1}{2\pi} \int d^2x dt \epsilon^{\mu\nu\tau} A_\mu \partial_\nu \Omega_\tau. \quad (115)$$

The physical consequences of effective theory (115) can be understood by studying the response equation,

$$j^\mu = \frac{1}{2\pi} \epsilon^{\mu\nu\rho} \partial_\nu \Omega_\rho. \quad (116)$$

As will be shown in the next section, Eq. (116) is the fundamental response equation for the QSH effect, which takes the form similar to the fundamental response equation for QH effect (8), with the replacement of the external gauge field by the effective Berry's phase gauge field. In this sense, our formalism provides a unifying theory for both effects. This type of relationship between different types of topological insulators will be discussed in more detail in Sec. VI.

It is worth noting that the response equation [Eq. (116)] can be expressed in an explicit form for Dirac model (108). According to Eq. (64), the momentum-space second Chern number of the (4+1)-dimensional Dirac model  $h(\mathbf{k}) = \sum_a d_a(\mathbf{k}) \Gamma^a$  is equal to the winding number of  $\hat{\mathbf{d}}(\mathbf{k})$  on the unit sphere  $S^4$ . Correspondingly, the Hamiltonian of the (2+1)-dimensional Dirac model (108) with constants  $\theta$  and  $\varphi$  has the form  $h(\mathbf{k}, \theta, \varphi) = \sum_a d_a(\mathbf{k}, \theta, \varphi) \Gamma^a$ , so the correlation function  $G_2$  defined in Eq. (110) can be obtained as

$$G_2(\theta, \varphi) = \frac{3}{4\pi} \int d^2k \epsilon^{abcde} \frac{d_a \partial_k d_b \partial_k d_c \partial_\theta d_d \partial_\varphi d_e}{|\mathbf{d}(\mathbf{k}, \theta, \varphi)|^5}.$$

Thus the curvature of effective gauge vector potential  $\Omega_\mu$  is expressed as

$$\partial_\mu \Omega_\nu - \partial_\nu \Omega_\mu = 3 \epsilon^{abcde} \int \frac{d^2k d_a \partial_k d_b \partial_k d_c \partial_\theta d_d \partial_\varphi d_e}{4\pi |\mathbf{d}(\mathbf{k}, \theta, \varphi)|^5}. \quad (117)$$

Now consider a slightly different version of lattice Dirac model given by

$$h(\mathbf{k}, \mathbf{n}) = \sin k_x \Gamma^1 + \sin k_y \Gamma^2 + (\cos k_x + \cos k_y - 2) \Gamma^0 + m \sum_{a=0,3,4} \hat{n}_a \Gamma^a,$$

in which  $m > 0$  and  $\hat{\mathbf{n}} = (\hat{n}_0, \hat{n}_3, \hat{n}_4)$  is a 3D unit vector. For such a model the  $\mathbf{d}$  vector can be decomposed as

$$\mathbf{d}(\mathbf{k}, \theta, \varphi) = \mathbf{d}_0(\mathbf{k}) + \begin{pmatrix} 0 \\ 0 \\ m\hat{\mathbf{n}} \end{pmatrix},$$

with  $\mathbf{d}_0(\mathbf{k}) = (\sin k_x, \sin k_y, 0, 0, \cos k_x + \cos k_y - 2)$ . In the limit  $m \ll 2$ , the Hamiltonian has the continuum limit  $h(\mathbf{k}, \hat{\mathbf{n}}) \simeq \sum_{a=1,2} k_a \Gamma^a + \sum_{b=0,3,4} m \hat{n}_b \Gamma^b$ , which is the continuum  $4 \times 4$  Dirac model with three possible mass terms. In this limit the integral over  $\mathbf{k}$  in Eq. (117) can be explicitly carried out, leading to the following expression:

$$\partial_\mu \Omega_\nu - \partial_\nu \Omega_\mu = \frac{1}{2} \hat{\mathbf{n}} \cdot \partial_\mu \hat{\mathbf{n}} \times \partial_\nu \hat{\mathbf{n}},$$

which is the *skyrmion density* of the unit vector  $\hat{\mathbf{n}}$ . Combined with Eq. (116) we obtain the response equation for the Dirac model in the continuum limit

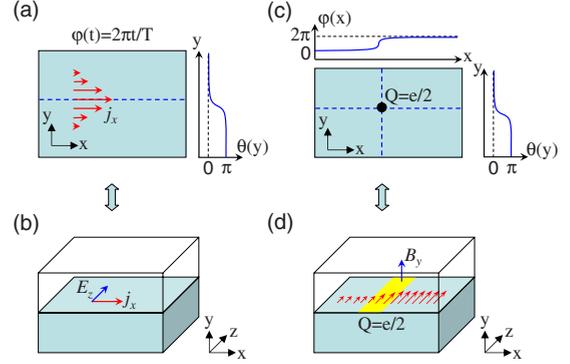


FIG. 17. (Color online) (a) Illustration of the charge pumping effect in the (2+1)-dimensional lattice Dirac model with spatial dependent  $\theta = \theta(y)$  and time dependent  $\varphi = \varphi(t)$ . This effect is the dimensional reduction in the domain-wall QH effect in the (3+1)-dimensional system as shown in (b), in which  $\partial_t \varphi$  plays the role of electric field  $E_z$ . (c) Illustration of the half charge trapped at the crossing point of the  $\theta$  and  $\varphi$  domain walls. This effect is the dimensional reduction in the charge trapped on the domain wall by a magnetic field, as shown in (d).  $\varphi(x)$  corresponds to the gauge vector potential  $A_z$  as shown by red arrows lying in the top surface, the curvature of which leads to a magnetic field  $B_y$  in the yellow shaded region with net flux  $2\pi$ . Due to the half quantized Hall conductance of the  $\theta$  domain wall, such a magnetic flux induces a half charge.

$$j^\mu = \frac{1}{8\pi} \epsilon^{\mu\nu\tau} \hat{\mathbf{n}} \cdot \partial_\mu \hat{\mathbf{n}} \times \partial_\nu \hat{\mathbf{n}}. \quad (118)$$

Equation (118) describes a topological response where the charge density and current are equal to the skyrmion density and current, respectively. Such an equation can be considered as a (2+1)-dimensional version of Goldstone-Wilczek formula (31), which has been studied extensively in the literature.<sup>13,66–69</sup> Thus, through the discussion above we have shown that the topological response formula (118) of (2+1)-dimensional Dirac fermions is a special example of the generic response equation [Eq. (116)].

To understand the physics described by Eq. (116), consider the lattice Dirac model in Eq. (108) with an adiabatic time evolution of  $\varphi(t) = 2\pi t/T$  and a spatial domain wall configuration of  $\theta(\vec{x})$ ,

$$\theta(y) = \frac{\pi}{2} \left[ 1 + \tanh\left(\frac{y}{\xi}\right) \right],$$

as shown in Fig. 17(a). According to Eq. (116), the charge current along the domain wall is given by

$$j_x = \frac{1}{2\pi} (\partial_y \Omega_t - \partial_t \Omega_y).$$

When the parameter  $\varphi$  evolves adiabatically from 0 to  $2\pi$ , the net charge flowing across the line  $x=0$  is given by  $\int dt I_x = \int dt dy (\partial_y \Omega_t - \partial_t \Omega_y) / 2\pi = \int_0^\pi d\theta \int_0^{2\pi} d\varphi (\partial_\theta \Omega_\varphi - \partial_\varphi \Omega_\theta) / 2\pi$ , which is the integration of the second Chern form  $\text{Tr}[\epsilon^{ABCD} f_{AB} f_{CD}] / 32\pi^2$  over the parameter range  $\theta \in [0, \pi]$ ,  $\varphi \in [0, 2\pi]$ , where  $A, B, C, D = k_x, k_y, \theta, \varphi$ . According to the discussion in the (3+1)-dimensional case, we know that a

magnetoelectric polarization  $P_3$  can be defined as

$$P_3(\theta) = \int d^2k d\varphi \mathcal{K}^\theta = - \int d\varphi \Omega_\varphi / 2\pi,$$

which implies  $\int dt I_x = - \int_0^\pi d\theta \partial_\theta P_3(\theta)$ . For  $-4c < m < -2c$ , we have  $\int_0^\pi dP_3(\theta) = 1/2$ , corresponding to the pumped charge  $\Delta Q = \int dt I_x = 1/2$ . In summary, the physical consequence of the topological response equation [Eq. (116)] is to induce a topological charge pumping effect during the adiabatic evolution of  $\varphi$ , in which the charge pumped in one period is proportional to the spatial gradient of the magnetoelectric polarization  $P_3$  defined in  $(k_x, k_y, \varphi)$  space. Specifically, a charge  $e/2$  is pumped along each  $\Delta\theta = \pi$  domain wall of  $\theta$  when  $\varphi$  evolves from 0 to  $2\pi$ .

Such a charge pumping effect can also be viewed as the dimensional reduction of the half QH effect on the  $\theta$  domain wall of (3+1)-dimensional lattice Dirac model, which has been studied in Sec. IV B. This dimensional reduction procedure is in exact analogy with the usual reduction from the (2+1)-dimensional QH effect to (1+1)-dimensional quantized pumping effect studied in Sec. II C. Similar to the latter case, a fractional charge effect can also be proposed in (2+1) dimensions according to Eq. (116). To show this effect, one can consider the same  $\theta$  domain wall as shown above, and a  $2\pi$  domain wall of  $\varphi$  along the  $y$  direction  $\varphi(\mathbf{x}) = \pi[1 + \tanh(x/\xi)]$ , as shown in Fig. 17(b). The charge density is given by  $\rho = (\partial_x \Omega_y - \partial_y \Omega_x) / 2\pi$ . By integrating over the  $x$  direction we obtain  $\int \Omega_x dx = \int \Omega_y d\varphi = -2\pi P_3(\theta)$ , such that  $\rho_{1D} = \int dx \rho = \partial_y P_3(\theta)$  and  $\int dy \rho_{1D} = 1/2$ . Thus, a half charge is localized at the crossing of  $\theta$  and  $\varphi$  domain walls.<sup>70</sup> Such a fractional charge effect can also be understood through the dimensional reduction from (3+1) dimensions. The spatial dependence of  $\varphi(x)$  corresponds to the spatial dependence of  $k_z - A_z(x)$ , which describes a magnetic field perpendicular to the 2D domain wall in the (3+1)-dimensional system. When  $\varphi(x)$  has a  $2\pi$  domain wall, the net flux of the corresponding magnetic field is  $2\pi$ , which thus induces a half charge as shown in Fig. 17(b).

In summary, we have studied the physical consequences of the topological effective action (115) in a spatially and/or temporally inhomogeneous insulator. In the rest of this section we will show how to define a  $Z_2$  topological invariant in (2+1)-dimensional TRI insulators and study the physical properties of the  $Z_2$  nontrivial phase—QSH phase—by applying effective theory (115).

### B. $Z_2$ topological classification of TRI insulators

In Secs. II D and II E we have shown how a  $Z_2$  classification of particle-hole invariant insulators can be defined in both (1+1) dimensions and (0+1) dimensions through dimensional reduction from the (2+1)-dimensional QH effect. The second-Chern-class analogy of the (2+1)-dimensional QH effect is the (4+1)-dimensional QH effect<sup>14</sup> described by the Chern-Simons theory (52), which then leads to the  $Z_2$  classification of TRI insulators in (3+1)-dimensional, as shown in Sec. IV C. Following this line of reasoning, it is straightforward to see that a  $Z_2$  classification can be defined for (2+1)-dimensional TRI insulators, as an analog of (0

+1)-dimensional particle-hole symmetric insulators. In this section we will sketch the demonstration of such a topological classification without going into detail since the derivation here is exactly parallel to that in Sec. II E.

First of all, for two TRI (2+1)-dimensional insulators  $h_1(\mathbf{k})$  and  $h_2(\mathbf{k})$  an adiabatic interpolation  $h(\mathbf{k}, \theta)$  can be defined, satisfying

$$h(\mathbf{k}, 0) = h_1, \quad h(\mathbf{k}, \pi) = h_2,$$

$$T^\dagger h(-\mathbf{k}, -\theta) T = h^T(\mathbf{k}, \theta). \quad (119)$$

Since  $h(\mathbf{k}, \theta)$  corresponds to the Hamiltonian of a (3+1)-dimensional TRI insulator, a  $Z_2$  topological quantity  $N_3[h(\mathbf{k}, \theta)] = \pm 1$  can be defined as shown in Sec. IV C. The key point to defining a  $Z_2$  invariant for the (2+1)-dimensional Hamiltonians  $h_1, h_2$  is to demonstrate the independence of  $N_3[h(\mathbf{k}, \theta)]$  on the choice of  $h(\mathbf{k}, \theta)$ . Consider two different parametrizations  $h(\mathbf{k}, \theta)$  and  $h'(\mathbf{k}, \theta)$ . An interpolation  $g(\mathbf{k}, \theta, \varphi)$  can be defined between them which satisfies

$$g(\mathbf{k}, \theta, 0) = h(\mathbf{k}, \theta), \quad g(\mathbf{k}, \theta, \pi) = h'(\mathbf{k}, \theta),$$

$$g(\mathbf{k}, 0, \varphi) = h_1(\mathbf{k}), \quad g(\mathbf{k}, \pi, \varphi) = h_2(\mathbf{k}),$$

$$g^T(\mathbf{k}, \theta, \varphi) = T^\dagger g(-\mathbf{k}, -\theta, -\varphi) T.$$

Here,  $g(\mathbf{k}, \theta, \varphi)$  corresponds to a (4+1)-dimensional insulator Hamiltonian, for which a second Chern number  $C_2[g]$  is defined. By its definition, the ‘‘second Chern parity’’  $N_3$  of  $h(\mathbf{k}, \theta)$  and  $h'(\mathbf{k}, \theta)$  satisfies  $N_3[h]N_3[h'] = (-1)^{C_2[g]}$ . At the same time,  $g(\mathbf{k}, \theta, \varphi)$  can also be considered as an interpolation between  $\theta=0$  and  $\theta=\pi$  systems, i.e., between  $g(\mathbf{k}, 0, \varphi) \equiv h_1(\mathbf{k})$  and  $g(\mathbf{k}, \pi, \varphi) \equiv h_2(\mathbf{k})$ . Since  $h_{1,2}(\mathbf{k})$  are both independent of  $\varphi$ , the  $\varphi$  component of Berry’s phase gauge field vanishes for  $g(\mathbf{k}, 0, \varphi)$  and  $g(\mathbf{k}, \pi, \varphi)$ . Consequently, it can be shown that  $(-1)^{C_2[g]} = N_3[g(\mathbf{k}, 0, \varphi)] = N_3[g(\mathbf{k}, \pi, \varphi)] = 1$ , so that  $N_3[h]N_3[h'] = 1$  for any two interpolations  $h$  and  $h'$ . Thus, we have shown that the  $Z_2$  quantity  $N_2[h_1(\mathbf{k}), h_2(\mathbf{k})] \equiv N_3[h(\mathbf{k}, \theta)]$  only depends on the (2+1)-dimensional Hamiltonians  $h_1$  and  $h_2$ . By defining a constant Hamiltonian  $h_0(\mathbf{k}) = h_0$  as reference, all (2+1)-dimensional TRI insulators are classified by the value of  $N_2[h_0, h(\mathbf{k})]$ . An insulator with  $N_2[h_0, h] = -1$  cannot be adiabatically deformed to the trivial Hamiltonian  $h_0$  without breaking time-reversal symmetry.

In the next section, the physical properties of the  $Z_2$  nontrivial insulator defined here will be studied. We will see that the  $Z_2$  nontrivial insulator defined here has nontrivial edge dynamics and corresponds to the QSH insulator studied in the literature.<sup>16,17,20,21</sup> Compared to the former definition of the  $Z_2$  topological classification,<sup>16,23,26</sup> our definition has the advantage of providing a direct relationship between the topological quantum number and the physical response properties of the system.

### C. Physical properties of the $Z_2$ nontrivial insulators

Similar to the (3+1)-dimensional case, the topological properties of a  $Z_2$  nontrivial insulator lead to nontrivial edge

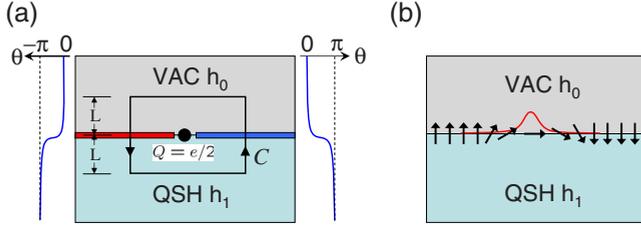


FIG. 18. (Color online) (a) Illustration of an interface between the vacuum (VAC) with Hamiltonian  $h_0(\mathbf{k})$  and a QSH insulator (QSH) with Hamiltonian  $h_1(\mathbf{k})$ . An interpolation  $h(\mathbf{k}, \theta)$ ,  $\theta \in [0, 2\pi]$  can be defined between  $h_0$  and  $h_1$ . On the left (right) half of the interface marked by red (blue),  $\theta$  has a domain wall from 0 to  $-\pi(\pi)$ . A fractional charge  $Q=e/2$  is trapped on the domain wall between the left and right half of the interface, which can be calculated by an integration of  $\Omega$  along the loop  $L$  (see text). (b) Physical realization of the domain wall between two interfaces in (a) by an antiphase domain wall of magnetic field or magnetization. The red curve shows schematically the charge-density distribution.

state dynamics described by effective theory (115) or, equivalently, the response equation [Eq. (116)]. The edge of a  $Z_2$  nontrivial insulator is equivalent to a domain wall between a nontrivial insulator and a trivial insulator (since the vacuum can be considered as a trivial insulator with a large gap). Thus, in the following we will focus on the domain wall between a nontrivial system with Hamiltonian  $h_1(\mathbf{k})$  and a trivial system with Hamiltonian  $h_0$ .

As discussed in the last section, an interpolation  $h(\mathbf{k}, \theta)$  can be defined between  $h_0$  and  $h_1$  satisfying  $h(\mathbf{k}, 0)=h_0$ ,  $h(\mathbf{k}, \pi)=h_1(\mathbf{k})$  and  $T^\dagger h(-\mathbf{k}, -\theta)T=h^T(\mathbf{k}, \theta)$ . Since  $h_1$  is nontrivial,  $h(\mathbf{k}, \theta)$  has to break time-reversal symmetry for general  $\theta$  to adiabatically connect  $h_1$  to  $h_0$ . Making use of  $h(\mathbf{k}, \theta)$ , two different interfaces between  $h_1$  and  $h_0$  can be defined. Consider a spatially dependent  $\theta$  given by

$$\theta(x, y) = \frac{\pi}{2} \left[ 1 - \tanh\left(\frac{y}{\xi}\right) \right].$$

Then the spatially dependent Hamiltonians  $h[\mathbf{k}, \theta(y)]$  and  $h[\mathbf{k}, -\theta(y)]$  both describe a spatial domain wall between  $h_1$  (for  $y \ll -\xi$ ) and  $h_0$  (for  $y \gg \xi$ ). The only differences between these two Hamiltonians are the time-reversal symmetry breaking terms around the interface. Now consider a more complicated interface, with

$$h(\mathbf{k}, \mathbf{x}) = \begin{cases} h[\mathbf{k}, \theta(y)], & x < 0 \\ h[\mathbf{k}, -\theta(y)], & x > 0, \end{cases}$$

as shown in Fig. 18(a). In such a system, the time-reversal symmetry on the interface is broken in *opposite* ways for  $x > 0$  and  $x < 0$  in the sense that  $h^T[\mathbf{k}, (x, y)] = T^\dagger h[-\mathbf{k}, (-x, y)]T$ . Now we study the charge localized around the point  $x=0, y=0$ . For a loop  $C$  enclosing this point as shown in Fig. 18(a), the charge in the region  $A$  enclosed by  $C$  is given by Eq. (116) as

$$Q = \frac{1}{2\pi} \int_A d^2x (\partial_x \Omega_y - \partial_y \Omega_x) = \frac{1}{2\pi} \oint_C \Omega \cdot d\mathbf{l}.$$

When the size of the loop is large enough compared to the boundary width  $\xi$ , such a loop integration is equivalent to an integration over  $\theta$  from 0 to  $2\pi$ , which leads to  $Q = \oint \Omega_\theta d\theta / 2\pi = P_3[h(\mathbf{k}, \theta)]$ . According to the definition of a  $Z_2$  nontrivial insulator in the last section,  $P_3[h(\mathbf{k}, \theta)] = 1/2 \bmod 1$  for any interpolation  $h(\mathbf{k}, \theta)$  between  $h_0$  and  $h_1$ . Consequently, the charge confined on the domain wall is  $Q = (n+1/2)e$  with  $n$  as an integer depending on the details of the interface.<sup>70</sup>

To summarize, a time-reversal symmetry breaking term can be applied at the interface of trivial and nontrivial insulators. For a given interface described by Hamiltonian  $h(\mathbf{k}, y)$ , its time-reversal partner,  $h'(\mathbf{k}, y) = Th^T(-\mathbf{k}, y)T^\dagger$ , describes a different connecting condition at the interface. If the 1D interface is described by  $h$  in one region and by  $h'$  in another region, then the domain wall between these two regions will trap a half-charge as a consequence of the nontrivial topology. To understand such a domain wall better, we can consider the case with a magnetic field as the time-reversal symmetry breaking term on the interface. When the magnetic field has an antiphase domain wall as shown in Fig. 18(b), a half-charge must be trapped on the domain wall.

Moreover, one can also obtain the distribution of 1D charge density and current density on the interface by integrating Eq. (116) only along the  $y$  direction,

$$\rho_{1D}(x) = \frac{1}{2\pi} \int_{-L}^L dy (\partial_x \Omega_y - \partial_y \Omega_x),$$

$$j_{1D}(x) = \frac{1}{2\pi} \int_{-L}^L dy (\partial_y \Omega_t - \partial_t \Omega_y),$$

in which  $L$  is a cutoff in the  $y$  direction, satisfying  $L \gg \xi$  so that the contribution to  $\rho_{1D}$  and  $j_{1D}$  from the region  $|y| > L$  is negligible. According to definitions (112) and (113) of the effective gauge vector potential  $\Omega_\mu$ , we know that  $\Omega_\mu(x, y, t) \rightarrow 0$  for a point deep in the QSH or VAC region, i.e., when  $|y| \rightarrow \infty$ . Thus the expression of 1D density and current can be simplified to

$$\rho_{1D}(x, t) = \partial_x P_3(x, t), \quad j_{1D}(x, t) = -\partial_t P_3(x, t) \quad (120)$$

with  $P_3(x, t) = \int_{-L}^L dy \Omega_y(x, y, t) / 2\pi$  as the magnetoelectric polarization defined in  $(k_x, k_y, y)$  space. Equation (120) is exactly the Goldstone-Wilczek formula<sup>13</sup> describing the charge fractionalization effect in the (1+1)-dimensional Dirac model, and  $2\pi P_3(x, t)$  plays the role of the phase angle of the Dirac mass term. When the interfaces on the left and right sides of the domain wall are related by time-reversal symmetry, the change in  $2\pi P_3$  through the domain wall must be  $(2n+1)\pi$ , which gives the half charge on the domain wall.

Such a relation between the interface and the (1+1)-dimensional Dirac model can be understood more intuitively in terms of the edge effective theory.<sup>70</sup> Similar to the relation between the edge theory of a (4+1)-dimensional topological insulator and the reduced edge theory of a (3+1)-dimensional  $Z_2$  nontrivial insulator, we can obtain the

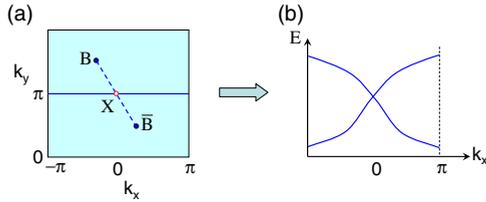


FIG. 19. (Color online) Illustration of the dimensional reduction from the surface of (a) a (3+1)-dimensional topological insulator to the edge of (b) a (2+1)-dimensional QSH insulator. The red hollow and blue filled circles in (a) are the positions of gapless (2+1)-dimensional Dirac cones in the surface BZ. The blue horizontal line at  $k_y = \pi$  defines the edge theory of a QSH insulator, as shown in (b).

edge theory of a (2+1)-dimensional  $Z_2$  nontrivial insulator from that of a (3+1)-dimensional nontrivial insulator. The interpolation  $h(\mathbf{k}, \theta)$  between  $h_1$  and  $h_0$  can be viewed as the Hamiltonian of a (3+1)-dimensional TRI insulator, in which  $\theta$  plays the role of  $k_z$ . Consider a specific point of the (2+1)-dimensional boundary of the (3+1)-dimensional system, say  $y=0$ . According to the discussion in Sec. IV C, an odd number of (2+1)-dimensional Dirac fermions are propagating on the boundary of the (3+1)-dimensional system. Due to time-reversal symmetry there must be an odd number of Dirac cones on the four time-reversal symmetric points, as shown in Fig. 12(b). For a slice at  $y=0$ , the wave vector of the surface state is  $(k_x, \theta)$ . Consequently, when  $\theta$  is considered to be a parameter, the surface energy spectrum  $E(k_x, \theta)$  for a given  $\theta$  describes the dispersion of (1+1)-dimensional edge states of (2+1)-dimensional insulators. Specifically,  $\theta = 0$  corresponds to the vacuum Hamiltonian  $h_0$ , which cannot support any nontrivial edge states. Thus, the Dirac cones at the time-reversal symmetric points can only appear at  $(k_x, \theta) = (0, \pi)$  and  $(\pi, \pi)$ . To have the minimal *odd* number of Dirac cones in the  $(k_x, \theta)$  BZ, there must be one gapless Dirac cone at  $(0, \pi)$  or  $(\pi, \pi)$ , but not both, as shown in Fig. 19. In other words, the edge effective theory of a (2+1)-dimensional nontrivial insulator is given by a gapless Dirac theory,

$$H = \int \frac{dk}{2\pi} v \sum_{\sigma=\pm 1} \sigma k \psi_{k\sigma}^\dagger \psi_{k\sigma}, \quad (121)$$

where  $\sigma = \pm 1$  means left and right movers, respectively. This edge theory agrees with the former descriptions of QSH edge states<sup>18,19,26</sup> and shows the equivalence of the  $Z_2$  nontrivial insulator defined in this section and the QSH insulator.

Once the edge theory is obtained, it is easy to understand the charge fractionalization proposed above. Due to the Kramers degeneracy, any TRI perturbation cannot open a gap at the edge. Only when a magnetic field or other time-reversal symmetry breaking term is applied can a mass term  $m_x \sigma_x + m_y \sigma_y$  be generated in the edge theory. Time-reversal symmetry also guarantees that the mass induced by opposite magnetic fields is exactly opposite. This implies that an antiphase domain wall of the magnetic-field corresponds to a sign change in the mass of the Dirac fermion. Thus, the edge state theory is described by the well-known Jackiw-Rebbi

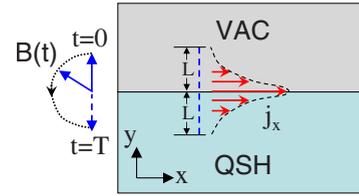


FIG. 20. (Color online) Illustration of quantized charge pumping on the boundary of a QSH insulator induced by a rotating magnetic field. During the time period  $t \in [0, T]$  the magnetic field is rotated from  $B$  to  $-B$ , and a half charge  $Q = \int dt \int_{-L}^L dy j_x = e/2$  is pumped along  $x$  direction.  $L$  is taken as a cutoff with  $j_x$  vanishingly small for  $|y| > L$ .

model<sup>40</sup> or equivalently the Su-Schrieffer-Heeger model.<sup>9</sup> The study of the fractional charge in the edge theory approach and its experimental consequences has been presented in Ref. 70. Thus, we have seen that the effective theory we obtained from dimensional reduction agrees with the edge theory analysis, just like in the (3+1)-dimensional case. The effective theory correctly describes the half-charge associated with a magnetic domain wall, which is a direct physical manifestation of the  $Z_2$  quantum number.

A quantized charge pumping effect always accompanies a fractional charge effect and can be realized when a *time-dependent* T-breaking field is applied at the edge. If the system is described by a time-dependent Hamiltonian  $h(\mathbf{k}, \mathbf{x}, t)$  which satisfies  $h(\mathbf{k}, \mathbf{x}, t=0) = h[\mathbf{k}, \theta(y)]$  and  $h(\mathbf{k}, \mathbf{x}, t=T) = h[\mathbf{k}, -\theta(y)]$  with  $\theta(y)$  as the domain wall configuration discussed above, then the charge pumped through the interface during the time  $t \in [0, T]$  is given by

$$Q_{\text{pump}} = \int_0^T dt j_{\text{ID}} = -[P_3(T) - P_3(0)] = -\left(n + \frac{1}{2}\right).$$

In the example of an applied magnetic field, such a pumping process implies that a half charge is pumped when a magnetic field rotates adiabatically from  $\mathbf{B}$  to  $-\mathbf{B}$ , as shown in Fig. 20. The experimental proposal of such a charge pumping effect is also discussed in Ref. 70.

Besides providing a quantized response property of the QSH insulators, the fractional charge and charge pumping effects proposed here are a (1+1)-dimensional version of electromagnetic duality. In (3+1)-dimensions the electromagnetic duality gives rise to the Witten effect,<sup>62</sup> where a magnetic monopole carries a charge  $\Theta/2\pi$  and becomes a “dyon” when a topological  $\Theta$  term is introduced in the Lagrangian.<sup>62,71</sup> Such an effect can occur in a (3+1)-dimensional topological insulator where the charge of the dyon is  $1/2$ , as studied in Sec. IV B. In comparison, the magnetic domain wall on the boundary of (2+1)-dimensional QSH insulator can be considered as a topological point defect of magnetic field in (1+1) dimensions, which also carries a half-charge. In this sense, if we consider the magnetic domain wall as a dynamical degree of freedom of the system (e.g., when the magnetic domain wall is generated by a ferromagnetic stripe on top of the 2D QSH system), it can be considered as the (1+1)-dimensional manifestation of dyons.

TABLE I. Summary on the properties of the topological insulators. In the effective Lagrangians, the indices  $\mu, \nu, \dots = 0, 1, \dots, d$  with  $d$  as the spatial dimension. The vector  $A_\mu$  stands for the gauge vector of the external electromagnetic field and  $a_\mu$  that for the Berry phase gauge field.  $C$  and  $T$  stands for particle-hole symmetry and time-reversal symmetry, respectively. The quantities  $C_1, P_1, C_2, P_3$ , and  $\Omega_\rho$  are defined by Eqs. (5), (24), (54), (79), and (112) respectively (see text of Sec. VI for explanations of the effective actions).

	Dimension	Topological quantum number	Effective theory	Symmetry requirement	Physical properties
Family 1	2+1	First Chern number $C_1 \in \mathbb{Z}$	$\mathcal{L} = \frac{C_1}{2\pi} A_\mu \varepsilon^{\mu\nu\rho} \partial_\nu A_\rho$	Not required	QH effect; axial anomaly on the boundary
	1+1	First Chern parity $N_1 = (-1)^{C_1[h_0, h]} \in \mathbb{Z}_2$	$\mathcal{L} = P_1 \varepsilon^{\mu\nu} \partial_\mu A_\nu$	$C$	Half charge on the boundary
	0+1	First Chern parity $N_0 = N_1[h_0, h] \in \mathbb{Z}_2$	$\mathcal{L} = \text{Tr}[a_0]$	$C$	Not applicable
Family 2	4+1	Second Chern number $C_2 \in \mathbb{Z}$	$\mathcal{L} = \frac{C_2}{24\pi^2} A_\mu \varepsilon^{\mu\nu\rho\sigma\tau} \partial_\nu A_\rho \partial_\sigma A_\tau$	Not required	4DQH effect; chiral anomaly on the boundary
	3+1	Second Chern parity $N_3 = (-1)^{C_2[h_0, h]} \in \mathbb{Z}_2$	$\mathcal{L} = \frac{1}{4\pi} P_3 \varepsilon^{\mu\nu\sigma\tau} \partial_\mu A_\nu \partial_\sigma A_\tau$	$T$	Half QH effect on the boundary, TME effect
	2+1	Second Chern parity $N_2 = N_3[h_0, h] \in \mathbb{Z}_2$	$\mathcal{L} = \frac{1}{2\pi} A_\mu \varepsilon^{\mu\nu\rho} \partial_\nu \Omega_\rho$	$T$	QSH effect; half charge at antiphase domain wall on the boundary

Interestingly, such an analogy can also be generalized to (2+1) dimensions, where the topological defect of magnetic field is a flux tube. Recently it has been shown that a  $\pi$  flux tube threaded into a QSH insulator carries either charge  $\pm e$ , spin 0 or charge 0, spin 1/2, where the spin 0 (1/2) is generically defined as a Kramers singlet (doublet) under time-reversal symmetry.<sup>31,32</sup> In other words, the  $\pi$  flux tube becomes a dyonlike object and realizes spin-charge separation in (2+1) dimensions. Such a spin-charge separation phenomenon also provides an alternative definition of the  $\mathbb{Z}_2$  topological insulators in (2+1) dimensions.<sup>31</sup>

In summary, we have shown how a  $\mathbb{Z}_2$  classification of (2+1)-dimensional TRI insulators is obtained and how the physical properties of the  $\mathbb{Z}_2$  nontrivial insulator are described by the effective theory derived from dimensional reduction. Together with the discussion of (3+1)-dimensional topological insulators in Sec. IV, we have seen that the nontrivial topology and its consequences in both (3+1)-dimensional and (2+1)-dimensional TRI systems have their origin in the nontrivial second Chern number in (4+1) dimensions. The dimensional reduction series (4+1)dimensions  $\rightarrow$  (3+1)dimensions  $\rightarrow$  (2+1) dimensions is in exact analogy of the lower dimensional one (2+1)dimensions  $\rightarrow$  (1+1)dimensions  $\rightarrow$  (0+1) dimensions. In the next section, we will develop the unified framework of dimensional reduction in generic dimensions, which contains the two series as simplest examples.

## VI. UNIFIED THEORY OF TOPOLOGICAL INSULATORS

### A. Phase space Chern-Simons theories

Up to now, we have systematically studied several related topological phenomena, including the (2+1)-dimensional QH insulator with nontrivial first Chern number, the (4+1)-dimensional topological insulator with nontrivial second Chern number, and their dimensional reductions. Com-

paring Sec. II with Secs. III–V one can easily see the exact analogy between the two series of topological insulators: the (2+1)-dimensional and (4+1)-dimensional *fundamental* topological insulators are characterized by an integer—the first and second Chern numbers, respectively. Under a discrete symmetry [particle-hole symmetry for the (2+1)-dimensional family and time-reversal symmetry for the (4+1)-dimensional family], a  $\mathbb{Z}_2$  topological classification can be defined for the lower dimensional descendent systems, the physical properties of which can be described by effective theories obtained from the dimensional reduction procedure. The main facts about these topological phenomena are summarized in Table I. In this section, we will show that the effective theories for all these systems share a universal form when written in phase space.

As a simple example, we first consider effective theory (9) of the QH effect. Expanding the expression of the first Chern number explicitly, Eq. (9) can be expressed as

$$S = \frac{1}{4\pi} \int \frac{d^2k}{2\pi} \epsilon^{ij} \text{Tr}[\partial_i a_j] \int dt d^2x \epsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho, \quad (122)$$

in which  $i, j$  are indices 1, 2 in momentum space and  $\mu, \nu, \rho = 0, 1, 2$  are space-time indices. Here and below,  $A_\mu$  and  $a_i$  stand for the external electromagnetic gauge field in real space and Berry's phase gauge field in momentum space, respectively. The trace is carried over all occupied energy levels. If we define the phase-space coordinate as  $\mathbf{q} = (t, x, y, k_x, k_y)$  and the phase-space gauge potentials  $\mathbf{A} = (A_0, A_1, A_2, 0, 0)$ ,  $\mathbf{a} = (0, 0, 0, a_1, a_2)$ , then the action above is equivalent to the following second Chern-Simons term:

$$S_{2+1} = \frac{1}{8\pi^2} \int d^5q \epsilon^{ABCDE} A_A \partial_B A_C \text{Tr}[\partial_D a_E], \quad (123)$$

where all capital roman indices, e.g.,  $A, B, C, \dots$ , run over the appropriate phase space coordinates. Since in this system  $\mathbf{A}$

and  $\mathbf{a}$  are always orthogonal to each other, such a reformulation seems trivial. However, it turns out to be helpful when considering the dimensional reduction procedure. As discussed earlier, dimensional reduction in the (2+1)-dimensional system to (1+1) dimensions is defined by replacing, say,  $k_y + A_y$  by a parameter  $\theta(x, t)$ , which in general is space-time dependent. Four changes are induced by this substitution: (1)  $A_y$  is replaced by  $\theta(x, t) - \theta_0$  with  $\theta_0$  playing the role of  $k_y$ ; (2)  $\partial/\partial k_y$  is replaced by  $\partial/\partial\theta$ ; (3) Berry's phase gauge field  $a_k^{\alpha\beta} = -i\langle\alpha; \mathbf{k}|\partial_{k_y}|\beta; \mathbf{k}\rangle$  is replaced by  $a_\theta^{\alpha\beta} = -i\langle\alpha; k_x, \theta|\partial_\theta|\beta; k_x, \theta\rangle$ ; and (4) The integrations  $\int dy$  and  $\int dk_y/2\pi$  are removed from the effective action.

By making these substitutions, the effective action of the (1+1)-dimensional system can be obtained. To help us understand the general dimensional reduction procedure, we show the derivation of the (1+1)-dimensional effective theory explicitly. For simplicity, one can start from Eq. (122). Note that the gauge field  $A_\mu$  depends only on the (1+1)-dimensional coordinates  $(t, x)$  after dimensional reduction. Consequently, in the Chern-Simons form  $\varepsilon^{\mu\nu\tau} A_\mu \partial_\nu A_\tau$  the terms containing  $\partial_y$  are identically zero, so that

$$\varepsilon^{\mu\nu\tau} A_\mu \partial_\nu A_\tau = 2(A_t \partial_x A_y - A_x \partial_t A_y) = 2(A_t \partial_x \theta - A_x \partial_t \theta)$$

after an integration by parts. Effective action (122) after making all four substitutions above is expressed as

$$S_{1+1} = \frac{1}{2\pi} \int dk_x \text{Tr}[\partial_{k_x} a_\theta - \partial_\theta a_{k_x}] \int dt dx (A_t \partial_x \theta - A_x \partial_t \theta).$$

With a smooth space-time dependent  $\theta(x, t)$  field, the eigenstates  $|\alpha; k_x, \theta\rangle$  can be considered as space-time dependent local eigenstates, whose space-time dependence originates only from that of  $\theta(x, t)$ . In this way, each state  $|\alpha; k_x, \theta(x, t)\rangle = |\alpha; k_x, x, t\rangle$  is defined in the full phase space  $(t, x, k_x)$ , and Berry's phase gauge field can gain real-space components defined as  $a_\mu^{\alpha\beta} = -i\langle\alpha; k_x, x, t|\partial_\mu|\beta; k_x, x, t\rangle = a_\theta \partial_\mu \theta$ , in which  $\mu = t, x$ . Similarly, the space-time derivative of  $a_{k_x}$  is given by  $\partial_\mu a_{k_x} = \partial_\theta a_{k_x} \partial_\mu \theta$ . By making use of these observations, the effective action can be simplified to

$$S_{1+1} = \frac{1}{2\pi} \int dk_x \int dt dx \{A_t \text{Tr}[\partial_{k_x} a_x - \partial_x a_{k_x}] - A_x \text{Tr}[\partial_{k_x} a_t - \partial_t a_{k_x}]\}. \quad (124)$$

By generalizing the definition of the gauge vector potential  $(A_t, A_x)$  to the phase space vector  $(A_t, A_x, 0)$ , the equation above can be expressed as the mixed Chern-Simons term in the phase space,

$$S_{1+1} = \frac{1}{2\pi} \int d^3q \varepsilon^{ABC} A_A \text{Tr}[\partial_B a_C], \quad (125)$$

which describes an inhomogeneous (1+1)-dimensional insulator. Note that momentum derivatives acting on  $(A_t, A_x)$  vanish. This effective action agrees our discussion in Sec. II C as one can see by taking  $\frac{\delta S_{1+1}}{\delta A_A}$  and comparing the resulting response equations with Eq. (28).

It should be noted that such a phase space formalism is only applicable when the space-time variation in  $\theta$  is smooth and can be approximated by a constant in the neighborhood of a space-time point  $(t, x, y)$ . More quantitatively, the characteristic frequency  $\omega$  and wave vector  $k$  of  $\theta(x, y, t)$  should satisfy

$$\hbar\omega, \quad \hbar vk \ll E_g, \quad (126)$$

where  $v$  is a typical velocity scale of the system. For example, in the lattice Dirac model (12),  $v$  is the speed of light [which is normalized to be 1 in Eq. (12)]. Under condition (126) the space-time variation in the  $\theta$  field does not generate excitations across the gap, and the system can be viewed locally as a band insulator with a "local Hamiltonian"  $h(k_x; x, t)$ .

Carrying out such a procedure once more to action (125) one can obtain the (0+1)-dimensional action. We will show the derivation explicitly. First one must take  $a_x = 0$ ,  $\partial_x a_{k_x} = 0$  in Eq. (124) since nothing can depend on the spatial  $x$  coordinate after it is dimensionally reduced. Next we replace  $A_x$  by the parameter  $\phi(t) - \phi_0$ , which leads to

$$\begin{aligned} S_{0+1} &= - \int dt (\phi - \phi_0) \text{Tr}[\partial_\phi a_t - \partial_t a_\phi] \\ &= - \int dt \text{Tr}(\partial_t \phi \tilde{a}_\phi + \tilde{a}_t), \end{aligned}$$

where the integration  $\int dx dk_x/2\pi$  has been removed and  $a_t, a_{k_x}$  in (1+1) dimensions are relabeled as  $\tilde{a}_t, \tilde{a}_\phi$  for later convenience and to obtain the second equality, an integration by parts is carried out. It should be noted that  $\partial_\phi$  comes from  $\partial/\partial k_y$ , which actually means  $\partial/\partial\phi_0$  since it is  $\phi_0$  that is replacing  $k_y$ . Compared to the (2+1)-dimensions  $\rightarrow$  (1+1) dimensions reduction, the difference here is that the wave functions are, in general, already time dependent in (1+1) dimensions. This comes from the dimensional reduction from (2+1) dimensions. Consequently, Berry's phase gauge potential in the (0+1)-dimensional system consists of two terms as shown below:

$$\begin{aligned} a_t^{\alpha\beta} &= -i\langle\alpha; t, \phi(t)|\partial_t|\beta; t, \phi(t)\rangle \\ &= -i\langle\alpha; t, \phi|\left[\left(\frac{\partial}{\partial t}\right)_\phi + \frac{\partial\phi}{\partial t}\left(\frac{\partial}{\partial\phi}\right)_t\right]|\beta; t, \phi\rangle = \tilde{a}_t + \partial_t \phi \tilde{a}_\phi, \end{aligned}$$

in which  $(\partial/\partial t)_\phi$  means to take the  $t$  derivative while keeping  $\phi$  constant. Both of these terms are necessary for the correct topological response and similar terms (including spatially dependent ones) will be present in all higher dimensions when more than one reduction is carried out. Combining the two equations above we finally obtain

$$S_{0+1} = - \int dt \text{Tr}[a_t], \quad (127)$$

which has the form of a "zeroth" Chern-Simons term and describes the "(0+1)-dimensional insulator" discussed in Sec. II E, i.e., a single-site fermion system. In this case the only gauge-invariant quantity remaining is Berry's phase the

single-site system obtains during a closed path of adiabatic evolution.

For the second family of topological insulators we discussed, effective theory (52) can be expressed in the following phase space form:

$$S_{4+1} = \frac{1}{192\pi^4} \int d^9 q \epsilon^{ABCDEFGH} A_A \partial_B A_C \partial_D A_E \text{Tr}[D_F a_G D_H a_I], \quad (128)$$

in which the covariant derivative  $D_B = \partial_B + ia_B$  is introduced for the non-Abelian Berry phase gauge field. The dimensional reduction to (3+1) dimensions can be performed similarly to the (2+1)-dimensional case. Denoting the fourth spatial dimension as  $w$ , which is the dimension to be reduced, then any term with  $\partial_w$  vanishes and so does  $a_w$ . Consequently, the only nonvanishing terms in effective action (128) are those with  $A_w$ , which now is replaced by the parameter  $\theta(x, y, z, t)$ . On the other hand, one of  $F, G, H, I$  in the form  $D_F a_G D_H a_I$  has to be  $k_w$ , which is now replaced by  $\theta$ . In summary the theory can be rewritten as

$$S_{3+1} = \frac{3}{96\pi^3} \int d^7 q \epsilon^{\mu\nu\sigma\tau} \epsilon^{ijk} A_\mu \partial_\nu A_\sigma \partial_\tau \theta \text{Tr}[D_\theta a_i D_j a_k + \text{cycl.}]$$

where  $\mu, \nu, \dots = 0, 1, 2, 3$ ,  $i, j, k = 1, 2, 3$  are space-time and momentum indices of (3+1)-dimensional system, and cycl. denotes the other three terms obtained from cyclicly permuting  $\theta$  and  $i, j, k$ . The integration  $\int dk_w dw / 2\pi$  has been removed and a prefactor 3 appears due to the fact that there are three  $A_A$ 's in effective action (128). In the same way as in (2+1)-dimensional to (1+1)-dimensional case,  $\partial_\mu \theta D_\theta$  can be replaced by  $D_\mu$ , so that the effective theory of the (3+1)-dimensional insulator is finally obtained,

$$S_{3+1} = \int \frac{d^7 q}{32\pi^3} \epsilon^{AB\dots G} A_A \partial_B A_C \text{Tr}[D_D a_E D_F a_G]. \quad (129)$$

According to the definition of  $P_3$  in Eq. (79) we know that

$$\partial_\ell P_3 = \frac{1}{8\pi^2} \int d^3 k \epsilon^{\theta ijk} \text{Tr}[f_\ell i f_{jk}],$$

which shows the equivalence of action (129) to action (80) we derived earlier.

Now from the two examples of (2+1) dimensions  $\rightarrow$  (1+1) dimensions  $\rightarrow$  (0+1) dimensions and (4+1) dimensions  $\rightarrow$  (3+1) dimensions  $\rightarrow$  (2+1) dimensions, one can easily obtain the general rule of dimensional reduction to the phase-space Chern-Simons theories. For a  $(d+1)$ -dimensional system, the phase space dimension is  $2d+1$ , and the dimensional reduction in corresponding phase-space Chern-Simons theory is defined as follows.

(1) Remove a term  $\partial_A A_B$  from the action and correspondingly replace the  $(2d+1)$ -dimensional antisymmetric tensor by the one in  $(2d-1)$  dimensions.

(2) Remove the integration  $\int dx_d dk_d / 2\pi$  from the action when  $x_d, k_d$  are the spatial and momentum indices to be reduced.

(3) Multiply the action by a factor  $n$  when the power of external gauge field  $A_A$  in the original action is  $n$ .

Following these rules, the effective action for (2+1)-dimensional TRI insulator can be easily obtained by one more step of dimensional reduction from Eq. (129),

$$S_{2+1} = \int \frac{d^5 q}{8\pi^2} \epsilon^{ABCDE} A_A \text{Tr}[D_B a_C D_D a_E], \quad (130)$$

where the coefficient is determined by  $8\pi^2 = 32\pi^3 / (2 \cdot 2\pi)$ . By considering the space-time and momentum indices separately, one can easily confirm that Eq. (130) is equivalent to Eq. (115) we obtained earlier.

In summary, we have shown that all the known topological insulators are described by a Chern-Simons effective theory in phase space, and the topological theories in different dimensions can be related by the dimensional reduction procedure. It is straightforward to generalize this formalism to arbitrary dimensions. As shown in Appendix C, the relation between the nonlinear-response function (53) and the corresponding Chern number in momentum space can be generically proven for any odd space-time dimension. The effective theory of such a  $(2n+1)$ -dimensional topological insulator is given by<sup>41</sup>

$$S_{2n+1} = \frac{C_n}{(n+1)!(2\pi)^n} \int d^{2n+1} x \epsilon^{\mu_1 \mu_2 \dots \mu_{2n+1}} \times A_{\mu_1} \partial_{\mu_2} A_{\mu_3} \dots \partial_{\mu_{2n}} A_{\mu_{2n+1}} \quad (131)$$

with the  $n$ th Chern number in momentum space defined as

$$C_n = \frac{1}{n!2^n(2\pi)^n} \int d^{2n} k \epsilon^{i_1 i_2 \dots i_{2n}} \text{Tr}[f_{i_1 i_2} f_{i_3 i_4} \dots f_{i_{2n-1} i_{2n}}]. \quad (132)$$

Thus the  $(4n+1)$  dimensional phase space formula for this effective action can be written as

$$S_{2n+1} = \frac{1}{n!(n+1)!(2\pi)^{2n}} \int d^{4n+1} q \epsilon^{A_1 A_2 \dots A_{4n+1}} A_{A_1} \times \partial_{A_2} A_{A_3} \dots \partial_{A_{2n}} A_{A_{2n+1}} \text{Tr}[D_{A_{2n+2}} a_{A_{2n+3}} \dots D_{A_{4n}} a_{A_{4n+1}}]. \quad (133)$$

Following the general rules of dimensional reduction procedure discussed above, one can obtain the effective actions for lower-dimensional topological insulators as ‘‘descendants’’ of topological theory (133). From the examples discussed above it can be seen that the number of Berry's phase gauge vectors  $a_i$  in the effective action remains invariant during dimensional reduction, while the number of external gauge-field insertions  $A_\mu$  decreases by one at each step of dimensional reduction. After  $n+1$  reductions we obtain an effective action of an  $[(n-1)+1]$ -dimensional system that contains no  $A_i$ . Just like in the (0+1)-dimensional case, such an effective action does not result in any response equation of the system but only describes Berry's phase the system obtains during adiabatic evolution. Obviously the dimensional reduction cannot be carried out again on such an  $[(n-1)+1]$ -dimensional system. Thus, the  $(2n+1)$ -dimensional topological insulator with a nontrivial  $n$ th Chern number only has  $(n+1)$  descendants under dimensional reduction. It

is straightforward to show that the  $m$ th descendant ( $1 \leq m \leq n+1$ ) of the  $(2n+1)$ -dimensional topological insulator has the effective action,

$$\begin{aligned}
 S_{2n+1-m}^{(m)} &= \frac{\binom{2n+1-m}{n}}{(2n+1-m)!} \int \frac{d^{4n+1-2m}q}{(2\pi)^{2n-m}} \epsilon^{A_1 A_2 \dots A_{4n+1-2m}} \\
 &\quad \times A_{A_1} \partial_{A_2} A_{A_3} \dots \partial_{A_{2n-2m}} A_{A_{2n+1-2m}} \\
 &\quad \times \text{Tr} [D_{A_{2n+2-2m}} a_{A_{2n+3-2m}} \dots D_{A_{4n-2m}} a_{A_{4n+1-2m}}] \\
 &\equiv \text{CS}_{4n-2m+1}^{n-m+1}, \quad (134)
 \end{aligned}$$

in which  $\text{CS}_s^t$  stands for the mixed Chern-Simons action in  $s$  phase-space dimensions with  $t$  powers of the external  $A_A$  field. Specifically,  $\text{CS}_s^0$  is a pure non-Abelian Chern-Simons term of Berry's phase gauge field  $a_A$ , which cannot be reduced to a function of  $f_{AB} = D_A a_B - D_B a_A$  alone. Thus we have seen that a whole family of topological phenomena are described by phase-space Chern-Simons theories with different  $s, t$  values. In a given spatial dimension  $d$ , all possible topological phenomena in band insulators are given by the actions  $\text{CS}_{2d+1}^t$  with all possible values of  $t$ . It should be noted that the external gauge field  $A_A$  is only defined in real space, meaning that  $A_A$  and  $\partial_B A_A$  are both vanishing if  $A$  or  $B$  is a momentum index  $A, B = d+2, d+3, \dots, 2d+1$ . Since in  $\text{CS}_{2d+1}^t$  there are  $t$   $A_A$ 's and at least  $t-1$  partial derivative operators acting on  $A_A$ 's, the Chern-Simons action  $\text{CS}_{2d+1}^t$  vanishes if  $2t-1 > d+1$ . Consequently, there are in total  $[d/2]+2$  available Chern-Simons terms in the phase space of a  $(d+1)$ -dimensional system, which are

$$\text{CS}_{2d+1}^t, t = 0, 1, \dots, [d/2] + 1.$$

Here  $[d/2]$  denotes the maximal integer that does not exceed  $d/2$ . For example, in  $(2+1)$  dimensions there are three available phase space Chern-Simons terms, two of which are  $\text{CS}_5^2$  describing a QH insulator and  $\text{CS}_5^1$  describing a QSH insulator. The third one is given by

$$\begin{aligned}
 \text{CS}_5^0 &= \frac{1}{3!(2\pi)^2} \int d^5 q \epsilon^{ABCDE} \text{Tr} \left[ a_A \partial_B a_C \partial_D a_E \right. \\
 &\quad \left. + \frac{3}{2} a_A a_B a_C \partial_D a_E + \frac{3}{5} a_A a_B a_C a_D a_E \right],
 \end{aligned}$$

which contains no  $A_A$  and thus does not describe any electromagnetic response properties of the system. The information contained in the effective action  $\text{CS}_5^0$  is Berry's phase the system obtains during adiabatic evolution, just like the effective action of the  $(0+1)$ -dimensional system  $\text{CS}_1^0 = \int dt a_0$ . We have grouped the phase-space Chern-Simons theories based on the parent theories and their descendants. The relationships are summarized in a "family tree" in Fig. 21. Similar to the generalization of the  $(2+1)$ -dimensional QH insulator to any odd space-time dimension, the  $Z_2$  topological insulators we have studied can also be generalized to higher dimensions, which will be explained in the next section.

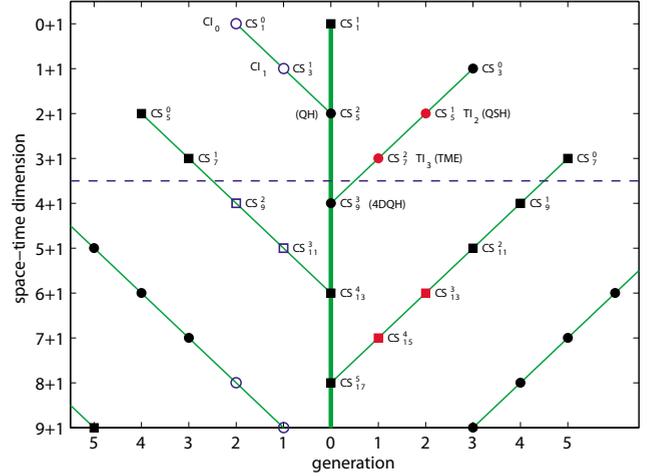


FIG. 21. (Color online) The family tree of topological insulators. The black points on the "trunk" (i.e., zeroth generation) stand for the fundamental topological insulators in odd space-time dimensions characterized by a nontrivial Chern number in momentum space. The blue and red markers show the descendants of the  $(4n-1)$ -dimensional and  $(4n+1)$ -dimensional insulators, respectively. Physical effects associated with some of the prominent topological insulators are indicated in parenthesis. A  $Z_2$  classification is defined for each blue hollow circle (square) under (pseudo-) particle-hole symmetry  $C$  ( $\tilde{C}$ ), and for each red filled circle (square) under (pseudo-) time-reversal symmetry  $T$  ( $\tilde{T}$ ). The definitions of  $C$ ,  $\tilde{C}$ ,  $T$ , and  $\tilde{T}$  are given in Sec. VI B. For all the physically realizable systems with spatial dimensions  $d \leq 3$ , the names of topological insulators are labeled, where  $\text{CI}_n$  ( $\text{TI}_n$ ) stands for a particle-hole symmetric (TRI) topological insulator in  $n+1$  dimension. The black circles and squares stand for other topological phenomena obtained from dimensional reduction, which are also described by the phase space Chern-Simons theories but do not correspond to  $Z_2$  topological insulators. The phase space Chern-Simons theory  $\text{CS}_{2n+1}^t$  [as defined in Eq. (134)] corresponding to each topological phenomenon is also specified on the figure.

Before moving on to that, we would like to point out an interesting mathematical fact about the phase-space Chern-Simons theories. For an  $(n+1)$  dimensional system with  $N$  occupied bands, a  $U(N)$  gauge vector potential can be defined in phase space as

$$\mathcal{A}_A^\lambda = \lambda A_A + a_A,$$

with  $A_A$  being the external gauge potential and  $a_A$  being the Berry phase gauge potential. The non-Abelian Chern-Simons term for  $\mathcal{A}_A^\lambda$  can be expressed as

$$\begin{aligned}
 \text{CS}_{2n+1}(\lambda) &\equiv \text{CS}_{2n+1}[\mathcal{A}_A^\lambda] \\
 &= \frac{1}{(n+1)!(2\pi)^n} \int d^{2n+1}q \epsilon^{A_1 A_2 \dots A_{2n+1}} \\
 &\quad \times \text{Tr} [\mathcal{A}_{A_1}^\lambda \partial_{A_2} \mathcal{A}_{A_3}^\lambda \dots \partial_{A_{2n}} \mathcal{A}_{A_{2n+1}}^\lambda + \text{N.T.}], \quad (135)
 \end{aligned}$$

Here N.T. stands for the non-Abelian terms containing commutators of  $\mathcal{A}_A^\lambda$ , which can be determined by the relation of

the  $(2d+1)$ -dimensional *Chern-Simons* form to the  $[(2d+1)+1]$ -dimensional *Chern* form. For more details, see Sec. 11.5 of Ref. 72. By expanding the parametrized action  $CS_{2n+1}(\lambda)$  over  $\lambda$ , the following equality can be obtained:

$$\begin{aligned} CS_{2n+1}(\lambda) &= \sum_{t=0}^{[n/2]+1} \lambda^t CS'_{2n+1} \Rightarrow CS'_{2n+1} \\ &= \frac{1}{t!} \left. \frac{\partial^t}{\partial \lambda^t} CS_{2n+1}(\lambda) \right|_{\lambda=0}. \end{aligned} \quad (136)$$

This implies that all possible phase space Chern-Simons terms can be obtained from a single “generating functional”  $CS_{2n+1}(\lambda)$ . We present Eq. (135) as the unified theory of all topological insulators.

### B. $Z_2$ topological insulator in generic dimensions

For the descendants of the  $(2+1)$ -dimensional and  $(4+1)$ -dimensional insulators, we have defined a  $Z_2$  classification under the constraint of a discrete symmetry. For the descendants of the  $(2+1)$ -dimensional QH insulator, the  $Z_2$  classification is defined for particle-hole symmetric insulators satisfying Eq. (32),

$$C^\dagger h(-\mathbf{k})C = -h^T(\mathbf{k}), \quad C^\dagger C = C^*C = \mathbb{I}. \quad (137)$$

The key point of this classification is to show that an interpolation between two particle-hole symmetric Hamiltonians  $h_1(k)$  and  $h_2(k)$  forms a closed path when combined with its particle-hole transformed partner. The Chern number enclosed in such a closed path always has a certain parity which does not depend on the choice of the path. In the same way, a  $Z_2$  classification of particle-hole symmetric insulators is also defined in  $(0+1)$  dimensions. For the family of  $(4+1)$ -dimensional insulators it is the same story except that the particle-hole symmetry is replaced by time-reversal symmetry,

$$T^\dagger h(-\mathbf{k})T = h^T(\mathbf{k}), \quad T^\dagger T = -T^*T = \mathbb{I}. \quad (138)$$

Following this one can easily generalize such  $Z_2$  classifications to higher dimensions. To do that, one first needs to understand what is the difference between  $(2+1)$  and  $(4+1)$  dimensions that requires the choice of different discrete symmetries. The easiest way to see such a difference is to study the transformation of the corresponding Chern-Simons theories under particle-hole symmetry ( $C$ ) and time-reversal symmetry ( $T$ ). Under particle-hole symmetry, the charge density and charge current both change sign. The vector potential does as well, as required by the invariance of the minimal coupling  $A_\mu j^\mu$ . In the same way one can obtain the time-reversal property of  $A_\mu$ , as summarized below:

$$C:A_\mu \rightarrow -A_\mu, \quad T:A_\mu \rightarrow \begin{cases} A_0 \\ -A_j \end{cases} \quad (139)$$

In both cases of  $C$  and  $T$ , the momentum operator  $-i\partial_\mu$  has the same transformation property as  $A_\mu$ . Based on these facts the transformation properties of Chern-Simons Lagrangian (131) are

$$C:S_{2n+1}^{\text{CS}} \rightarrow (-1)^{n+1} S_{2n+1}^{\text{CS}},$$

$$T:S_{2n+1}^{\text{CS}} \rightarrow (-1)^n S_{2n+1}^{\text{CS}}. \quad (140)$$

Thus, we see that  $S_{4n+1}^{\text{CS}}$  is  $T$  even but  $C$  odd, while  $S_{4n-1}^{\text{CS}}$  is  $T$  odd but  $C$  even. In other words, a  $(4n+1)$ -dimensional topological insulator has to break particle-hole symmetry but can be time-reversal invariant, just like the case of  $(4+1)$  dimensions; a  $[(4n-2)+1]$ -dimensional topological insulator has to break time-reversal symmetry but can be particle-hole symmetric, just like the case of  $(2+1)$  dimensions. Consequently, for the descendants of  $(4n+1)$ -dimensional  $[(4n-1)$ -dimensional] topological insulators, it is only possible to define  $Z_2$  topological classifications by the dimensional reduction procedure under the constraint of  $T(C)$  symmetry.

Naively, it seems that the procedure we introduced to define the  $Z_2$  classification by dimensional reduction could be applied recursively to all the descendants of a  $(2n+1)$ -dimensional topological insulator. However, this turns out to be incorrect. As an example, we can study the  $(1+1)$ -dimensional TRI insulator as a descendant of the  $(2+1)$ -dimensional QSH insulator [not the  $(2+1)$ -dimensional QH insulator]. In the dimensional reduction from  $(3+1)$  dimensions to  $(2+1)$  dimensions discussed in Sec. V B, we define an interpolation  $h(\mathbf{k}, \theta)$  between two  $(2+1)$ -dimensional TRI Hamiltonians  $h_1(\mathbf{k})$  and  $h_2(\mathbf{k})$ . When the interpolation  $h(\mathbf{k}, \theta)$  is required to satisfy time-reversal symmetry [Eq. (119)], it corresponds to the Hamiltonian of a  $(3+1)$ -dimensional topological insulator, for which a  $Z_2$  index  $N_3[h(\mathbf{k}, \theta)]$  can be defined. In Sec. V B we have shown that  $N_3[h(\mathbf{k}, \theta)]$  does not depend on the choice of the interpolation  $h(\mathbf{k}, \theta)$ , which thus provides a criteria on whether  $h_1(\mathbf{k})$  and  $h_2(\mathbf{k})$  are topologically equivalent. If we carry out the same procedure on  $(1+1)$ -dimensional TRI insulators, it seems that for two Hamiltonians  $h_1(k)$  and  $h_2(k)$  a  $Z_2$  topological classification can be defined in the same way. To see if this is true, one can again take the lattice Dirac model as an example. The single-particle Hamiltonian of  $(2+1)$ -dimensional  $4 \times 4$  lattice Dirac model with time-reversal symmetry is written as

$$h_{2D}(\mathbf{k}) = \Gamma^1 \sin k_x + \Gamma^2 \sin k_y + \Gamma^0 [m + c(\cos k_x + \cos k_y)], \quad (141)$$

which is in the topological nontrivial phase for  $0 < m < 2|c|$  or  $-2|c| < m < 0$ . By dimensional reduction,  $h_{2D}(\mathbf{k})$  can be considered as the interpolation between two  $(1+1)$ -dimensional TRI Hamiltonians  $h_1(k) = h_{2D}(k, 0)$  and  $h_2(k) = h_{2D}(k, \pi)$ . Thus if the  $Z_2$  classification procedure applied to  $(1+1)$ -dimensional systems,  $h_1(k)$  and  $h_2(k)$  should be topologically distinct when  $h_{2D}(\mathbf{k})$  is in the nontrivial phase. In other words, it should be impossible to define another interpolation  $h_0(k, \theta)$  between  $h_1(k)$  and  $h_2(k)$ , which satisfies  $T^\dagger h_0(k, \theta)T = h_0^*(-k, -\theta) = h_0^*(-k, \theta)$ ,  $\forall \theta$ . However, such an interpolation can, in fact, be constructed as follows:

$$h_0(k, \theta) = \Gamma^1 \sin k + \Gamma^{02} \sin^2 \theta + \Gamma^0 (m + c \cos k_x + c \cos \theta), \quad (142)$$

in which  $\Gamma^{02} = i\Gamma^0\Gamma^2$  is even under time reversal. The existence of two topologically distinct interpolations  $h_0(k, \theta)$  and  $h_{2D}(k, \theta)$  shows that it is not possible to define a  $Z_2$  classifi-

cation of (1+1)-dimensional TRI insulators in the same way as in (2+1) dimensions and (3+1) dimensions. The main reason for the failure is that the proof in Sec. V B requires the parametrized Hamiltonian manifold to be *simply connected*. In other words, an interpolation  $h(\mathbf{k}, \theta, \varphi)$  can always be defined for two interpolations  $h(\mathbf{k}, \theta)$  and  $h'(\mathbf{k}, \theta)$ . Similar arguments do not work in the classification of (1+1)-dimensional Hamiltonians because it may not be possible to adiabatically deform one path to the other. In the example of the lattice Dirac model, the paths  $h_{2D}$  and  $h_0$  cannot be adiabatically connected because the combined path

$$g(k, \theta) = \begin{cases} h_{2D}(k, \theta), & \theta \in [0, \pi] \\ h_0(k, 2\pi - \theta), & \theta \in [\pi, 2\pi] \end{cases}$$

is a (2+1)-dimensional Hamiltonian that breaks time-reversal symmetry and has a nontrivial first Chern number  $C_1 = -1$ . Consequently, the path  $g$  cannot be contracted to a point, and the definition of a path-independent  $Z_2$  invariant fails.

From this example we have seen that the definition of a  $Z_2$  topological classification for the descendants of (2n+1)-dimensional topological insulators fails when the dimension is reduced to  $[(2n-3)+1]$  dimensions, since the lower Chern number  $C_{n-1}$  is defined for each closed path of  $[(2n-3)+1]$ -dimensional Hamiltonians, thus obstructing the adiabatic connection between two different paths. In other words, the  $Z_2$  topological insulators as descendants of (2n+1)-dimensional topological insulators can only be defined in  $[(2n-1)+1]$  and  $[(2n-2)+1]$  dimensions. Since a bulk topological insulator always corresponds to a topologically protected gapless edge theory, the validity of a  $Z_2$  classification can also be justified by studying the stability of edge theories. As discussed in Sec. III B, the boundary theory of a (4+1)-dimensional topological insulator with second Chern number  $C_2$  contains  $|C_2|$  flavors of chiral (Weyl) fermions. In the simplest example of the lattice Dirac model (61) with  $-4c < m < -2c$ , the boundary single-particle Hamiltonian is

$$H_{\partial(4+1)} = v \vec{\sigma} \cdot \vec{\mathbf{p}},$$

which is topologically stable since no mass term is available for the edge system. Under dimensional reduction the boundary theory of a (3+1)-dimensional  $Z_2$  nontrivial insulator is simply given by taking  $p_z = 0$  in the above equation,

$$H_{\partial(3+1)} = v(\sigma_x p_x + \sigma_y p_y),$$

which is stable *in the presence of time-reversal symmetry* since no T-invariant mass terms are available. The same analysis shows the stability of the edge theory of a (2+1)-dimensional topological insulator, given by  $H_{\partial(2+1)} = v \sigma_z p_z$ . When dimensional reduction is carried out once more, we obtain the (0+1)-dimensional edge of the (1+1)-dimensional insulator described by  $H_{\partial(1+1)} = 0$ . This just describes a Kramers pair of localized states on the boundary. Since such a pair of midgap states can be easily removed by a constant energy shift without breaking time-reversal symmetry, the (1+1)-dimensional TRI insulator does not have a topologically nontrivial class. Different edge

state stabilities for effective theories in different dimensions are illustrated in Fig. 21.

Such an edge theory analysis can be easily generalized to higher dimensions. The boundary states of a  $[2n+1]$ -dimensional topological insulator with nontrivial Chern number are described by a  $[(2n-1)+1]$ -dimensional chiral fermion theory with the Hamiltonian

$$H_{2n-1}(\mathbf{p}) = v \sum_{i=1}^{2n-1} p_i \Gamma^i, \quad (143)$$

in which  $\Gamma^i$  are  $2^{n-1} \times 2^{n-1}$  matrices forming a representation of the  $so(2n-1)$  Clifford algebra. The boundary theory of the  $m$ th descendant of the (2n+1)-dimensional system is given by simply taking  $p_i = 0$  for  $i = 2n-m, 2n-m+1, \dots, 2n-1$ . The symmetry properties and stability of the theory can be studied by studying the properties of the  $\Gamma^a$  matrices. Here we will display the conclusions of the edge state analysis, with the details presented in Appendix E.

(1) The Chiral Hamiltonians [Eq. (143)] in different dimensions satisfy different discrete symmetry properties, as listed below:

$$C^\dagger H_{2n-1}(\mathbf{p}) C = -H_{2n-1}(-\mathbf{p}), \quad C^* C = \mathbb{I}, \quad n = 4m - 3, \\ m \in \mathbb{N},$$

$$T^\dagger H_{2n-1}(\mathbf{p}) T = H_{2n-1}(-\mathbf{p}), \quad T^* T = -\mathbb{I}, \quad n = 4m - 2, \\ m \in \mathbb{N},$$

$$\tilde{C}^\dagger H_{2n-1}(\mathbf{p}) \tilde{C} = -H_{2n-1}(-\mathbf{p}), \quad \tilde{C}^* \tilde{C} = -\mathbb{I}, \quad n = 4m - 1, \\ m \in \mathbb{N}$$

$$\tilde{T}^\dagger H_{2n-1}(\mathbf{p}) \tilde{T} = H_{2n-1}(-\mathbf{p}), \quad \tilde{T}^* \tilde{T} = \mathbb{I}, \quad n = 4m, \quad m \in \mathbb{N}. \quad (144)$$

(2) Only the first two descendants of  $H_{2n-1}$ , i.e.,  $H(\mathbf{p}) = \sum_{i=1}^{2n-2} p_i \Gamma^i$  and  $H(\mathbf{p}) = \sum_{i=1}^{2n-3} p_i \Gamma^i$ , are topologically stable under the constraint of the discrete symmetry in the given dimension ( $C, \tilde{C}, T$ , or  $\tilde{T}$ ). Consequently, the  $Z_2$  topologically nontrivial insulators descending from the (2n+1)-dimensional topological insulator only exist in  $[(2n-1)+1]$  dimensions and  $[(2n-2)+1]$  dimensions.

The edge state analysis confirms our insight from the bulk picture, that is, the  $Z_2$  topological classification is only well defined for the first two generations of descendants of the (2n+1)-dimensional topological insulator. Moreover, it also provides more information about the discrete symmetries in different dimensions. In the (6+1)-dimensional topological insulator and its descendants, the correct discrete symmetry is  $\tilde{C}$  which is similar to particle-hole symmetry  $C$  but satisfies  $\tilde{C}^* \tilde{C} = -\mathbb{I}$ . This is necessary since a usual particle-hole symmetry cannot be defined for the (5+1)-dimensional chiral fermion  $\sum_{i=1}^5 p_i \Gamma^i$ . Such a symmetry  $\tilde{C}$  can be called a ‘‘pseudo-particle-hole symmetry.’’ Similarly, in the

(8+1)-dimensional topological insulator and its descendants, the discrete symmetry is a “pseudo-time-reversal symmetry” satisfying  $T^*T=I$ . In Fig. 21, the dimensions with “true”  $C$  or  $T$  symmetry are labeled with filled circles and those with  $\tilde{C}$  or  $\tilde{T}$  symmetry are labelled with squares.

This paper is partly inspired by work in high-energy physics. The study of topological insulators with nontrivial boundary states is analogous to the generation of massless fermions on higher dimensional domain walls.<sup>44,73</sup> In a high energy context these surface states would be subsequently used to mimic chiral fermions in lattice gauge theories. In our case they become the gapless boundary liquids that generate novel transport properties and characterize the topological stability of the state. Our picture of a (4+1)-dimensional topological insulator characterized by the Chern-Simons term has a very special meaning in high-energy physics when reduced to (3+1) dimensions. The Chern-Simons term becomes a  $\theta$  term which is related to the so-called vacuum angle. If  $\theta$  is a constant then this term does not contribute to the equations of motion. In our case the  $\theta$  term has a solitonic structure with a domain wall at the surface of the topological insulator. Inside the insulator  $\theta$  jumps by  $\pi$  which still preserves CP if the original vacuum does. Thus, the only effect on the system is a nonzero boundary term at the  $\theta$  domain wall. From a high-energy perspective we have introduced an axionic domain wall and a topological insulator exists in one domain while a trivial insulator exists in the other.

The dimensional reduction procedure we introduced here is not new to physics and was first used in the 1920s in an attempt to unify gravity and electromagnetism in (4+1) dimensions by Kaluza and Klein.<sup>74,75</sup> Basically, the dimensional reduction amounts to compactifying the “extra dimension” with periodic boundary conditions, e.g., on a circle, and shrinking its radius to zero. The compactification creates a tower of modes labeled by a discrete index, but as the circle is shrunk only one low-energy mode from each field remains (which can be seen by taking the Fourier transform of a field on a circle with periodic boundary conditions). These “zero modes” become the propagating fields in the lower dimensional space. Carrying out this procedure via our method or by compactification yields the same results. The adiabatic parameter we introduced is connected with the flux threading, the higher dimensional circle. One could imagine compactifying using higher dimensional manifolds, such as a sphere or something more exotic, and threading various fluxes through nontrivial cycles of the compact space. The zero-mode structures of these manifolds are more complicated and we will not deal with them here, but perhaps other interesting theories can arise.

Additionally, we have unearthed a ladder of topological insulators with gapless fermionic boundary states whose stability depends on the presence of a discrete symmetry. The discrete symmetries repeat mod 8, which should be no surprise since their form is derived from the representation theory of real Clifford algebras which exhibit Bott periodicity with period 8.<sup>76</sup> Due to this periodicity we can analyze the types of gapless fermions allowed to exist at the boundaries of topological insulators. Any boundary can support

Dirac fermions and any even dimensional space-time boundary can have Weyl (chiral) fermions as well. For Majorana fermions we are restricted to boundaries with space-time dimensions  $\{(1+1), (2+1), (3+1), (7+1), (8+1)\} \bmod 8$ . Finally, there is a special representation, the Majorana-Weyl fermion, which is a real fermion with a definite handedness which can only exist in  $(1+1) \bmod 8$  dimensions, i.e.,  $(1+1), (9+1), (17+1), (25+1), \dots$ <sup>76</sup> It is this type of fermion which appears at the edge of  $(p+ip)$  superconductors.<sup>77-79</sup> However, because this representation is missing in (5+1) dimensions a (6+1)-dimensional “ $(p+ip)$  superconductor” which obeys the pseudo-particle-hole symmetry will not have single Majorana-Weyl boundary states. This can also be seen by the fact that  $\tilde{C}^2=-1$  which means that there would have to be at the minimum two Majorana-Weyl fermions at the boundary due to a Kramers-type theorem. Beginning with the parent  $(2n+1)$ -dimensional topological insulator we see that the boundary theories of itself and its stable descendants are massless fermions in  $[(2n-1)+1]$ ,  $[(2n-2)+1]$ , and  $[(2n-3)+1]$  dimensions. Theories of massless fermions often result in field theory anomalies and there is a deep connection between this boundary theory ladder and the corresponding anomaly ladder.<sup>72,80</sup>

In summary, we have provided a unified framework to describe a whole family of topological insulators in generic dimensions. All the topological effects are described by phase-space Chern-Simons theories, which either describe the topological insulators with nontrivial Chern number in odd space-time dimensions or describe their lower dimensional descendants through dimensional reduction.  $Z_2$  topologically nontrivial insulators exist in  $[(2n-1)+1]$  and  $[(2n-2)+1]$  space-time dimensions and are protected by a given discrete symmetry that is preserved by the parent  $(2n+1)$ -dimensional topological insulator. We found that in  $(2n+1)$  dimensions there are two types of topological insulators, one of which is characterized by the Chern number  $C_n$  and the other by the  $Z_2$  invariant, as a descendant of  $[(2n+2)+1]$  dimensional topological insulator. In comparison, in  $[(2n-1)+1]$  dimensions there is only one type of topological insulator, which is characterized by a  $Z_2$  invariant, as a descendant of  $(2n+1)$  dimensional topological insulator. There are many tantalizing connections of our work with well-developed sectors of high-energy physics, and diving deeper into these subjects is sure to benefit both condensed-matter and high-energy physics.

## VII. CONCLUSION AND DISCUSSIONS

In conclusion we have constructed the topological field theory of TRI insulators. We showed that the fundamental TRI insulator naturally exists in (4+1) dimensions, and the effective topological field theory is the Chern-Simons theory in (4+1) dimensions. We introduced the concept of dimensional reduction for microscopic fermion models, where some spatial dimensions are compactified and the associated momentum variables are replaced by adiabatic fields. This method enables us to obtain the topological field theory for the 3D and 2D TRI insulators from the dimensional reduction in the (4+1) dimensional Chern-Simons field theory. In

TABLE II. Table of conventions used in the paper.

Symbol	Explanation
$(n+1)$ dimension	Space-time dimension $(n+1)$
$nD$	Spatial dimension $n$
$\mu, \nu, \rho, \sigma, \dots$	Space-time or frequency-momentum indices $0, 1, \dots, n$
$i, j, k, l, \dots$	Spatial or momentum indices $1, 2, \dots, n$
$A, B, C, D, \dots$	Phase space indices $0, 1, \dots, n, n+1, \dots, 2n+1$
$a, b, c, d, \dots$	Indices for anticommuting $\Gamma^a$ matrices and corresponding coefficients $d_a$
$\alpha, \beta, \gamma, \delta, \dots$	Energy band indices
$A_\mu, F_{\mu\nu}$	Gauge vector potential and gauge curvature of external electromagnetic field
$\alpha_i, f_{ij}$	(Generally non-Abelian) Berry phase gauge vector potential and gauge curvature in momentum space
$A_A, a_A$	Electromagnetic or Berry phase gauge vector potential defined in phase space, respectively

particular, we obtain the ‘‘axion’’ field theory for the 3D insulator, and many experimental consequences follow directly from this topological field theory. The most striking prediction is the TME effect, where an electric field induces a magnetic field along the same direction, with a universal constant of proportionality quantized in odd multiples of the fine-structure constant. The role of the axion or the adiabatic field is played by a magnetoelectric polarization, whose change is quantized when an adiabatic process is completed. The topological field theory for the 2D TRI insulator involves two adiabatic fields, and this theory directly predicts the fractional charge of a magnetic domain wall at the edge of the QSH insulator. These topological effects illustrate the predictive power of the topological field theory constructed in this work, and many more experimental consequences can be obtained by the proper generalization of the concepts introduced here.

Our work also presents the general classification of topological insulators in various dimensions. The fundamental topological insulators are described by the topological Chern-Simons field theory, and the effective topological field theory of their descendants can be obtained by the procedure of dimensional reduction. The descendent topological insulators are generally classified by discrete symmetries such as the charge conjugation and the time reversal symmetries. This way, the Chern number classification of fundamental topological insulators and the  $Z_2$  classifications of their descendants are unified. Finally, we present a framework in terms of the Chern-Simons field theory in phase space, which gives a unified theory of all topological insulators and contains all experimentally observable topological effects.

#### ACKNOWLEDGMENTS

We would like to thank B. A. Bernevig for many insightful discussions on this subject and collaborations at the early stage of this project. We would like to thank H. D. Chen, S. Kachru, C. X. Liu, J. Maciejko, M. Mulligan, S. Raghu, S.

Shenker, and J. Zaanen for helpful discussions. This work is supported by the NSF under Grant No. DMR-0342832 and the US Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC03-76SF00515.

#### APPENDIX A: CONVENTIONS

Due to the special importance of dimensionality in this work we tried to be very consistent with our dimension and index conventions. In addition, gauge fields of all types appear and we have selected a convention for the electromagnetic vector potential and the adiabatic (Berry’s phase) connection. The conventions we used in this paper are summarized in Table II and also explained below.

For space-time conventions we have chosen the form  $(n+1)$  dimension where  $n$  can be 0. For spatial dimensions only we use  $nD$  with a capital D.

For indices, space-time and frequency-momentum indices share the same convention. Greek indices from the middle of the alphabet such as  $\mu, \nu, \rho, \sigma$ , and  $\tau$  run from 0 to the spatial dimension in the current context, examples being 0, 1, 0, 1, 2, etc. Additionally, when indexing momentum space objects 0 is frequency and 1, 2, 3, ... are the momentum components in the 1, 2, 3, ... directions. Latin indices from the middle of the alphabet such as  $i, j, k$ , and  $\ell$  are purely spatial indices and run from 1 onward to the spatial dimension of the current context. For momentum space objects they index the spatial momenta, e.g.,  $k_x, k_y, k_z, k_w, \dots$ . We always use the Einstein summation convention unless stated otherwise. Since we are considering flat space we make no distinction between raised and lowered indices.

In some special contexts we will need two more sets of indices. Several Hamiltonian models we use can be written in terms of the  $2 \times 2$  or  $4 \times 4$  Dirac matrices. In these cases where you see vectors  $d^a$  indexed by lowercase Latin letters from the beginning of the alphabet they run from 1, 2, 3 and 0, 1, 2, 3, 4, respectively. Finally, for cases where the indices do not just run over coordinate or momentum space separately

but instead cover phase-space coordinates, we use capital Latin letters from the beginning of the alphabet such as  $A$ ,  $B$ ,  $C$ ,  $D$ , and  $E$ . These run over the phase space variables in the current context. Some examples being  $q^A=(t,x,y,k_x,k_y)$ ,  $q^A=(t,x,k_x)$ , or  $A$  running over  $(k_x,k_y,\theta,\phi)$ . For orbital, band, or state labels we use Greek letters from the beginning of the alphabet such as  $\alpha$ ,  $\beta$ , and  $\gamma$ .

For the electromagnetic  $U(1)$  gauge field we use  $A^\mu$  where  $A$  is capitalized. For Berry's phase gauge field we use  $a^i$  with a lower case  $a$ . Note the different index labels. The electromagnetic gauge field has a zero component, while Berry's phase has no "frequency component." For the curvatures we use  $F_{\mu\nu}$  and  $f_{ij}$ , respectively.

### APPENDIX B: DERIVATION OF Eq. (36)

Since Eq. (36) is such an important result we provide a detailed derivation in this section. First we recall the follow-

ing results from Sec. II D for a particle-hole symmetric insulator,

$$E_{\bar{\alpha}}(-k, 2\pi - \theta) = -E_{\alpha}(k, \theta),$$

$$C|k, \theta; \alpha\rangle^* = C \sum_{\beta} (\langle m, \beta | k, \theta; \alpha \rangle)^* |m, \beta\rangle \equiv |-k, 2\pi - \theta; \bar{\alpha}\rangle,$$

$$C^\dagger C = 1.$$

Also, to make the derivation clearer we will make a distinction between the charge polarization  $P_{occ}$  calculated from the Berry phase of all the occupied bands and  $P_{unocc}$  the charge polarization calculated from the Berry phase of the unoccupied bands.

Following the notation introduced in Sec. II D we can perform the following manipulations:

$$\begin{aligned} P_{occ}(\theta) &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\alpha}(k, \theta) < 0} (-i) \langle k, \theta; \alpha | \partial_k | k, \theta; \alpha \rangle \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\alpha}(k, \theta) < 0} (-i) \left[ \sum_{m, n, \beta, \gamma} \langle m, \beta | \partial_k | n, \gamma \rangle \langle k, \theta; \alpha | m, \beta \rangle^* \langle n, \gamma | k, \theta; \alpha \rangle^* \right]^* \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\alpha}(k, \theta) < 0} (-i) \left[ \sum_{m, n, \beta, \gamma} \langle m, \beta | C^\dagger \partial_k C | n, \gamma \rangle \langle k, \theta; \alpha | m, \beta \rangle^* \langle n, \gamma | k, \theta; \alpha \rangle^* \right]^* \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\alpha}(k, \theta) < 0} (-i) [\langle -k, 2\pi - \theta; \bar{\alpha} | \partial_k | -k, 2\pi - \theta; \bar{\alpha} \rangle]^* \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\alpha}(k, \theta) < 0} i [\langle -k, 2\pi - \theta; \bar{\alpha} | \partial_k | -k, 2\pi - \theta; \bar{\alpha} \rangle] \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{-E_{\bar{\alpha}}(-k, 2\pi - \theta) < 0} i [\langle -k, 2\pi - \theta; \bar{\alpha} | \partial_k | -k, 2\pi - \theta; \bar{\alpha} \rangle] \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\bar{\alpha}}(-k, 2\pi - \theta) > 0} (-i) [\langle -k, 2\pi - \theta; \bar{\alpha} | \partial_k | -k, 2\pi - \theta; \bar{\alpha} \rangle] \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{E_{\bar{\alpha}}(k, 2\pi - \theta) > 0} (-i) [\langle k, 2\pi - \theta; \bar{\alpha} | \partial_k | k, 2\pi - \theta; \bar{\alpha} \rangle] \equiv P_{unocc}(2\pi - \theta). \end{aligned} \quad (B1)$$

On the other hand, we have

$$\begin{aligned} P_{occ}(\theta) + P_{unocc}(\theta) &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{\alpha} (-i) \langle k, \theta; \alpha | \partial_k | k, \theta; \alpha \rangle \\ &= \int_0^{2\pi} \frac{dk}{2\pi} \sum_{\alpha, \beta} (-i) \langle k, \theta; \alpha | m, \beta \rangle \partial_k \langle m, \beta | k, \theta; \alpha \rangle \\ &= \int_0^{2\pi} \frac{dk}{2\pi} (-i) \text{Tr}[u^{-1}(k) \partial_k u(k)] \\ &= \int_0^{2\pi} \frac{dk}{2\pi} (-i) \partial_k \text{Tr}[\log u(k)] \end{aligned} \quad (B2)$$

with  $u(k) \in U(N)$  as a unitary matrix defined by  $u_{\beta\alpha}(k) = \langle m, \beta | k, \theta; \alpha \rangle$ . Consequently,  $P_{occ} + P_{unocc}$  is an integer given by the winding number of the phase,

$$\varphi(k) = -i \text{Tr}[\log u(k)] = \text{Im} \log[\det u(k)]. \quad (\text{B3})$$

Thus we have the following relation:

$$P_{occ}(\theta) + P_{unocc}(\theta) = 0 \pmod{1}. \quad (\text{B4})$$

Combined with Eq. (B1) we obtain  $P_{occ}(\theta) = P_{unocc}(2\pi - \theta) = -P_{occ}(2\pi - \theta) \pmod{1}$ , which is what we wished to show.

### APPENDIX C: DERIVATION OF Eq. (54)

In this appendix we will prove conclusion (54). As is briefly sketched in Sec. III A, the demonstration consists of three steps: (1) topological invariance of Eq. (53); (2) any Hamiltonian  $h(\mathbf{k})$  is adiabatically connected to an  $h_0(\mathbf{k})$  with the form of Eq. (55); and (3) for such an  $h_0(\mathbf{k})$  the nonlinear

correlation function (53) is equal to the second Chern number. In the following we will demonstrate these three steps separately.

#### 1. Topological invariance of Eq. (53)

To prove the topological invariance of Eq. (53), we just need to prove that any infinitesimal deformation of Green's function  $G(\mathbf{k}, \omega)$  leads to a vanishing variation in  $C_2$ . Under a variation of  $G(\mathbf{k}, \omega)$  we have

$$\begin{aligned} \delta(G\partial_\mu G^{-1}) &= \delta G\partial_\mu G^{-1} + G\partial_\mu(\delta G^{-1}) \\ &= \delta G\partial_\mu G^{-1} - G\partial_\mu(G^{-1}\delta G G^{-1}) \\ &= -G(\partial_\mu G^{-1})\delta G G^{-1} - \partial_\mu(\delta G)G^{-1}. \end{aligned}$$

Thus the variation in  $C_2$  is

$$\begin{aligned} \delta C_2 &= -\frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \text{tr}[\delta(G\partial_\mu G^{-1})(G\partial_\nu G^{-1})(G\partial_\rho G^{-1})(G\partial_\sigma G^{-1})(G\partial_\tau G^{-1})] \\ &= \frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \text{tr}[(G\partial_\mu G^{-1}\delta G G^{-1})(G\partial_\nu G^{-1})(G\partial_\rho G^{-1})(G\partial_\sigma G^{-1})(G\partial_\tau G^{-1})] \\ &\quad + \frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \text{tr}[(\partial_\mu \delta G G^{-1})(G\partial_\nu G^{-1})(G\partial_\rho G^{-1})(G\partial_\sigma G^{-1})(G\partial_\tau G^{-1})] \\ &= \frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \text{tr}[\partial_\mu(G^{-1}\delta G)(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)] \\ &= \frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \partial_\mu \text{tr}[(G^{-1}\delta G)(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)] \\ &\quad - \frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \{ \text{tr}[(G^{-1}\delta G)\partial_\mu(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)] \\ &\quad + \text{tr}[(G^{-1}\delta G)(\partial_\nu G^{-1}G)\partial_\mu(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)] \\ &\quad + \text{tr}[(G^{-1}\delta G)(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)\partial_\mu(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)] \\ &\quad + \text{tr}[(G^{-1}\delta G)(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)\partial_\mu(\partial_\tau G^{-1}G)] \\ &\quad + \text{tr}[(G^{-1}\delta G)(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)\partial_\mu] \} \\ &= \frac{\pi^2}{15} \epsilon^{\mu\nu\rho\sigma\tau} \int \frac{d^4 k d\omega}{(2\pi)^5} \partial_\mu \text{tr}[(G^{-1}\delta G)(\partial_\nu G^{-1}G)(\partial_\rho G^{-1}G)(\partial_\sigma G^{-1}G)(\partial_\tau G^{-1}G)] \equiv 0. \quad (\text{C1}) \end{aligned}$$

Thus Eq. (53) is topologically invariant. The topological invariance of the second Chern number defined in Eq. (54) is a well-known mathematical fact. Such a topological invariance is quite helpful for showing the equivalence between the second Chern number and Eq. (53).

#### 2. Adiabatic deformation of arbitrary $h(\mathbf{k})$ to $h_0(\mathbf{k})$

An adiabatic deformation  $h(k, t), t \in [0, 1]$  can be written down, which connects an arbitrary gapped Hamiltonian  $h(\mathbf{k})$  to a "maximally degenerate Hamiltonian"  $h_0(\mathbf{k})$  in the form of Eq. (55). Any single particle Hamiltonian  $h(\mathbf{k})$  can be diagonalized as

$$h(\mathbf{k}) = U(\mathbf{k})D(\mathbf{k})U^\dagger(\mathbf{k})$$

with  $U(\mathbf{k})$  as unitary and  $h_0(\mathbf{k}) = \text{diag}[\epsilon_1(\mathbf{k}), \epsilon_2(\mathbf{k}), \dots, \epsilon_N(\mathbf{k})]$  as the diagonal matrix of energy eigenvalues. Without loss of generality, the chemical potential can be defined to be zero and the eigenvalues can be arranged in ascending order. For an insulator with  $M$  bands filled, one has

$$\epsilon_1(\mathbf{k}) \leq \epsilon_2(\mathbf{k}) \leq \dots \leq \epsilon_M(\mathbf{k}) < 0 < \epsilon_{M+1}(\mathbf{k}) \leq \dots \leq \epsilon_N(\mathbf{k}). \quad (\text{C2})$$

For  $t \in [0, 1]$ , defining

$$E_\alpha(\mathbf{k}, t) = \begin{cases} \epsilon_\alpha(\mathbf{k})(1-t) + \epsilon_G t, & 1 \leq \alpha \leq M \\ \epsilon_\alpha(\mathbf{k})(1-t) + \epsilon_E t, & M < \alpha \leq N \end{cases} \quad (\text{C3})$$

and  $D_0(\mathbf{k}, t) = \text{diag}[E_1(\mathbf{k}, t), E_2(\mathbf{k}, t), \dots, E_N(\mathbf{k}, t)]$ , then we have

$$D_0(\mathbf{k}, 0) = D(\mathbf{k}), \quad D_0(\mathbf{k}, 1) = \begin{pmatrix} \epsilon_G \mathbb{I}_{M \times M} & \\ & \epsilon_E \mathbb{I}_{N-M \times N-M} \end{pmatrix}.$$

As long as  $\epsilon_G < 0 < \epsilon_H$ ,  $D_0(\mathbf{k}, t)$  remains gapped for  $t \in [0, 1]$ . Thus by defining

$$h(\mathbf{k}, t) = U(\mathbf{k})D_0(\mathbf{k}, t)U^\dagger(\mathbf{k}), \quad (\text{C4})$$

we obtain an adiabatic interpolation between  $h(\mathbf{k}, 0) = h(\mathbf{k})$  and  $h(\mathbf{k}, 1) = U(\mathbf{k})D_0(\mathbf{k}, 1)U^\dagger(\mathbf{k})$ . Since the matrix  $U(\mathbf{k})$  can be written in the eigenstates of  $h(\mathbf{k})$  as  $U(\mathbf{k}) = (|1, \mathbf{k}\rangle, |2, \mathbf{k}\rangle, \dots, |N, \mathbf{k}\rangle)$ , we have

$$h(\mathbf{k}, 1) = \epsilon_G \sum_{\alpha=1}^M |\alpha, \mathbf{k}\rangle \langle \alpha, \mathbf{k}| + \epsilon_E \sum_{\beta=M+1}^N |\beta, \mathbf{k}\rangle \langle \beta, \mathbf{k}| = \epsilon_G P_G(\mathbf{k}) + \epsilon_E P_E(\mathbf{k}). \quad (\text{C5})$$

In summary, we have proven that each gapped Hamiltonian

$h(\mathbf{k})$  can be adiabatically connected to a Hamiltonian with the form of Eq. (55).

### 3. Calculation of correlation function (53) for $h_0(\mathbf{k})$

For Hamiltonian of form (C5) Green's function is written in the simple form,

$$G(\mathbf{k}, \omega) = [\omega + i\delta - \epsilon_G P_G(\mathbf{k}) - \epsilon_E P_E(\mathbf{k})]^{-1} = \frac{P_G(\mathbf{k})}{\omega + i\delta - \epsilon_G} + \frac{P_E(\mathbf{k})}{\omega + i\delta - \epsilon_E}. \quad (\text{C6})$$

On the other hand, we have

$$\frac{\partial G^{-1}(\mathbf{k}, \omega)}{\partial \omega} = 1,$$

$$\frac{\partial G^{-1}(\mathbf{k}, \omega)}{\partial k_i} = -\epsilon_G \frac{\partial P_G(\mathbf{k})}{\partial k_i} - \epsilon_E \frac{\partial P_E(\mathbf{k})}{\partial k_i} = (\epsilon_E - \epsilon_G) \frac{\partial P_G(\mathbf{k})}{\partial k_i}, \quad (\text{C7})$$

where  $i=1, 2, 3, 4$ . Thus Eq. (53) can be written as

$$C_2 = -\frac{\pi^2}{3} \epsilon^{ijkl} \int \frac{d^4 k d\omega}{(2\pi)^5} \sum_{n,m,s,t=1,2} \frac{\text{Tr} \left[ P_n \frac{\partial P_G}{\partial k_i} P_m \frac{\partial P_G}{\partial k_j} P_s \frac{\partial P_G}{\partial k_k} P_t \frac{\partial P_G}{\partial k_\ell} \right] (\epsilon_E - \epsilon_G)^4}{(\omega + i\delta - \epsilon_n)^2 (\omega + i\delta - \epsilon_m) (\omega + i\delta - \epsilon_s) (\omega + i\delta - \epsilon_t)}, \quad (\text{C8})$$

in which  $\epsilon_{1,2} = \epsilon_{G,E}$  and  $P_{1,2}(\mathbf{k}) = P_{G,E}(\mathbf{k})$ , respectively. From the identity  $P_E + P_G = \mathbb{I}$  and  $P_E^2 = P_E$ ,  $P_G^2 = P_G$ , we obtain

$$P_E \frac{\partial P_G}{\partial k_i} = -\frac{\partial P_E}{\partial k_i} P_G = \frac{\partial P_G}{\partial k_i} P_G, \\ P_G \frac{\partial P_G}{\partial k_i} = -P_G \frac{\partial P_E}{\partial k_i} = \frac{\partial P_G}{\partial k_i} P_E. \quad (\text{C9})$$

Consequently,  $P_G \partial_i P_G P_G = P_E \partial_i P_G P_E = 0$ , so that the trace in Eq. (C8) can be nonzero only when  $n \neq m$ ,  $m \neq s$ ,  $s \neq t$ , and  $t \neq n$ . In other words, only two terms are left out of the 16 terms summed over in Eq. (C8)

$$C_2 = -\frac{\pi^2}{3} \epsilon^{ijkl} \int \frac{d^4 k d\omega}{(2\pi)^5} \left\{ \frac{\text{Tr} \left[ P_G \frac{\partial P_G}{\partial k_i} P_E \frac{\partial P_G}{\partial k_j} P_G \frac{\partial P_G}{\partial k_k} P_E \frac{\partial P_G}{\partial k_\ell} \right]}{(\omega + i\delta - \epsilon_G)^3 (\omega + i\delta - \epsilon_E)^2} \right. \\ \left. + \frac{\text{Tr} \left[ P_E \frac{\partial P_G}{\partial k_i} P_G \frac{\partial P_G}{\partial k_j} P_E \frac{\partial P_G}{\partial k_k} P_G \frac{\partial P_G}{\partial k_\ell} \right]}{(\omega + i\delta - \epsilon_G)^2 (\omega + i\delta - \epsilon_E)^3} \right\} (\epsilon_E - \epsilon_G)^4.$$

Carrying out the integral over  $\omega$  and using identities [Eq. (C9)] again, we obtain

$$C_2 = \frac{1}{8\pi^2} \int d^4 k \epsilon^{ijkl} \text{Tr} \left[ P_E \frac{\partial P_G}{\partial k_i} \frac{\partial P_G}{\partial k_j} P_E \frac{\partial P_G}{\partial k_k} \frac{\partial P_G}{\partial k_\ell} \right]. \quad (\text{C10})$$

Now we will show that this is just the second Chern number. The Berry phase gauge field is defined by

$$a_i^{\alpha\beta}(\mathbf{k}) = -i \left\langle \alpha, \mathbf{k} \left| \frac{\partial}{\partial k_i} \right| \beta, \mathbf{k} \right\rangle,$$

in which  $\alpha, \beta = 1, 2, \dots, M$  are the occupied bands. The  $U(M)$  gauge curvature is given by

$$f_{ij}^{\alpha\beta} = \partial_i a_j^{\alpha\beta} - \partial_j a_i^{\alpha\beta} + i[a_i, a_j]^{\alpha\beta} \\ = -i \left( \frac{\partial \langle \alpha, \mathbf{k} | \partial | \beta, \mathbf{k} \rangle}{\partial k_i \partial k_j} - (i \leftrightarrow j) \right) + i \left( \frac{\partial \langle \alpha, \mathbf{k} |}{\partial k_i} \sum_{\gamma=1}^M |\gamma, \mathbf{k} \rangle \right. \\ \left. \times \langle \gamma, \mathbf{k} | \frac{\partial | \beta, \mathbf{k} \rangle}{\partial k_j} - (i \leftrightarrow j) \right) \\ = -i \left( \frac{\partial \langle \alpha, \mathbf{k} |}{\partial k_i} P_E(\mathbf{k}) \frac{\partial | \beta, \mathbf{k} \rangle}{\partial k_j} - (i \leftrightarrow j) \right),$$

in which  $(i \leftrightarrow j)$  means the term with  $i, j$  exchanged. In operator form, we have

$$\sum_{\alpha,\beta=1}^M |\alpha, \mathbf{k}\rangle f_{ij}^{\alpha\beta} \langle \beta, \mathbf{k}| = -i \frac{\partial P_G(\mathbf{k})}{\partial k_i} P_E \frac{\partial P_G(\mathbf{k})}{\partial k_j} - (i \leftrightarrow j).$$

Thus Eq. (C10) can be written in terms of the Berry phase curvature as<sup>43,47</sup>

$$C_2 = \frac{1}{32\pi^2} \int d^4k \epsilon^{ijkl} \text{Tr}[f_{ij} f_{kl}]. \quad (\text{C11})$$

In conclusion, the equivalence of nonlinear correlation function (53) and second Chern number (54) has been proven for the specific models of form (55). Due to the topological invariance of both Eq. (53) and the second Chern number, such a relation holds for any band insulator in (4+1) dimensions.

The procedure in this section can be easily generalized to any odd space-time dimensions. In (2n+1)-dimensional space-time, a *n*th Chern number  $C_n$  is defined in the BZ of a band insulator, which appears as a topological response coefficient to an external gauge field. Equivalently, the coefficient of the *n*th Chern-Simons term in the effective action of an external gauge field obtained from integrating out the fermions.<sup>41</sup>

#### APPENDIX D: THE WINDING NUMBER IN THE NONLINEAR RESPONSE OF DIRAC-TYPE MODELS

Starting from the symmetric form, in terms of general Green's functions, of Eq. (53) we want to calculate  $C_2$  for the generalized lattice Dirac model  $H(k) = d_a(k)\Gamma^a$ . Green's function for this model can be simplified to<sup>41</sup>

$$G(k, \omega) = \frac{\omega + d^a(k)\Gamma^a}{[\omega^2 - d^a(k)d_a(k)]}. \quad (\text{D1})$$

Inserting this into Eq. (53) yields the following calculation:

$$\begin{aligned} C_2 &= -\frac{\pi^2}{3} \epsilon_{ijkl} \int \frac{d^4k d\omega}{(2\pi)^5} \frac{1}{(\omega^2 - |d|^2)^5} \text{Tr}[(\omega + d_m \Gamma^m)(\omega + d_n \Gamma^n) \\ &\quad \times (\partial_i d^o \Gamma^o)(\omega + d_p \Gamma^p)(\partial_j d^q \Gamma^q) \\ &\quad \times (\omega + d_r \Gamma^r)(\partial_k d^s \Gamma^s)(\omega + d_t \Gamma^t)(\partial_l d^u \Gamma^u)] \\ &= -\frac{\pi^2}{3} \epsilon_{ijkl} \int \frac{d^4k d\omega}{(2\pi)^5} \frac{\partial_i d^o \partial_j d^q \partial_k d^s \partial_l d^u}{(\omega^2 - |d|^2)^5} \text{Tr}[(\omega + d_m \Gamma^m) \\ &\quad \times (\omega + d_n \Gamma^n) \Gamma^o (\omega + d_p \Gamma^p) \Gamma^q (\omega + d_r \Gamma^r) \Gamma^s (\omega + d_t \Gamma^t) \Gamma^u], \end{aligned}$$

where  $m, n, o, p, q, r, s, t, u = 0, 1, 2, 3, 4$ . Among all the terms in the bracket, the only ones with nonzero traces are those of 5, 7, and 9 $\Gamma$  matrices<sup>43</sup> and the complete trace simplifies nicely to  $-4\epsilon_{ioqsu} d^l (\omega^2 - |d|^2)^2$ , which simplifies  $C_2$  as

$$\begin{aligned} C_2 &= \frac{\pi^2}{3} \epsilon_{ijkl} \int \frac{d^4k d\omega}{(2\pi)^5} 4\epsilon_{ioqsu} \frac{d^l \partial_i d^o \partial_j d^q \partial_k d^s \partial_l d^u}{(\omega^2 - |d|^2)^3} \\ &= \frac{\pi^2}{4} \epsilon_{ijkl} \int \frac{d^4k}{(2\pi)^4} \epsilon_{ioqsu} \frac{d^l \partial_i d^o \partial_j d^q \partial_k d^s \partial_l d^u}{|d|^5} \\ &= \frac{3}{8\pi^2} \int d^4k \epsilon_{ioqsu} \hat{d}^l \partial_x \hat{d}^o \partial_y \hat{d}^q \partial_z \hat{d}^s \partial_w \hat{d}^u. \quad (\text{D2}) \end{aligned}$$

Thus we have proved that the winding number given in ex-

pression (64) is equal to the second Chern number defined in Eq. (53) for generic (4+1)-dimensional Dirac models.

#### APPENDIX E: STABILITY OF EDGE THEORIES IN GENERIC DIMENSIONS

In this appendix, we will determine the existence or absence of  $Z_2$  topological insulators in generic dimensions by studying the stability of boundary theories. First of all, the boundary theory of a topological insulator in (2n+1)-dimensions with nontrivial *n*th Chern number  $C_n$  is  $|C_n|$  copies of chiral fermions,

$$H = \text{sgn}(C_n) \sum_{s=1}^{|C_n|} \sum_{a=1}^{2n-1} \psi_s^\dagger v p_a \Gamma^a \psi_s. \quad (\text{E1})$$

The  $2n-1$   $\Gamma^a$  matrices are  $2^{n-1}$  dimensional and form an  $so(2n-1)$  Clifford algebra. We will study the lattice Dirac model since other systems with the same Chern number can be obtained by an adiabatic deformation from the lattice Dirac model and the topological stability of edge states does not depend on the adiabatic deformation, as will be discussed below. For simplicity, in the following we will focus on the case  $C_n=1$ , in which case the single particle Hamiltonian can be written as  $h(\mathbf{p}) = v p_a \Gamma^a$ . We have the following theorem about the symmetry of the Hamiltonian:

**Theorem I.** In a  $2^{n-1}$  dimensional representation of an  $so(2n-1)$  Clifford algebra generated by  $\Gamma^a$ ,  $a=1, 2, \dots, 2n-1$ , a unitary matrix  $M_{(n)}$  can be defined such that

$$M_{(n)}^\dagger \Gamma^a M_{(n)} = (-1)^{n-1} \Gamma^{a*}, \quad \forall a = 1, 2, \dots, 2n-1, \quad (\text{E2})$$

$$M_{(n)}^* M_{(n)} = (-1)^{[n/2]} \mathbb{I}. \quad (\text{E3})$$

Consequently the chiral Hamiltonian  $h(\mathbf{p}) = v p_a \Gamma^a$  satisfies

$$M_{(n)}^\dagger h(\mathbf{p}) M_{(n)} = (-1)^n h^*(-\mathbf{p}). \quad (\text{E4})$$

In Eq. (E3)  $[n/2]$  means the maximal integer that does not exceed  $n/2$ .

Theorem I can be proved by induction. First, the  $\Gamma^a$  matrices for  $n=2$  are the Pauli matrices  $\Gamma^a = \sigma_a$ ,  $a=1, 2, 3$ , and  $M_{(2)} = i\sigma_2$  satisfies the theorem. Suppose the theorem is true for case  $n$  with the matrix  $M_{(n)}$  and the  $2n-1$  gamma matrices  $\Gamma_{(2n-1)}^a$ , then the  $\Gamma_{(2n+1)}^a$  matrices for the  $so(2n+1)$  Clifford algebra can be generated by

$$\Gamma_{(2n+1)}^a = \begin{cases} \Gamma_{(2n-1)}^a \otimes \tau_y, & a = 1, 2, \dots, 2n-1 \\ \mathbb{I} \otimes \tau_x, & a = 2n \\ \mathbb{I} \otimes \tau_z, & a = 2n+1. \end{cases} \quad (\text{E5})$$

It is straightforward to check the anticommutation relations. Defining

$$M_{(n+1)} = \begin{cases} M_{(n)} \otimes i\tau_y, & n \text{ odd} \\ M_{(n)} \otimes \mathbb{I}, & n \text{ even}, \end{cases} \quad (\text{E6})$$

we find  $M_{(n+1)}$  satisfies the Theorem I.

From Eqs. (E3) and (E4) one finds different properties of  $M_{(n)}$  in different dimensions. (i) For  $n=4k-3, k \in \mathbb{Z}$ ,  $M_{(n)}$  reverses the sign of energy and satisfies  $M_{(n)}^* M_{(n)} = \mathbb{I}$ , which can be identified as particle-hole symmetry; (ii) for  $n=4k-2, k \in \mathbb{Z}$ ,  $M_{(n)}$  preserves the energy and satisfies  $M_{(n)}^* M_{(n)} = -\mathbb{I}$ , which can be identified as the time-reversal symmetry; (iii) for  $n=4k-1, k \in \mathbb{Z}$ ,  $M_{(n)}$  reverses the sign of energy but satisfies  $M_{(n)}^* M_{(n)} = -\mathbb{I}$ , which we call pseudo-particle-hole symmetry and denote by  $\tilde{C}$ ; and (iv) for  $n=4k, k \in \mathbb{Z}$ ,  $M_{(n)}$  preserves the energy but satisfies  $M_{(n)}^* M_{(n)} = \mathbb{I}$ , which behaves like the time-reversal symmetry of an integer-spin particle, and we call pseudo-time-reversal symmetry  $\tilde{T}$ .

In the following, we will denote the symmetries  $C, \tilde{C}, T$ , and  $\tilde{T}$ , defined by Eq. (E4), by  $M$  symmetry. Now we study what other terms can be added in the Hamiltonian without breaking the  $M$  symmetry. Given  $\Gamma^a$  for the  $so(2n-1)$  case, all the  $2^{n-1} \times 2^{n-1}$  Hermitian matrices can be expanded in the basis

$$\{\mathbb{I}, \Gamma^{a_1 a_2 \dots a_m} = i^{m(m-1)/2} \Gamma^{a_1} \Gamma^{a_2} \dots \Gamma^{a_m}, m = 1, 2, \dots, n-1\}.$$

As expected, the total number of matrices forming the basis is  $1 + \sum_{m=1}^{n-1} \binom{m}{2n-1} = 2^{2n-2}$ . The  $M_{(n)}$  transformation property of  $\Gamma^{a_1 a_2 \dots a_m}$  can be determined by that of  $\Gamma^a$  as

$$M_{(n)}^\dagger \Gamma_{(n)}^{a_1 a_2 \dots a_m} M_{(n)} = (-1)^{m(2n+m-3)/2} \Gamma_{(n)}^{a_1 a_2 \dots a_m*}. \quad (\text{E7})$$

If we have a constant term  $m_{a_1 a_2 \dots a_m} \Gamma^{a_1 a_2 \dots a_m}$  in the Hamiltonian without breaking the  $M$  symmetry, the following condition must be satisfied:

$$M^\dagger \Gamma^{a_1 a_2 \dots a_m} M = (-1)^n \Gamma^{a_1 a_2 \dots a_m*}. \quad (\text{E8})$$

Equation (E7) and (E8) thus require

$$(-1)^{m(2n+m-3)/2} = (-1)^n, \quad (\text{E9})$$

so that  $(m-1)(2n-2+m)/2$  must be odd. This condition can be satisfied by several possibilities: (1) when  $m$  is odd,  $(m-1)/2$  must be odd, which means  $m=4k-1, k \in \mathbb{N}$ . (2) When  $m$  is even,  $n-1+m/2$  must be odd, which means  $m=4k$  if  $n$  is even or  $m=4k-2$  if  $n$  is odd. In summary, the terms available in the Hamiltonian are given by

$$m = \begin{cases} 4k-1 \text{ or } 4k, & n \text{ even} \\ 4k-1 \text{ or } 4k-2, & n \text{ odd,} \end{cases} \quad (\text{E10})$$

in which  $k \in \mathbb{N}$  and  $m$  is also bounded by  $1 \leq m < n$ . For example, the first nontrivial case is  $n=2$  in which  $\Gamma^a, a=1, 2, 3$ , are Pauli matrices. Since  $n$  is even,  $m$  is required to be  $4k-1$  or  $4k$ , in which the lowest value is  $m=3$ . Consequently there is no mass term available. When  $n=3(2n-1=5)$  the  $\Gamma^a, a=1, 2, \dots, 5$  matrices are the usual Dirac matrices, and the only value of  $m$  satisfying Eq. (E10) is  $m=2$ . In other words, the terms  $\Gamma^{ab} = i\Gamma^a \Gamma^b$  do not break the corre-

sponding symmetry—the pseudo-particle-hole symmetry  $\tilde{C}$ .

Though there are all these constant terms available, the perturbed Hamiltonian  $h(k) = k_a \Gamma^a + \sum_m m_{a_1 a_2 \dots a_m} \Gamma^{a_1 a_2 \dots a_m}$  remains gapless because each  $\Gamma^{a_1 a_2 \dots a_m}$  commutes with some  $\Gamma^a$ . If  $m$  is even, it commutes with  $\Gamma^b, b \neq a_s \forall s=1, \dots, m$ . If  $m$  is odd, it commutes with  $\Gamma^{a_s}, \forall s=1, \dots, m$ . For the Hamiltonian with only one constant term  $m_{a_1 a_2 \dots a_m} \Gamma^{a_1 a_2 \dots a_m}$ , one can take  $p_i=0$  for all  $i$  except for  $i=a$ , where  $a$  is chosen such that  $\Gamma^a$  commutes with  $\Gamma^{a_1 a_2 \dots a_m}$ . Due to the commutativity, the Hamiltonian can be diagonalized along the  $a$ th axis, with the eigenvalues  $p_a \pm m$ . Consequently, we know the Hamiltonian is gapless. In other words, we have shown that no perturbation with the  $2^{n-1}$  band theory can open a gap for the  $2n$ -dimensional chiral fermion, which agrees with the topological stability of the  $(2n+1)$ -dimensional bulk system characterized by the  $n$ th Chern number.

Starting from the  $(2n+1)$ -dimensional topological insulator, dimensional reduction procedures can be carried out to obtain a  $[(2n-1)+1]$ -dimensional topological insulator. Correspondingly, the edge theory of the  $[(2n-2)+1]$ -dimensional topological insulator is given by the dimensional reduction in the  $[(2n-1)+1]$ -dimensional chiral fermion. If a nontrivial topological insulator can be defined, it has the boundary theory  $h(\mathbf{p}) = \sum_{a=1}^{2n-2} p_a \Gamma^a$ . Compared to the  $[(2n-1)+1]$ -dimensional chiral fermion, one  $\Gamma^a$  matrix is absent in the theory. In the same way, the boundary theory of lower dimensional descendants can be obtained by removing more momenta and  $\Gamma^a$  matrices from the chiral fermion Hamiltonian. Obviously, if too few  $\Gamma^a$  matrices are left, a mass term will be available, which anticommutes with all the rest of the  $\Gamma^a$  matrices and thus can make gap the whole edge spectrum. The upper critical dimension where the edge states become unstable is determined by the maximal number of  $\Gamma^a$  matrices that anticommute with some  $\Gamma^{a_1 a_2 \dots a_m}$ . For  $m$  even, the maximal number of  $\Gamma^a$ 's that anticommute with  $\Gamma^{a_1 a_2 \dots a_m}$  is  $m$ , while for  $m$  odd, the maximal number is  $2n-1-m$ . On the other hand, the available values of  $m$  are defined by the constraint [Eq. (E10)]. By studying the cases  $n=4k-3, 4k-2, 4k-1, 4k$ , and  $k \in \mathbb{N}$  separately, we obtain that the space-time dimension in which the chiral theory becomes unstable is given by  $d=2n-3$ . For example, in the case  $n=3$  discussed earlier, the mass terms  $\Gamma^{ab}$  are permitted by the pseudo-particle-hole symmetry. When the dimension is reduced from  $(5+1)$  dimensions to  $(2+1)$  dimensions, only two  $\Gamma^a$  matrices are used in the  $k$ -linear terms, so that some  $\Gamma^{ab}$  can be found which anticommutes with the gapless Hamiltonian and thus can make the surface theory gapped. Consequently, topologically stable boundary theories as descendants of  $[(2n-1)+1]$ -dimensional chiral fermions can only exist in  $[(2n-2)+1]$  and  $[(2n-3)+1]$  dimensions. Correspondingly, the  $Z_2$  topological insulators as descendants of the  $(2n+1)$ -dimensional topological insulator (with  $n$ th Chern number) can only be defined in  $[(2n-1)+1]$  and  $[(2n-2)+1]$  dimensions.

- <sup>1</sup>K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).
- <sup>2</sup>D. C. Tsui, H. L. Stormer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982).
- <sup>3</sup>R. B. Laughlin, Phys. Rev. B **23**, 5632 (1981).
- <sup>4</sup>R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
- <sup>5</sup>D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
- <sup>6</sup>Q. Niu, D. J. Thouless, and Y.-S. Wu, Phys. Rev. B **31**, 3372 (1985).
- <sup>7</sup>S. C. Zhang, T. H. Hansson, and S. Kivelson, Phys. Rev. Lett. **62**, 82 (1989).
- <sup>8</sup>S. C. Zhang, Int. J. Mod. Phys. B **6**, 25 (1992).
- <sup>9</sup>W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
- <sup>10</sup>R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
- <sup>11</sup>G. Ortiz and R. M. Martin, Phys. Rev. B **49**, 14202 (1994).
- <sup>12</sup>D. J. Thouless, Phys. Rev. B **27**, 6083 (1983).
- <sup>13</sup>J. Goldstone and F. Wilczek, Phys. Rev. Lett. **47**, 986 (1981).
- <sup>14</sup>S. C. Zhang and J. P. Hu, Science **294**, 823 (2001).
- <sup>15</sup>B. A. Bernevig, C. H. Chern, J. P. Hu, N. Toumbas, and S. C. Zhang, Ann. Phys. **300**, 185 (2002).
- <sup>16</sup>C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
- <sup>17</sup>B. A. Bernevig and S. C. Zhang, Phys. Rev. Lett. **96**, 106802 (2006).
- <sup>18</sup>C. Wu, B. A. Bernevig, and S. C. Zhang, Phys. Rev. Lett. **96**, 106401 (2006).
- <sup>19</sup>C. Xu and J. E. Moore, Phys. Rev. B **73**, 045322 (2006).
- <sup>20</sup>B. A. Bernevig, T. L. Hughes, and S. C. Zhang, Science **314**, 1757 (2006).
- <sup>21</sup>M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
- <sup>22</sup>L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).
- <sup>23</sup>J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306(R) (2007).
- <sup>24</sup>R. Roy, arXiv:cond-mat/0604211 (unpublished).
- <sup>25</sup>S. Murakami, N. Nagaosa, and S. C. Zhang, Phys. Rev. Lett. **93**, 156804 (2004).
- <sup>26</sup>C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
- <sup>27</sup>L. Fu and C. L. Kane, Phys. Rev. B **74**, 195312 (2006).
- <sup>28</sup>L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
- <sup>29</sup>R. Roy, arXiv:cond-mat/0607531 (unpublished).
- <sup>30</sup>R. Roy, arXiv:cond-mat/0608064 (unpublished).
- <sup>31</sup>X. L. Qi and S. C. Zhang, Phys. Rev. Lett. **101**, 086802 (2008).
- <sup>32</sup>Y. Ran, A. Vishwanath, and D.-H. Lee, Phys. Rev. Lett. **101**, 086801 (2008).
- <sup>33</sup>A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett. **51**, 2077 (1983).
- <sup>34</sup>J. von Neumann and E. Wigner, Phys. Z. **30**, 467 (1929).
- <sup>35</sup>X. L. Qi, Y. S. Wu, and S. C. Zhang, Phys. Rev. B **74**, 045125 (2006).
- <sup>36</sup>G. E. Volovik, *The Universe in a Helium Droplet* (Oxford University Press, New York City, 2003).
- <sup>37</sup>F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
- <sup>38</sup>Chaoxing Liu, Xiaoliang Qi, Xi Dai, Zhong Fang, and Shoucheng Zhang, Phys. Rev. Lett. **101**, 146802 (2008).
- <sup>39</sup>Y. Hatsugai, Phys. Rev. Lett. **71**, 3697 (1993).
- <sup>40</sup>R. Jackiw and C. Rebbi, Phys. Rev. D **13**, 3398 (1976).
- <sup>41</sup>M. F. L. Golterman, K. Jansen, and D. B. Kaplan, Phys. Lett. B **301**, 219 (1993).
- <sup>42</sup>G. E. Volovik, JETP Lett. **75**, 63 (2002).
- <sup>43</sup>S. Murakami, N. Nagaosa, and S. C. Zhang, Phys. Rev. B **69**, 235206 (2004).
- <sup>44</sup>M. Creutz, Rev. Mod. Phys. **73**, 119 (2001).
- <sup>45</sup>We use the capital  $\Gamma^a$  here since they are defined differently from the usual Dirac matrix  $\gamma^a$ ,  $a=0,1,2,3,5$ . The relation between  $\Gamma^a$  and  $\gamma^a$  is  $\Gamma^0=\gamma^0$ ,  $\Gamma^{1,2,3,4}=-i\gamma^0\gamma^{1,2,3,5}$ .
- <sup>46</sup>J. E. Avron, L. Sadun, J. Segert, and B. Simon, Phys. Rev. Lett. **61**, 1329 (1988).
- <sup>47</sup>E. Demler and S. C. Zhang, Ann. Phys. **271**, 83 (1999).
- <sup>48</sup>S. Adler, Phys. Rev. **177**, 2426 (1969).
- <sup>49</sup>J. S. Bell and R. Jackiw, Nuovo Cimento A **60**, 47 (1969).
- <sup>50</sup>D. Xiao, J. Shi, D. P. Clougherty, and Q. Niu, arXiv:cond-mat/0711.1855 (unpublished).
- <sup>51</sup>M. C. Huang and P. Sikivie, Phys. Rev. D **32**, 1560 (1985).
- <sup>52</sup>F. Wilczek, Phys. Rev. Lett. **58**, 1799 (1987).
- <sup>53</sup>K. Lee, Phys. Rev. D **35**, 3286 (1987).
- <sup>54</sup>R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).
- <sup>55</sup>P. Sikivie, Phys. Lett. B **137**, 353 (1984).
- <sup>56</sup>C. G. Callan, R. F. Dashen, and D. J. Gross, Phys. Lett. B **63**, 334 (1976).
- <sup>57</sup>R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976).
- <sup>58</sup>X. Dai, T. L. Hughes, X. L. Qi, Z. Fang, and S.-C. Zhang, Phys. Rev. B **77**, 125319 (2008).
- <sup>59</sup>E. Fradkin, E. Dagotto, and D. Boyanovsky, Phys. Rev. Lett. **57**, 2967 (1986).
- <sup>60</sup>N. Hur, S. Park, P. A. Sharma, J. S. Ahn, S. Guha, and S.-W. Cheong, Nature (London) **429**, 392 (2004).
- <sup>61</sup>W. Eerenstein, N. D. Mathur, and J. F. Scott, Nature (London) **442**, 759 (2006).
- <sup>62</sup>E. Witten, Phys. Lett. B **86**, 283 (1979).
- <sup>63</sup>More explicitly, if we choose the representation  $\Gamma^0=\sigma^3\otimes\mathbb{I}$ ,  $\Gamma^{1,2,3}=\sigma^1\otimes\tau^{1,2,3}$ , and  $\Gamma^4=\sigma^2\otimes\mathbb{I}$ , then the time-reversal transformation matrix  $T$  is defined as  $T=\mathbb{I}\otimes i\tau^2$ .
- <sup>64</sup>B. Halperin, Jpn. J. Appl. Phys., Suppl. **26**, 1913 (1987).
- <sup>65</sup>A. N. Redlich, Phys. Rev. D **29**, 2366 (1984).
- <sup>66</sup>T. Jaroszewicz, Phys. Lett. B **146**, 337 (1984).
- <sup>67</sup>A. G. Abanov and P. B. Wiegmann, J. High Energy Phys. **10**, 30 (2001).
- <sup>68</sup>C. Chamon, C.-Y. Hou, R. Jackiw, C. Mudry, S.-Y. Pi, and A. P. Schnyder, Phys. Rev. Lett. **100**, 110405 (2008).
- <sup>69</sup>T. Grover and T. Senthil, Phys. Rev. Lett. **100**, 156804 (2008).
- <sup>70</sup>X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Nat. Phys. **4**, 273 (2008).
- <sup>71</sup>B. Julia and A. Zee, Phys. Rev. D **11**, 2227 (1975).
- <sup>72</sup>M. Nakahara, *Geometry, Topology, and Physics* (Hilger, London, 1990).
- <sup>73</sup>D. B. Kaplan, Phys. Lett. B **288**, 342 (1992).
- <sup>74</sup>T. Kaluza, Sitzungsber. Preuss. Akad. Wiss. Berlin. (Math. Phys.) 966 (1921).
- <sup>75</sup>O. Klein, Z. Phys. **37**, 895 (1926).
- <sup>76</sup>H. B. Lawson and M.-L. Michelsohn, *Spin Geometry* (Princeton, Princeton, NJ, 1989).
- <sup>77</sup>N. Read and D. Green, Phys. Rev. B **61**, 10267 (2000).
- <sup>78</sup>D. A. Ivanov, Phys. Rev. Lett. **86**, 268 (2001).
- <sup>79</sup>M. Stone and R. Roy, Phys. Rev. B **69**, 184511 (2004).
- <sup>80</sup>C. G. Callan and J. A. Harvey, Nucl. Phys. B **250**, 427 (1985).