

**Quantization of topological invariants under symmetry-breaking disorder**Juntao Song<sup>1</sup> and Emil Prodan<sup>2,\*</sup><sup>1</sup>*Department of Physics, Hebei Normal University, Hebei 050024, China*<sup>2</sup>*Department of Physics, Yeshiva University, New York, New York 10016, USA*

(Received 13 July 2015; revised manuscript received 22 September 2015; published 10 November 2015)

In the strictly periodic setting, the electric polarization of inversion-symmetric solids with and without time-reversal symmetry and the isotropic magnetoelectric response function of time-reversal-symmetric insulators are known to be topological invariants displaying an exact  $\mathbb{Z}_2$  quantization. This quantization is stabilized by the symmetries. In the present work, we investigate the fate of such symmetry-stabilized topological invariants in the presence of a disorder which breaks the symmetries but restores them on average. Using a rigorous analysis, we conclude that the strict quantization still holds in these conditions. Numerical calculations confirm this prediction.

DOI: [10.1103/PhysRevB.92.195119](https://doi.org/10.1103/PhysRevB.92.195119)

PACS number(s): 73.43.Cd, 73.43.Nq, 74.62.En, 75.85.+t

An important issue in the field of symmetry-protected topological phases is the fate of the topological invariants in the presence of bulk and surface disorder which can break the symmetries [1]. For the time-reversal-symmetric topological insulators, this issue is related to the contamination of the samples with magnetic impurities and it has been addressed experimentally and theoretically in a number of works [2]. While this contamination can be fully controlled in laboratory conditions, this is not the case in the real world conditions. Still, if the contamination is small, the magnetic impurities are in a noncorrelated disordered phase and, on average, the time-reversal symmetry is preserved. Then a question which is extremely relevant for the practical applications of these materials is whether the topological characteristics, such as the extended character of the surface states and the quantization of the bulk topological invariant, are preserved under such “average” time-reversal-symmetry conditions. The first characteristic has been shown in Ref. [3] to hold if the disorder is not too strong and here we show that the isotropic magneto-electric response function [4] remains quantized in such disordered regimes, provided a spectral gap is present. The antiferromagnetic topological insulators introduced in Ref. [5] are another class where the topology is stabilized by an average time-reversal symmetry. In the presence of disorder, this class of insulators was recently shown [6] to possess distinct topological phases surrounded by a sharp phase boundary which can be detected by transport measurements. Furthermore, the weak topological insulators can be thought of as protected by the translational symmetry, and disorder definitely breaks this symmetry but the symmetry is restored on average. References [7,8] gave evidence that, in certain conditions, this is enough for the surface states of weak topological insulators to avoid Anderson localization.

When some of the stabilizing symmetries are space symmetries, such as for the crystalline topological insulators [9,10] or the inversion-symmetric [11,12], the reflection-symmetric [13,14], or the spin-orbit free topological insulators [15] in general, these issues are actually of central importance because the space disorder of the atoms is inevitable and impossible to control even in the laboratory conditions. If, for example, one considers the random displacements of the atoms due to finite

temperature, there can be no expectations that the disordered potential respects the underlying symmetry of the lattice, but one can be sure that, on average, the symmetries are strictly preserved since the thermodynamic state (including the nuclei) is symmetric. Such restoration of the symmetries by averaging also occurs when the disorder is induced by defects which do not destroy the crystalline order completely. Reference [1] introduced a  $\mathbb{Z}_2$  topological invariant, which in principle covers all classes of symmetry-protected topological insulators (and more) in such disordered conditions. This topological invariant was put to work for a disordered model with averaged reflection symmetry and a localization-delocalization transition was observed numerically exactly at the point where the invariant changed its value. In the present work, we consider disordered insulators with averaged inversion symmetry and show that the classical electric polarization assumes strict quantized values 0 or  $\frac{1}{2}$ , provided a spectral gap is present.

Now, the existing definitions of the symmetry-stabilized topological invariants depend fundamentally on the exactness of the symmetries, hence, at first sight, it seems impossible to define them for a concrete disorder configuration which, just by itself, breaks the symmetries. However, as noted in Ref. [1], when the symmetries, which are preserved only average, are combined with the translations, then a certain self-averaging property takes place, enabling one to define *exact* topological invariants for such realistic conditions. In the present work, we provide a precise formulation of this self-averaging property within the framework of homogeneous disordered systems [16,17]. While at this moment we cannot make any statements about the boundary states, our findings definitely contribute to the growing body of evidence that the symmetry-stabilized topological insulators are more robust than previously thought.

The paper is organized as follows. In the first section, we describe the homogeneous disordered systems and we formulate a precise relation between the disorder and symmetry, called here the compatibility condition, which ensures the restoration of the symmetry on average. An explicit yet very generic model for homogeneous disordered solids is presented, and the inversion as well as the time-reversal symmetries are shown to be compatible with the model. The second section discusses the electric polarization of homogeneous disordered systems with inversion symmetry and no time-reversal symmetry. The disorder is assumed to be compatible with the symmetry. The

\*prodan@yu.edu

polarization is shown to be an intensive macroscopic function with a self-averaging property (preventing fluctuations from one disorder configuration to another), and which takes only quantized values. The second section also reports a numerical analysis which confirms the theoretical predictions. The third section applies the arguments to the isotropic magnetoelectric response function of time-reversal-symmetric (TRS) insulators, leading to similar conclusions. The last section summarizes our conclusions.

## I. HOMOGENEOUS SOLIDS

Physically, a homogeneous system is defined as an extended system with translation invariance broken at the microscopic scale, but this symmetry breaking is undetectable at macroscopic scales. In other words, taking micron-size samples from a large piece of a homogeneous material will lead to same intensive thermodynamic parameters and coefficients (of course not exactly the same for finite pieces, but the differences are well below the experimental resolution). Formulating this property in a mathematically rigorous way was quite a challenge for the mathematical physics community, but these days the concept of a homogeneous solid state system has a very precise meaning and the mathematical framework built around this concept is natural and fruitful for a large class of aperiodic systems [16,17]. According to the precise definition, an aperiodic tight-binding Hamiltonian  $H$  over the lattice  $\mathbb{Z}^d$  is homogeneous if  $H$  together with its lattice translates  $T_a H T_a^{-1}$ , for all  $a \in \mathbb{Z}^d$ , define a set which has a compact closure in the strong topology of bounded operators. Here and throughout,  $T_a$  will denote the lattice translation by  $a$ . What we are going to present in the following applies strictly to the homogeneous system so defined, but for concreteness we will carry the discussion in the context of disordered lattice models, which are explicitly constructed in Example 1 below.

Henceforth, let us consider a generic finite-range disordered lattice model

$$H_\omega = H_0 + V_\omega \quad (1)$$

defined over the Hilbert space spanned by  $|n, \alpha\rangle$ , where  $n \in \mathbb{Z}^D$  represent the nodes of the lattice indexing the unit cells and  $\alpha = 1, \dots, Q$  the orbitals associated with a unit cell. The orbital index includes the spin degree of freedom, but the latter will be separated out when needed. The orbitals can and will be chosen to be real, that is, invariant to complex conjugation:  $\mathcal{K}|n, \alpha\rangle = |n, \alpha\rangle$ . The translationally invariant piece  $H_0$  of the Hamiltonian is assumed to depend on a set of  $N$  parameters  $\xi = (\xi_1, \dots, \xi_N)$ . When needed, we write this dependence explicitly as  $H_0(\xi)$  or  $H_\omega(\xi)$ . We include such dependence because both the electric polarization and the magnetoelectric response functions are defined through deformations of the system from a reference configuration. If the reader is more comfortable with an explicit representation of this parameter space, he may think of the vector  $\xi$  as the (always finite) collection of hopping amplitudes.

The random potential  $V_\omega$  depends on the disorder configuration  $\omega$ , which is seen here as a point in a disorder configuration space  $\Omega$ , which is compact and equipped with a probability measure  $d\omega$ . The system is assumed to be homogeneous, which

in this context is assured by the covariant property

$$T_a H_\omega T_a^{-1} = H_{t_a \omega}, \quad \forall a \in \mathbb{Z}^d, \quad (2)$$

where  $t$  represents a homeomorphic action of the group of lattice translations on  $\Omega$ . The action  $t$  is assumed ergodic and probability preserving so that Birkhoff's ergodic theorem [18] applies; namely, the following identity holds with probability 1 in  $\omega$ :

$$\lim_{V \rightarrow \infty} \frac{1}{V} \sum_{a \in V} f(t_a \omega) = \int_{\Omega} d\omega' f(\omega'). \quad (3)$$

It is precisely this identity which ensures the nonfluctuating character of the intensive thermodynamic functions from one disorder configuration to another.

We now formulate a precise condition which automatically leads to the restoration of the symmetry after the disorder average is taken. Consider a generic symmetry operation:

$$S|n, \alpha\rangle = \sum_{n', \alpha'} R_{n\alpha; n'\alpha'} |n', \alpha'\rangle, \quad (4)$$

which can be linear or antilinear. Recall that the orbitals are real, hence the symmetry is fully determined by the coefficients  $R_{n\alpha; n'\alpha'}$ . The symmetry generates an action on the parameter space,

$$S H_0(\xi) S^{-1} = H_0(S\xi), \quad (5)$$

and on the disorder configuration space,

$$S V_\omega S^{-1} = V_{S\omega}, \quad (6)$$

for which we use the same notation  $S$ . As a consequence, we have the following covariant property:

$$S H_\omega(\xi) S^{-1} = H_{S\omega}(S\xi). \quad (7)$$

In fact, such covariant property is obeyed by any function of the Hamiltonian:

$$S \Phi(H_\omega(\xi)) S^{-1} = \Phi(H_{S\omega}(S\xi)). \quad (8)$$

*Definition.* We say that the symmetry is compatible with the disorder if

$$d(S\omega) = d\omega, \quad (9)$$

that is, if the probability measure  $d\omega$  is invariant under the action on  $\Omega$  induced by the symmetry.

Perhaps the reader already noticed that we are merely restating the condition formulated in Ref. [1], Sec. II A, in a more concise mathematical notation. This definition is relevant when we consider the disorder average,

$$S \left( \int_{\Omega} d\omega \Phi(H_\omega(\xi)) \right) S^{-1} = \int_{\Omega} d\omega \Phi(H_{S\omega}(S\xi)), \quad (10)$$

because a change of variable  $S\omega \rightarrow \omega$  leads to

$$\dots = \int_{\Omega} d(S^{-1}\omega) \Phi(H_\omega(S\xi)) = \int_{\Omega} d\omega \Phi(H_\omega(S\xi)). \quad (11)$$

In other words,

$$S \left( \int_{\Omega} d\omega \Phi(H_\omega(\xi)) \right) S^{-1} = \int_{\Omega} d\omega \Phi(H_\omega(S\xi)). \quad (12)$$

In particular, if  $\xi$  is a fixed point for the symmetry,  $\mathcal{S}\xi = \xi$ , or equivalently if  $\mathcal{S}H_0(\xi)\mathcal{S}^{-1} = H_0(\xi)$ , then

$$\mathcal{S}\left(\int_{\Omega} d\omega \Phi(H_{\omega}(\xi))\right)\mathcal{S}^{-1} = \int_{\Omega} d\omega \Phi(H_{\omega}(\xi)), \quad (13)$$

which shows that Eq. (9) implies the restoration of the symmetry for the averaged quantities.

*Example 1.* Generic homogeneous disordered model.

$$H_{\omega} = \sum_{n,\alpha;m,\beta} [t_{n-m}^{\alpha\beta}(\xi) + W\omega_{n,m}^{\alpha\beta}]|n,\alpha\rangle\langle m,\beta|, \quad (14)$$

where  $\omega_{n,m}^{\alpha\beta}$  are independent random entries drawn uniformly from the interval  $[-\frac{1}{2}, \frac{1}{2}]$ . The collection of all random variables  $\omega = \{\omega_{n,m}^{\alpha\beta}\}$  can be viewed as a point in an infinite-dimensional configuration space  $\Omega$ , which is just an infinite product of intervals  $[-\frac{1}{2}, \frac{1}{2}]$ . The result is a compact and metrizable Tychonov space which can be equipped with the product probability measure:

$$d\omega = \prod_{n,\alpha;m,\beta} d\omega_{n,m}^{\alpha\beta}. \quad (15)$$

There is a natural action of the lattice translations on  $\Omega$ :

$$(t_a\omega)_{n,m}^{\alpha\beta} = \omega_{n-a,m-a}^{\alpha\beta}, \quad a \in \mathbb{Z}^D, \quad (16)$$

which leaves  $d\omega$  invariant and is known to act ergodically, hence Eq. (3) applies. It is straightforward to check that the Hamiltonian has indeed the covariant property:

$$T_a H_{\omega} T_a^{-1} = H_{t_a\omega}. \quad (17)$$

*Example 2.* Compatibility of the inversion symmetry with the model (14). This symmetry maps the unit cell  $\mathbf{n}$  into  $-\mathbf{n}$  and it can mix the orbitals in the process. However, we can always choose an orbital basis so that no such mixing occurs, and since  $\mathcal{I}^2 = 1$ :

$$\mathcal{I}|n,\alpha\rangle = \sum_{\alpha} \sigma_{\alpha} |-\mathbf{n},\alpha\rangle, \quad (18)$$

where all  $\sigma$ 's are signs. The induced action on  $\Omega$  is

$$(\mathcal{I}\omega)_{nm}^{\alpha\beta} = \sigma_{\alpha}\sigma_{\beta}\omega_{-n,-m}^{\alpha\beta}. \quad (19)$$

Now,

$$d\omega = \prod_{n,\alpha;m,\beta} d\omega_{n,m}^{\alpha\beta} = \prod_{n,m;\alpha} d\omega_{n,m}^{\alpha\alpha} \prod_{n,m;\alpha<\beta} d\omega_{n,m}^{\alpha\beta} d\omega_{n,m}^{\beta\alpha}, \quad (20)$$

and now one can see explicitly that  $d(\mathcal{I}\omega) = d\omega$ .

*Example 3.* Compatibility of the time-reversal symmetry with the model (14). The TRS is defined as the antilinear map,

$$\Theta|n,\alpha,\sigma\rangle = -\sigma|n,\alpha,-\sigma\rangle, \quad (21)$$

where the spin degree of freedom  $\sigma = \pm 1$  (for spin up/down) was separated out. Note that  $\Theta^2 = -1$ . The induced action on  $\Omega$  is

$$(\Theta\omega)_{nm}^{\alpha,\sigma;\beta,\sigma'} = \sigma\sigma'\omega_{n,m}^{\alpha,-\sigma;\beta,-\sigma'}. \quad (22)$$

Then,

$$\begin{aligned} d\omega &= \prod_{n,\alpha,\sigma;m,\beta,\sigma'} d\omega_{n,m}^{\alpha,\sigma;\beta,\sigma'} \\ &= \prod_{n,m} d\omega_{n,m}^{\alpha,1;\beta,1} d\omega_{n,m}^{\alpha,-1;\beta,-1} d\omega_{n,m}^{\alpha,-1;\beta,1} d\omega_{n,m}^{\alpha,1;\beta,-1}, \end{aligned} \quad (23)$$

and now one can see explicitly that  $d(\Theta\omega) = d\omega$ .

## II. ELECTRIC POLARIZATION OF INVERSION-SYMMETRIC INSULATORS

### A. Generic definition

By definition [19,20], the change in the electric polarization, as a result of a macroscopic deformation  $H_{\omega}(\xi_t)$  of the Hamiltonian, is

$$\Delta\mathcal{P}_{\omega} = \int_0^T \mathbf{j}_{\omega}(t) dt, \quad (24)$$

where  $\mathbf{j}_{\omega}(t)$  is the density of the charge current. The latter is microscopically defined as

$$\mathbf{j}_{\omega}(t) = \lim_{V \rightarrow \infty} \frac{1}{V} \sum_{n \in V} \sum_{\alpha=1}^Q \langle n,\alpha | \rho_{\omega}(t) \mathbf{J}_{\omega}(t) | n,\alpha \rangle, \quad (25)$$

where  $\rho_{\omega}(t)$  is the time-evolved density matrix and

$$\mathbf{J}_{\omega}(t) = ie[H_{\omega}(\xi_t), \mathbf{X}] \quad (26)$$

is the current operator. Here,  $\mathbf{X}$  denotes the position operator and, for convenience, the electron charge  $e$  will be set to unity in the following. If the deformation starts from the thermodynamic equilibrium state, then the time-evolved density matrix is

$$\rho_{\omega}(t) = U_t \Phi_{\text{FD}}(H_{\omega}(\xi_0)) U_t^{-1}, \quad (27)$$

with  $U_t$  being the unitary time evolution generated by  $H_{\omega}(\xi_t)$  and  $\Phi_{\text{FD}}$  the Fermi-Dirac distribution. One should not confuse  $\rho_{\omega}(t)$  and  $\Phi_{\text{FD}}(H_{\omega}(\xi_t))$ , because the time-evolved density matrix is no longer given by the Gibbs state.

### B. Self-averaging property

Now note that the fundamental formula in Eq. (24) involves the density of the current rather than the current itself, and this is why the trace per volume appears in Eq. (25). This is consistent with the fact that polarization is an intensive macroscopic function. From a technical point of view, this is an important observation because we can use Birkhoff's ergodic theorem [18] to demonstrate that  $\Delta\mathcal{P}_{\omega}$  is, with probability 1, independent of the disorder configuration  $\omega$ . Indeed, note that the operator inside the brackets of Eq. (25), called  $F_{\omega}$  in the lines below, is covariant:

$$T_a F_{\omega} T_a^{-1} = F_{t_a\omega}. \quad (28)$$

Then,

$$\frac{1}{V} \sum_{n \in V} \langle n | F_{\omega} | n \rangle = \frac{1}{V} \sum_{n \in V} \langle 0 | F_{t_n^{-1}\omega} | 0 \rangle = \int_{\Omega} d\omega \langle 0 | F_{\omega} | 0 \rangle \quad (29)$$

in the limit  $V \rightarrow \infty$ . The conclusion is that  $\Delta\mathcal{P}_{\omega}$  is self-averaging and its macroscopic value comes as an average over

the disorder configuration space. As such, we can drop the subscript  $\omega$  in  $\Delta\mathcal{P}_\omega$  from now on.

### C. Schulz-Baldes-Teufel formula

We consider now infinitely slow deformations of the Hamiltonian, which are better visualized as paths  $\gamma$  in the parameter space, parametrized as  $\{\xi_s\}_{s \in [0,1]}$ . By employing the adiabatic theorem, Schulz-Baldes and Teufel showed in Ref. [21] that in the extreme adiabatic limit and when the temperature goes to zero,

$$\Delta\mathcal{P}(\gamma) = \int_0^1 ds \mathcal{T}(P_\omega(\xi_s) i[\partial_s P_\omega(\xi_s), i[X, P_\omega(\xi_s)]]), \quad (30)$$

where  $\mathcal{T}$  denotes the trace per volume:

$$\mathcal{T}(\dots) = \lim_{V \rightarrow \infty} \frac{1}{V} \sum_{\mathbf{n} \in V} \sum_{\alpha=1}^Q \langle \mathbf{n}, \alpha | \dots | \mathbf{n}, \alpha \rangle, \quad (31)$$

and

$$P_\omega(\xi) = \chi_{[-\infty, \epsilon_F]}(H_\omega(\xi)) \quad (32)$$

is the Fermi projection onto the occupied states at coordinate  $\xi$ . Above, it is assumed that the spectral gap of  $H_\omega(\xi_s)$  remains open for all  $s \in [0,1]$  and that the Fermi level  $\epsilon_F$  is inside this gap. We will keep the  $i = \sqrt{-1}$  in front of the commutators in order to make them self-adjoint. Now, using Birkhoff's theorem as before, we can write, equivalently,

$$\Delta\mathcal{P}(\gamma) = \int_0^1 ds \int_\Omega d\omega \sum_{\alpha=1}^Q \langle 0, \alpha | P_\omega(\xi_s) i[\partial_s P_\omega(\xi_s), i[X, P_\omega(\xi_s)]] | 0, \alpha \rangle. \quad (33)$$

This is the Schulz-Baldes-Teufel formula, which can be regarded as the disordered version of the King-Smith-Vanderbilt formula for the static spontaneous orbital polarization [22].

### D. Quantization

The proof of quantization proceeds in two steps. First, based on the Schulz-Baldes-Teufel formula Eq. (33), we can demonstrate that the change in the electric polarization along the inverted path  $\mathcal{I}\gamma$  (see Fig. 1) is

$$\Delta\mathcal{P}(\mathcal{I}\gamma) = -\Delta\mathcal{P}(\gamma). \quad (34)$$

This equality is remarkable because we are computing the polarization of a system in an arbitrary disorder configuration

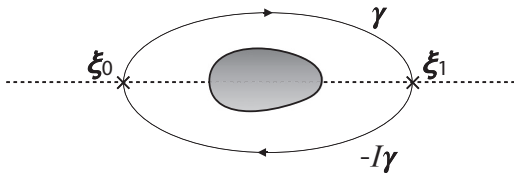


FIG. 1. The direct path  $\gamma$  (upper semiplane) and the inverted path  $-\mathcal{I}\gamma$  (lower semiplane) in the parameter space  $\xi$ . The horizontal axis represents the manifold which is left invariant by the symmetry. The interesting cases occur when the spectral gap closes along this manifold, which is schematically shown by the gray region.

which breaks the inversion symmetry. For the periodic case with strict inversion symmetry, this property is well known [11,12]. The keys to its proof are the self-averaging property, the compatibility between the inversion symmetry and disorder, and the behavior of the position operator under inversion,  $\mathcal{I}X\mathcal{I}^{-1} = -X$ . The proof proceeds as follows:

$$\Delta\mathcal{P}(\mathcal{I}\gamma) = \int_0^1 ds \int_\Omega d\omega \sum_{\alpha=1}^Q \langle 0, \alpha | P_\omega(\mathcal{I}\xi_s) \times i[\partial_s P_\omega(\mathcal{I}\xi_s), i[X, P_\omega(\mathcal{I}\xi_s)]] | 0, \alpha \rangle, \quad (35)$$

and observe that

$$P_\omega(\mathcal{I}\xi_s) = \mathcal{I}P_{\mathcal{I}^{-1}\omega}(\xi_s)\mathcal{I}^{-1}. \quad (36)$$

Then a change of variable  $\omega \rightarrow \mathcal{I}\omega$  and cancelations of terms such as  $\mathcal{I}\mathcal{I}^{-1}$  lead us to

$$\Delta\mathcal{P}(\mathcal{I}\gamma) = \int_0^1 ds \int_\Omega d(\mathcal{I}\omega) \sum_{\alpha=1}^Q \langle 0, \alpha | \mathcal{I}P_\omega(\xi_s) \times i[\partial_s P_\omega(\xi_s), i[-X, P_\omega(\xi_s)]] \mathcal{I}^{-1} | 0, \alpha \rangle. \quad (37)$$

The probability measure is invariant,  $d(\mathcal{I}\omega) = d\omega$ . Also, the 0 site is left invariant by  $\mathcal{I}$  and, since we are tracing over the orbital degrees of freedom, the action of the remaining  $\mathcal{I}$  operators have no effect and can be removed. Then Eq. (34) follows.

The second step of the proof consists of the following argument. Assume that the initial and final Hamiltonians are inversion symmetric; that is,  $\xi_0$  and  $\xi_1$  are fixed points of  $\mathcal{I}$ ,  $\mathcal{I}\xi_0 = \xi_0$ , and  $\mathcal{I}\xi_1 = \xi_1$ . Then any path  $\gamma$  joining the two Hamiltonians can be closed into a loop by augmenting  $-\mathcal{I}\gamma$ , where the minus sign implies that the path is walked in reverse (see Fig. 1). Note that the argument does not work if the end points of  $\gamma$  are not fixed points for  $\mathcal{I}$ . Given Eq. (34), this has the following effect:

$$\Delta\mathcal{P}(\gamma) = \frac{1}{2}[\Delta\mathcal{P}(\gamma) + \Delta\mathcal{P}(-\mathcal{I}\gamma)] = \frac{1}{2}\Delta\mathcal{P}(\gamma - \mathcal{I}\gamma). \quad (38)$$

But as already noted in Refs. [21,23], the change in the polarization along a closed loop leads to the noncommutative first Chern number defined in Ref. [24]. The conclusion is

$$\Delta\mathcal{P}_j(\gamma) = \frac{1}{2} C_1[(\gamma - \mathcal{I}\gamma) \times \mathcal{S}_j], \quad j = 1, \dots, D, \quad (39)$$

where on the right-hand side is the Chern number of the Fermi projection over the manifold  $(\gamma - \mathcal{I}\gamma)$  times the section of the noncommutative Brillouin torus along the  $j$ th direction:

$$C_1 = \int_{\gamma - \mathcal{I}\gamma} d\xi \mathcal{T}(P_\omega(\xi) i[\partial_\xi P_\omega(\xi), i[X_j, P_\omega(\xi)]]). \quad (40)$$

The usual constant in front is absent here because one derivative is taken in the  $k$ -space and the other in the real-space representation. As long as  $P_\omega(\xi)$  is smooth along the loop and its matrix elements  $\langle \mathbf{n} | P_\omega(\xi) | \mathbf{m} \rangle$  decay sufficiently fast, all these Chern numbers are integers for dimension  $D = 1$  or  $2$ . In  $D = 3$  these Chern numbers are weak topological invariants and they remain integer only if magnetic fields are not present. All these conditions are met if the spectral gap at the Fermi level remains open along the loop.



In conclusion, when insulators are deformed between fixed points of the inversion-symmetry operator, the change in the electric polarization is quantized in units of  $\frac{1}{2}$ , even though the symmetry exists only for disorder averages.

### E. Numerical confirmation

We consider the 1-dimensional Rice-Mele model [25]:

$$H_\omega(\xi) = \frac{1}{2} \sum_{n \in \mathbb{Z}} [1 + (-1)^n \xi_1] (|n\rangle \langle n+1| + |n+1\rangle \langle n|) + \frac{1}{2} \sum_{n \in \mathbb{Z}} [(-1)^n \xi_2 + W \omega_n] |n\rangle \langle n|, \quad (41)$$

with on-site disorder. As detailed in Ref. [26], this model was introduced originally by Rice and Mele to study solitons

in conducting polymers but it was subsequently used in other contexts too. It is a particular form of the generic model Eq. (14); hence everything we stated about Eq. (14) continues to hold here. In the absence of disorder, the Rice-Mele model is gapped if half filling is assumed, except at  $\xi = 0$ , and it is inversion symmetric whenever  $\xi_2 = 0$ . Also, it is well known [26] that adiabatic deformations around closed loops surrounding the origin lead to a quantized polarization change  $\Delta \mathcal{P} = \pm 1$  (or  $C_1 = \pm 1$ ).

The model has two states per unit cell and it can be brought to the form written in Eq. (14) by defining

$$|2n\rangle = |n, 1\rangle, \quad |2n+1\rangle = |n, -1\rangle, \quad (42)$$

in which case the Hamiltonian takes the form

$$H_\omega(\xi) = \frac{1}{2} \sum_{n \in \mathbb{Z}} \sum_{\alpha = \pm 1} (1 + \alpha \xi_1) \left( |n, \alpha\rangle \langle n + \frac{1}{2}(1 - \alpha), -\alpha| + |n + \frac{1}{2}(1 - \alpha), -\alpha\rangle \langle n, \alpha| \right) + \frac{1}{2} \sum_{n \in \mathbb{Z}} \sum_{\alpha = \pm 1} (\alpha \xi_2 + W \omega_{n, \alpha}) |n, \alpha\rangle \langle n, \alpha|. \quad (43)$$

If in the original rendering of Eq. (41) we choose the inversion point to be between the 0th and first site, then the inversion operation takes the form

$$\mathcal{I}|n, \alpha\rangle \rightarrow |-n, -\alpha\rangle. \quad (44)$$

It can be set in the diagonal form of Eq. (18) if we choose to work with the states  $\frac{1}{\sqrt{2}}(|n, 1\rangle \pm |n, -1\rangle)$ . However, we will not do that here. A direct computation will show that

$$\mathcal{I} H_\omega \mathcal{I}^{-1} = H_{\mathcal{I}\omega}(\mathcal{I}\xi), \quad (45)$$

with

$$\mathcal{I}(\xi_1, \xi_2) = (\xi_1, -\xi_2), \quad (\mathcal{I}\omega)_{n, \alpha} = \omega_{-n, -\alpha}. \quad (46)$$

Let us stress again that the on-site disorder breaks the inversion symmetry because we impose no correlation between  $\omega_{n, \alpha}$  and  $\omega_{-n, -\alpha}$ . However, the probability measure  $\prod_{n, \alpha} d\omega_{n, \alpha}$  is easily seen to be invariant under  $\mathcal{I}$ .

We now choose the path  $\gamma$  to be the semicircle:

$$\gamma = 0.1 \times \{\cos s, \sin s\}_{s \in [0, \pi]}, \quad (47)$$

which connects two fixed points of the inversion symmetry. By augmenting with  $-\mathcal{I}\gamma$  we obtain the full circle,  $s \in [0, 2\pi]$ . The numerical calculations are performed at finite size  $n \in \{0, \dots, L\}$  and with periodic boundary conditions. Formulas from Eqs. (30) and (40) are used for all the results reported here. The commutator with the position operator is implemented using the strategy developed in Ref. [27]. The path integral is discretized using 100 points for  $\gamma$  and 200 points for  $\gamma - \mathcal{I}\gamma$ . The derivative  $\partial_\xi$  with respect to  $\xi$  along the path is computed using the five-point stencil finite-difference approximation. The disorder strength  $W$  was increased gradually and the calculation was repeated 100 times with updated random potential for each  $W$ .

The results are reported in Fig. 2, where each dot seen in these plots represents a single disorder configuration. In other words, *no disorder average* has been performed on the data. In the first row we show the results for  $\Delta \mathcal{P}(\gamma)$  and Chern number or  $\Delta \mathcal{P}(\gamma - \mathcal{I}\gamma)$ . The first thing to notice is the clustering of the data for  $\Delta \mathcal{P}(\gamma)$  around the predicted quantized value of  $-\frac{1}{2}$ , at disorder strengths lower than a critical value  $W_c$ . The fluctuations around the quantized values can be attributed to the finite system sizes because the fluctuations can be seen to diminish as the size of the system gets larger. This confirms our theoretical prediction that  $\Delta \mathcal{P}(\gamma)$  is self-averaging hence nonfluctuating in the thermodynamic limit, and that, at least for small  $W$ 's, it takes quantized values in the unit of  $\frac{1}{2}$ .

The Chern number shows a very good quantization, with virtually no fluctuations from one disorder configuration to another below the same  $W_c$ . The reason for this difference is that the quantization of the Chern number does not require the restoring of the symmetry. It only requires that the exponential decay rate of the Fermi projection be large compared to  $1/L$ . The Fermi projection also needs to be smooth for  $\xi$  along the loop, but this is automatically the case if the spectral gap remains open (which is the case for small  $W$ 's). If, for example, we average  $\Delta \mathcal{P}(\gamma)$  over just two disorder configurations which are mirrored by inversion symmetry, then the quantization of  $\Delta \mathcal{P}(\gamma)$  will be as good as that of the Chern number (this can be shown exactly). While this could be a better numerical method to compute  $\Delta \mathcal{P}(\gamma)$ , it is irrelevant for the present discussion because here we want to demonstrate a principle, namely, that the quantization of  $\Delta \mathcal{P}(\gamma)$  occurs even for a single disorder configuration  $\omega$ , provided the system size is large enough. So, what is actually happening in this latter case? For this, let us note that the formula for  $\Delta \mathcal{P}(\gamma)$  remains invariant if the disorder is translated. Now, due to the ergodicity of the translations, when we translate the disorder we start exploring the disorder configuration space  $\Omega$ , and if the size

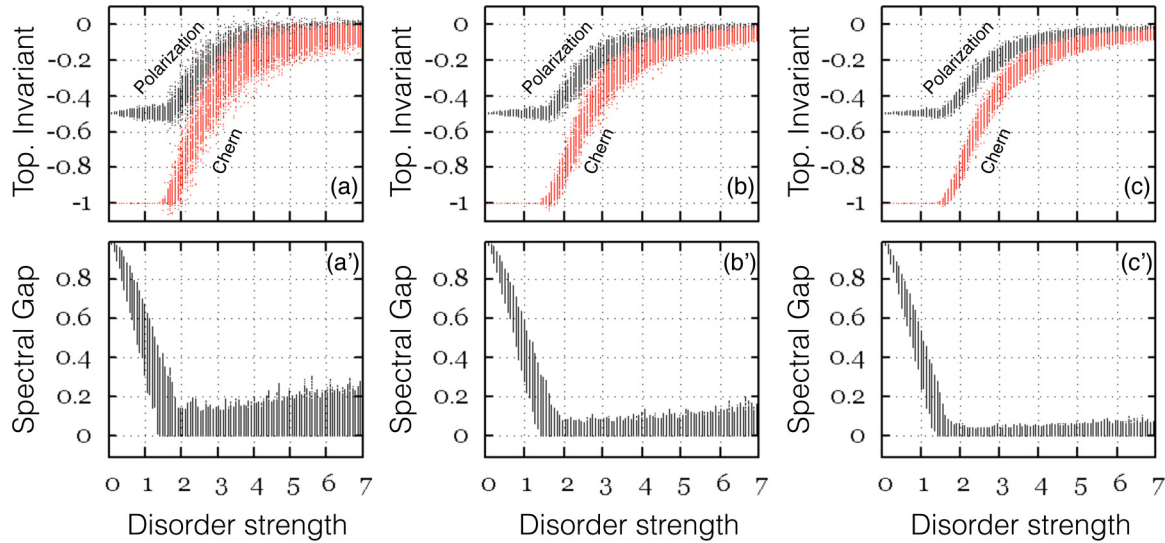


FIG. 2. (Color online) Top row: The variation of the electric polarization when the Rice-Mele model is adiabatically deformed along a semicircle (black dots) and along a full circle (red dots) in the two-parameter space. In the latter case, the variation is equal to a Chern number, hence the label. The calculations are performed in the presence of a disordered potential which breaks the inversion symmetry but restores it on the average. The disorder strength was increased from 0 to 7. Each dot corresponds to one disorder configuration and 100 disorder configurations were considered for each disordered strength. The average and statistical variance are not shown on purpose (see text). Bottom rows: The spectral gap of the Hamiltonian at half filling. A dot represents the spectral gap at one disorder configuration and one position along the adiabatic deformation path. The three columns correspond to the different system sizes:  $L = 100$  (a), 200 (b), and 400 (c).

of the system is large enough, at some point a translation will bring us close to a new disorder configuration which looks in many respects like the inverted disorder  $\mathcal{I}\omega$ . In other words, if the system size is large enough, we do not need to average over two mirrored disorder configurations because this already happens due to the ergodicity of the system.

The last issue we need to address is the value of the critical  $W_c$ . We know already that  $W_c$  should be larger or equal than the value of  $W$  where the spectral gap of  $H_\omega(\xi)$  closes somewhere along the loop. Can  $W_c$  be strictly larger than this value that we mentioned? To answer this question, we plotted in the second row of Fig. 2 the spectral gaps of  $H_\omega(\xi)$  at half filling, as  $\xi$  was varied along  $\gamma$  and the disorder configurations were updated. These many data were then collapsed and shown for increasing values of  $W$ . Whenever a dot in the second row of Fig. 2 touches the horizontal axis, the spectral gap of  $H_\omega(\xi)$  closed for some  $\xi$  along the loop and some disorder configuration. Note that the fluctuations of the gap due to disorder die out in the thermodynamic limit. By comparing the first and second rows of Fig. 2, we can conclude with confidence that the deviations from the quantized values in the top row of Fig. 2 occur exactly at the value of  $W$  where the spectral gap closes. The answer to our question is no, and the reason is because the Fermi projection fails to be smooth for  $\xi$  beyond that point.

### III. MAGNETOELECTRIC RESPONSE OF TRS INSULATORS

Consider now an insulator  $H_\omega(\xi)$  in dimension  $D = 3$ . In this section we investigate the isotropic part of the

magnetolectric response function:

$$\alpha = \frac{1}{3} \sum_{j=1}^3 \frac{\partial \mathcal{P}_j}{\partial B_j}. \quad (48)$$

If TRS is considered, then the partial derivatives with respect to the magnetic field are taken at  $\mathbf{B} = 0$ . Since the arguments are repetitive, we will expedite the exposition.

The main tool of our analysis is the formula derived in Ref. [28] for the change  $\Delta\alpha$  during a deformation of the system along a path  $\gamma = \{\xi_s\}_{s \in [0,1]}$  in the parameter space:

$$\Delta\alpha(\gamma) = \frac{1}{2} \epsilon^{i_1 \dots i_4} \int_0^1 ds \mathcal{T}(P_\omega(\xi_s) \partial_{i_1} P_\omega(\xi_s) \dots \partial_{i_4} P_\omega(\xi_s)),$$

where  $\mathcal{T}$  denotes the trace per volume,  $\partial_j$  is a shorthand for

$$\partial_j P_\omega(\xi_s) = i[X_j, P_\omega(\xi_s)], \quad \text{for } j = 1, 2, 3, \quad (49)$$

and  $\partial_4 P_\omega(\xi_s) = \partial_s P_\omega(\xi_s)$ . Also,  $\epsilon^{i_1 \dots i_4}$  is the antisymmetric tensor, and summation over repeating indices is assumed. Since the operator inside the trace per volume is covariant, we can apply Birkhoff's ergodic theorem [18] to equivalently write

$$\Delta\alpha(\gamma) = \frac{1}{2} \epsilon^{i_1 \dots i_4} \int_0^1 ds \int_\Omega d\omega \sum_\alpha \times \langle \mathbf{0}, \alpha | P_\omega(\xi_s) \partial_{i_1} P_\omega(\xi_s) \dots \partial_{i_4} P_\omega(\xi_s) | \mathbf{0}, \alpha \rangle. \quad (50)$$

This shows at once the self-averaging property of the magnetolectric response function.

Next, we show that if the deformation occurs along the TRS mirrored path  $\Theta\gamma$ , then

$$\Delta\alpha(\Theta\gamma) = -\Delta\alpha(\gamma). \quad (51)$$

This property is well known for periodic [4] and disordered [28] TRS insulators, but here we compute the magnetoelectric effect for a system in a disorder configuration which breaks the TRS. The proof proceeds in the same way as for polarization, using the self-averaging property, the compatibility between TRS and disorder, together with the behavior of the derivatives under TRS:  $\Theta \partial_j \Theta^{-1} = -\partial_j$ , for  $j = 1, 2, 3$ , and  $\Theta \partial_s \Theta^{-1} = \partial_s$ . The proof goes as follows:

$$\Delta\alpha(\Theta\gamma) = \frac{1}{2} \epsilon^{i_1 \dots i_4} \int_0^1 ds \int_{\Omega} d\omega \sum_{\alpha} \langle \mathbf{0}, \alpha | P_{\omega}(\Theta \xi_s) \partial_{i_1} P_{\omega}(\Theta \xi_s) \dots \partial_{i_4} P_{\omega}(\Theta \xi_s) | \mathbf{0}, \alpha \rangle. \quad (52)$$

Recall that

$$P_{\omega}(\Theta \xi_s) = \Theta P_{\Theta^{-1}\omega}(\xi_s) \Theta^{-1}. \quad (53)$$

Then, after a change of variable  $\omega \rightarrow \Theta\omega$ ,

$$\dots = -\frac{1}{2} \epsilon^{i_1 \dots i_4} \int_0^1 ds \int_{\Omega} d(\Theta\omega) \sum_{\alpha} \langle \mathbf{0}, \alpha | \Theta P_{\omega}(\xi_s) \partial_{i_1} P_{\omega}(\xi_s) \dots \partial_{i_4} P_{\omega}(\xi_s) \Theta^{-1} | \mathbf{0}, \alpha \rangle. \quad (54)$$

Then the statement follows from the explicit action of  $\Theta$  given in Eq. (21) and the compatibility between TRS and disorder,  $d(\Theta\omega) = d\omega$ .

Lastly, given a path  $\gamma$  between two TRS fixed points,  $\xi_0 = \Theta \xi_0$  and  $\xi_1 = \Theta \xi_1$ , one can close this path into a loop by augmenting with its TRS mirrored image  $-\Theta\gamma$ , taken with opposite orientation. Note that this argument will not work if the end points of  $\gamma$  are not TRS fixed points. Then,

$$\Delta\alpha(\gamma) = \frac{1}{2} [\Delta\alpha(\gamma) + \Delta\alpha(-\Theta\gamma)] = \frac{1}{2} \Delta\alpha(\gamma - \Theta\gamma). \quad (55)$$

For the periodic TRS case, it is well known [4] that the variation of the magnetoelectric response function along a closed loop is equal to a second Chern number. As demonstrated in Ref. [28], this remains true in the disordered case, in which case the connection is with the noncommutative second Chern number introduced in Ref. [29]. More precisely,

$$\Delta\alpha(\gamma - \Theta\gamma) = C_2[(\gamma - \Theta\gamma) \times \widetilde{\mathbb{T}}^3], \quad (56)$$

where on the right we have the second Chern number of the Fermi projection  $P_{\omega}(\xi)$  over the manifold  $(\gamma - \Theta\gamma)$  times the 3-dimensional noncommutative torus:

$$C_2 = \epsilon^{i_1 \dots i_4} \int_{\gamma - \Theta\gamma} d\xi \mathcal{T}(P_{\omega}(\xi) \partial_{i_1} P_{\omega}(\xi) \dots \partial_{i_4} P_{\omega}(\xi)). \quad (57)$$

This is a strong topological invariant which is known [29] to take only integer values. Note that the constant in front differs from the usual constant because some of the derivatives are taken in  $k$  space and some in real space.

In conclusion, when 3-dimensional insulators are deformed between fixed points of the time-reversal operation, the change in the isotropic magnetoelectric response function is quantized in units of  $\frac{1}{2}$ , even though TRS applies only for disorder averages. Recall that the existence of a spectral gap is required by our argument. The numerical simulations of  $\Delta\alpha$  performed for Ref. [30] indicate that, as in the previous case, the quantization does not survive beyond the spectral gap closing.

#### IV. CONCLUSIONS

The present work dealt exclusively with the bulk invariants, while the invariants formulated in Ref. [1] can be regarded as boundary invariants since they are computed from the boundary states. We can already foresee a connection between these invariants, which we would like to sketch briefly. The bulk-boundary principle developed in Ref. [31] provides an equality between the bulk first Chern number and a certain spectral flow of the boundary states. The latter has a self-averaging property similar to that of the first Chern number. When applied to the variation of the electric polarization  $\Delta P(\gamma - \mathcal{I}\gamma)$  of a 1-dimensional system with an edge, this bulk-boundary principle seems to lead precisely to the counting of the spectral features of the edge states performed in Ref. [1]. It will definitely be interesting to make this connection more precise and see it in action for concrete models.

The principle described in the present work seems to apply to any symmetry-stabilized topological invariant which can be formulated in a real-space representation. Unfortunately, there are important instances where a real-space representation is not yet available, such as the Kane-Mele  $\mathbb{Z}_2$  invariant [32] or the bulk topological invariants for point-symmetry-stabilized topological insulators [33]. For this reason, we have nothing to say about these invariants at this moment, but we hope our findings will spur a renewed effort in this direction. Nevertheless, the strategy does apply to the Loring-Hastings invariants [34,35], or to the spin-Chern numbers [36,37].

#### ACKNOWLEDGMENTS

This work was supported by US NSF Grant No. DMR-1056168 and NSFC Grant No. 11204065 and NSFHPC Grant No. A2013205168.

- [1] I. C. Fulga, B. van Heck, J. M. Edge, and A. R. Akhmerov, Statistical topological insulators, *Phys. Rev. B* **89**, 155424 (2014).
- [2] F. Ortman, S. Roche, and S. O. Valenzuela (eds.), *Topological Insulators: Fundamentals and Perspectives* (Wiley-VCH Verlag, Weinheim, Germany, 2015).
- [3] L. Fu and C. L. Kane, Topology, Delocalization via Average Symmetry and the Symplectic Anderson Transition, *Phys. Rev. Lett.* **109**, 246605 (2012).

- [4] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Topological field theory of time-reversal-invariant insulators, *Phys. Rev. B* **78**, 195424 (2008).
- [5] R. S. K. Mong, A. M. Essin, and J. E. Moore, Antiferromagnetic topological insulators, *Phys. Rev. B* **81**, 245209 (2010).
- [6] P. Baireuther, J. M. Edge, I. C. Fulga, and C. W. J. Beenakker, Quantum phase transitions of a disordered anti-ferromagnetic topological insulator, *Phys. Rev. B* **89**, 035410 (2014).

- [7] Z. Ringel, Y. E. Kraus, and A. Stern, Strong side of weak topological insulators, *Phys. Rev. B* **86**, 045102 (2012)
- [8] R. S. K. Mong, J. H. Bardarson, and J. E. Moore, Quantum Transport and Two-Parameter Scaling at the Surface of a Weak Topological Insulator, *Phys. Rev. Lett.* **108**, 076804 (2012).
- [9] L. Fu, Topological Crystalline Insulators, *Phys. Rev. Lett.* **106**, 106802 (2011).
- [10] Y. Ando and L. Fu, Topological crystalline insulators and topological superconductors: From concepts to materials, *Annu. Rev. Condens. Matter Phys.* **6**, 361 (2015).
- [11] A. M. Turner, Y. Zhang, R. S. K. Mong, and A. Vishwanath, Quantized response and topology of insulators with inversion symmetry, *Phys. Rev. B* **85**, 165120 (2012).
- [12] T. L. Hughes, E. Prodan, and B. A. Bernevig, Inversion-symmetric topological insulators, *Phys. Rev. B* **83**, 245132 (2011).
- [13] G. Y. Cho, C.-T. Hsieh, T. Morimoto, and S. Ryu, Topological phases protected by reflection symmetry and cross-cap states, *Phys. Rev. B* **91**, 195142 (2015).
- [14] M. Diez, D. I. Pikulin, I. C. Fulga, and J. Tworzydło, Extended topological group structure due to average reflection symmetry, *New J. Phys.* **17**, 043014 (2015).
- [15] A. Alexandradinata, C. Fang, M. J. Gilbert, and B. A. Bernevig, Spin-Orbit-Free Topological Insulators without Time-Reversal Symmetry, *Phys. Rev. Lett.* **113**, 116403 (2014).
- [16] J. Bellissard, Coherent and dissipative transport in aperiodic solids: An overview, *Lect. Notes Phys.* **597**, 413 (2002).
- [17] J. Bellissard,  $K$  theory of  $C^*$  algebras in solid state physics, *Lect. Notes Phys.* **257**, 99 (1986).
- [18] G. D. Birkhoff, Proof of the ergodic theorem, *Proc. Natl. Acad. Sci. USA* **17**, 656 (1931).
- [19] R. Resta and D. Vanderbilt, Theory of polarization: A modern approach, *Top. Appl. Phys.* **105**, 31 (2007).
- [20] R. Resta, Electrical polarization and orbital magnetization: The modern theories, *J. Phys.: Condens. Matter* **22**, 123201 (2010).
- [21] H. Schulz-Baldes and S. Teufel, *Commun. Math. Phys.* **319**, 649 (2013).
- [22] R. D. King-Smith and D. Vanderbilt, Theory of polarization of crystalline solids, *Phys. Rev. B* **47**, 1651 (1993).
- [23] G. De Nittis and M. Lein, Topological polarization in graphene-like systems, *J. Phys. A* **46**, 385001 (2013).
- [24] J. Bellissard, A. van Elst, and H. Schulz-Baldes, The noncommutative geometry of the quantum Hall effect, *J. Math. Phys.* **35**, 5373 (1994).
- [25] M. J. Rice and E. J. Mele, Elementary Excitations of a Linearly Conjugated Diatomic Polymer, *Phys. Rev. Lett.* **49**, 1455 (1982).
- [26] D. Xiao, Ming-Che Chang, and Qian Niu, Berry phase effects on electronic properties, *Rev. Mod. Phys.* **82**, 1959 (2010).
- [27] E. Prodan, Quantum transport in disordered systems under magnetic fields: A study based on operator algebras, *Appl. Math. Res. Express* **2013**, 176 (2013).
- [28] B. Leung and E. Prodan, A noncommutative formula for the isotropic magnetoelectric response, *J. Phys. A* **46**, 085205 (2012).
- [29] E. Prodan, B. Leung, and J. Bellissard, The noncommutative  $n$ th Chern number ( $n \geq 1$ ), *J. Phys. A* **46**, 485202 (2013).
- [30] B. Leung, Ph.D. thesis, Rutgers University, 2013.
- [31] H. Schulz-Baldes, J. Kellendonk, and T. Richter, Simultaneous quantization of edge and bulk Hall conductivity, *J. Phys. A* **33**, L27 (2000).
- [32] C. L. Kane and E. J. Mele,  $Z_2$  Topological Order and the Quantum Spin Hall Effect, *Phys. Rev. Lett.* **95**, 146802 (2005).
- [33] C. Fang, M. J. Gilbert, and B. A. Bernevig, Bulk topological invariants in noninteracting point group symmetric insulators, *Phys. Rev. B* **86**, 115112 (2012).
- [34] T. A. Loring and M. B. Hastings, Disordered topological insulators via  $C^*$  algebras, *Europhys. Lett.* **92**, 67004 (2010).
- [35] T. A. Loring and M. B. Hastings, Topological insulators and  $C^*$  algebras: Theory and numerical practice, *Ann. Phys.* **326**, 1699 (2011).
- [36] D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, Quantum Spin-Hall Effect and Topologically Invariant Chern Numbers, *Phys. Rev. Lett.* **97**, 036808 (2006).
- [37] E. Prodan, Robustness of the spin-Chern number, *Phys. Rev. B* **80**, 125327 (2009).