

Smooth gauge and Wannier functions for topological band structures in arbitrary dimensions

Georg W. Winkler, Alexey A. Soluyanov, and Matthias Troyer

Institute for Theoretical Physics and Station Q, ETH Zurich, 8093 Zurich, Switzerland

(Received 22 September 2015; revised manuscript received 14 January 2016; published 28 January 2016)

The construction of exponentially localized Wannier functions for a set of bands requires a choice of Bloch-like functions that span the space of the bands in question, and are smooth and periodic functions of \mathbf{k} in the entire Brillouin zone. For bands with nontrivial topology, such smooth Bloch functions can only be chosen such that they do not respect the symmetries that protect the topology. This symmetry breaking is a necessary, but not sufficient condition for smoothness, and, in general, finding smooth Bloch functions for topological bands is a complicated task. We present a generic technique for finding smooth Bloch functions and constructing exponentially localized Wannier functions in the presence of nontrivial topology, given that the net Chern number of the bands in question vanishes. The technique is verified against known results in the Kane-Mele model. It is then applied to the topological insulator Bi_2Se_3 , where the topological state is protected by two symmetries: time reversal and inversion. The resultant exponentially localized Wannier functions break both these symmetries. Finally, we illustrate how the calculation of the Chern-Simons orbital magnetoelectric response is facilitated by the proposed smooth gauge construction.

DOI: [10.1103/PhysRevB.93.035453](https://doi.org/10.1103/PhysRevB.93.035453)

I. INTRODUCTION

Discoveries of the past decade established that not only the geometry of electronic bands but also their topology is a fundamental material property [1,2]. Being at the root of a variety of physical effects [3,4], topologically nontrivial bands turned out to be ubiquitous, and many materials were found to host various topological states [5–9].

As a general principle, nontrivial topologies in band structures are a consequence of some symmetry. They are examples of symmetry protected topological states [10,11]. For instance, magnetic insulators in two dimensions (2D) realize the integer quantum Hall effect in the absence of external magnetic field [12,13]. In such insulators, dubbed Chern insulators, the nontrivial topology is captured by a Chern number topological invariant [14], which cannot be changed unless the $U(1)$ charge conservation symmetry is broken [15]. In the case of time-reversal (TR) symmetric insulators, a \mathbb{Z}_2 topological invariant is assigned [16] to the band structure that cannot be changed without breaking the TR symmetry (provided the band gap remains open). This concept of symmetry protection can also be generalized to crystalline symmetries, which protect the value of certain topological invariants, associated with these symmetries [17,18], giving rise to crystalline topological insulators [9,19,20].

The topology of a band structure in all these cases is associated with the Bloch bands and their Berry curvature [21,22] in momentum space. An alternative description of crystalline solids can be made in position space by means of localized Wannier functions (WFs) [23,24]. Such a local basis is often preferable to that of Bloch functions, for example for modeling finite size effects in materials [25], chemical bonding [26], computing electronic polarization [27], or ballistic transport [28]. This variety of applications and the existence of established numerical techniques for WF-based analysis of materials [24,29], makes it important to extend the phenomenology of WFs to materials with nontrivial topology.

Another motivation for finding Wannier representation of topological bands is the calculation of the isotropic contribu-

tion to the orbital magnetoelectric response [30]. The linear magnetoelectric coupling tensor α_{ij} is defined as

$$\alpha_{ij} = \left(\frac{\partial P_i}{\partial B_j} \right)_{\mathbf{E}=0} = \left(\frac{\partial M_j}{\partial E_i} \right)_{\mathbf{B}=0}, \quad (1)$$

where \mathbf{P} and \mathbf{M} are the polarization and magnetization of a material, \mathbf{E} and \mathbf{B} are electric and magnetic fields, and both derivatives are evaluated at zero fields. While lattice and spin degrees of freedom also contribute [31,32] to α_{ij} , here we focus solely on the orbital (frozen-lattice) part. The orbital magnetoelectric response can be further split into two parts [33,34]:

$$\alpha_{ij} = \tilde{\alpha}_{ij} + \frac{\theta e^2}{2\pi h} \delta_{ij}, \quad (2)$$

of which $\tilde{\alpha}_{ij}$ is traceless, and the second, isotropic, part is characterized by a dimensionless quantity θ , called the “axion angle” in high energy physics [3,35]. There are two contributions to θ [34]: one is an ordinary perturbative Kubo term and the other is a purely geometrical one, the Chern-Simons contribution θ_{CS} . The latter one can be evaluated from the ground-state electron wave functions, by computing an integral of the Chern-Simons 3-form over the entire Brillouin zone (BZ) [3,33,34]. However, this evaluation requires a choice of Bloch states that are smooth and periodic in \mathbf{k} space as detailed in Sec. V D. Alternative formulations of this term in position space require the existence of exponentially localized WFs (ELWFs) [30].

Magnetoelectric response was long thought to be observable only in materials that break TR and inversion symmetries. Remarkably, it turned out that TR- and inversion-symmetric topological insulators (TIs) are characterized by a nonzero quantized magnetoelectric response [3,18,33,36]. In the presence of one of these symmetries the Kubo contribution to θ vanishes and $\theta = \theta_{\text{CS}}$. The seeming contradiction is resolved by noticing that θ couples to the $\mathbf{E} \cdot \mathbf{B}$ term in the Lagrangian, so that the corresponding equations of motion are invariant under $\theta \rightarrow \theta + 2\pi$ [35]. Since, in addition, θ is odd under both

TR and inversion [35], the *two* values $\theta = 0, \pi$ are compatible with these symmetries.

In TR- or inversion-symmetric TIs these two values of θ can be used as an analog of the \mathbb{Z}_2 topological invariant [3,33], with $\theta = \pi$ corresponding to the TI phase and $\theta = 0$ to the normal insulator (NI) one. For the cases of TR- and inversion-symmetric insulators the quantization of θ is exact and since θ corresponds to the \mathbb{Z}_2 invariant, it can be obtained using the methods for computing this invariant [37–39]. However, for materials that lack these symmetries, θ is in general not quantized, and it becomes necessary to evaluate the θ_{CS} term directly. As mentioned above, this requires ELWFs. Since nontrivial band topologies can arise even when there are no symmetries that quantize the magnetoelectric response, computation of θ_{CS} gives yet another motivation for developing a method to obtain ELWFs for topologically nontrivial band structures.

WFs are constructed by Fourier transforming Bloch states, and thus the momentum space geometry of the Bloch function can strongly influence the properties of the resultant WFs, in particular the degree of its localization. An important example is that the construction of an ELWF for a Bloch state with nonzero Chern number is impossible [40]. This is a consequence of the fact that a Chern number represents an obstruction for choosing the Bloch state to be a smooth and periodic function of \mathbf{k} globally in the whole BZ [41].

ELWFs can only be constructed if all the Bloch states, for which a Wannier representation is constructed, are smooth and periodic in the whole BZ. For topologically trivial bands, the choice of smooth Bloch states, referred to as a smooth gauge choice, respecting all the symmetries of the underlying band structure is generally possible. For topological bands, however, the symmetry that protects the topology represents an obstruction to choosing a smooth gauge that respects this symmetry. Consequently, the smooth WFs have to break this symmetry. This was explicitly illustrated to be the case for the time-reversal (TR) symmetric \mathbb{Z}_2 TIs [37,42–44].

While it is clear that the construction of a smooth gauge on a lattice requires breaking certain symmetries in the gauge, finding an explicit representation of smooth Bloch states on a lattice of \mathbf{k} points, required for a numerical construction of ELWFs, is a nontrivial task. An explicit construction based on parallel transport of the occupied states was obtained in Ref. [45] for a 2D model of a quantum spin Hall insulator. That construction, however, is tedious to generalize to the many-band case, and especially to higher dimensions.

Another approach to finding a smooth gauge is based on projecting certain localized orbitals that break the topology-protecting symmetry onto the occupied states (see Sec. II for details) and is more appropriate for material calculations [43]. The problem of this method is that no specific algorithm for choosing the orbitals for the projection is presented. The requirement of breaking the topology-protecting symmetry in the initial projection is necessary but not sufficient for finding ELWFs.

In this work we develop a generic algorithm to construct ELWFs for a set of topologically nontrivial bands, with a zero net Chern number. The idea is to construct an adiabatic connection between trivial and TI phases by breaking the symmetries that protect the topology everywhere along the

connecting path, except the initial and final points. Since the topology-protecting symmetries are broken at intermediate steps, it becomes possible to find a path connecting these two phases in a parameter space such that the insulating gap remains open along the connection.

The construction of this path allows one to avoid the problem of finding the correct symmetry-breaking projection in the topological phase. Instead, one discretizes the path in parameter space into steps, and constructs ELWFs at each step by projecting onto ELWFs found at the previous step. The initial step requires finding ELWFs in a topologically trivial case, which is a standard task [26]. As a result, one naturally obtains ELWFs for the topological bands at the final point of the discretized path.

The breaking of symmetries plays a key role in adiabatically connecting two topologies. It is often the case that there are more than one symmetry protecting the topology of the band structure, and it is important that all of them are broken along the adiabatic path. For example, the \mathbb{Z}_2 topology of the TR-symmetric TIs can be additionally protected by inversion [18,46], in-plane mirror [47,48], and certain other point group symmetries [49]. These symmetries need to be broken in addition to TR along the adiabatic path to obtain a smooth gauge in such band structures.

This paper is organized as follows. In Sec. II we provide the theoretical background about construction of ELWFs. In this light we also discuss the topological obstruction for constructing ELWFs in TIs. In Secs. III and IV we apply our technique on the model of Kane and Mele [16]. Then in Sec. V we apply our technique onto Bi_2Se_3 and use it to calculate the Chern-Simons magnetoelectric coupling θ_{CS} . Finally, we summarize our findings and give an outlook in Sec. VI.

II. WANNIER FUNCTIONS AND TOPOLOGICAL OBSTRUCTION

A. Construction of Wannier functions

Here we provide a brief review of the methods introduced in the work of Ref. [26] used for an explicit construction of ELWFs starting from a set of Bloch states obtained by diagonalizing the Hamiltonian on a mesh of \mathbf{k} points. The construction requires a projection of a set of trial localized orbitals onto the Bloch states. The choice of these orbitals, which is simple in the case of topologically trivial bands, becomes problematic for the case of nontrivial topology. For the method we present in this work, localized trial orbitals need to be found only in the topologically trivial case.

Given a Bloch state $\psi_{n\mathbf{k}}(\mathbf{r})$, a corresponding WF is defined as

$$\langle \mathbf{r} | \mathbf{R} n \rangle = W_n(\mathbf{r} - \mathbf{R}) = \frac{V}{(2\pi)^d} \int_{\text{BZ}} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (3)$$

where V is the volume of the unit cell and d is the dimensionality. The Bloch wave functions $\psi_{n\mathbf{k}}$ are assumed to be normalized within the unit cell.

This definition, however, is not unique. The nonuniqueness, referred to as gauge freedom, is easily seen when constructing WFs for a set of N Bloch states. A general unitary

transformation $U(N)$ of the N occupied bands

$$|\tilde{\psi}_{n\mathbf{k}}\rangle = \sum_m U_{mn}(\mathbf{k}) |\psi_{m\mathbf{k}}\rangle, \quad (4)$$

results in a different set of Bloch states that span the same Hilbert space as the original ones, and hence can equally be used for constructing the Wannier representation. Depending on the particular gauge choice, the resultant WFs and their degree of localization can vary a lot. In order to obtain ELWFs, the gauge choice has to be smooth, meaning that all N Bloch states used to construct the WFs are smooth in the entire BZ, and obey periodic boundary conditions $\psi_{n\mathbf{k}} = \psi_{n\mathbf{k}+\mathbf{G}}$ upon translation by any reciprocal lattice vector \mathbf{G} .

But even such a smooth gauge choice is not unique, since a smooth gauge transformation performed on a set of smooth Bloch states will result in a different set of smooth Bloch states and different ELWFs. Additional constraints can be put on the gauge to reduce the gauge freedom. A very common choice of such a constraint is the requirement of maximal localization of the resultant WFs in position space proposed in the work of Ref. [26]. This gauge is obtained by minimizing the spread functional

$$\Omega = \sum_{n=1}^N (\langle \mathbf{r}^2 \rangle_n - \langle \mathbf{r} \rangle_n^2) = \Omega_I + \tilde{\Omega}, \quad (5)$$

where $\langle \mathbf{r} \rangle_n = \int |\mathbf{r}| W_n^2 d\mathbf{r}$, and Ω_I and $\tilde{\Omega}$ stand for the gauge-independent

$$\Omega_I = \sum_{n=1}^N \left[\langle \mathbf{r}^2 \rangle_n - \sum_{\mathbf{R}m} |\langle \mathbf{R}m | \mathbf{r} | 0n \rangle|^2 \right] \quad (6)$$

and gauge-dependent parts of the spread

$$\tilde{\Omega} = \sum_{n=1}^N \sum_{\mathbf{R}m \neq 0n} |\langle \mathbf{R}m | \mathbf{r} | 0n \rangle|^2. \quad (7)$$

The goal of maximal localization is to find a $U(N)$ transformation that, when applied to some initial set of Bloch states according to Eq. (4), minimizes $\tilde{\Omega}$ to produce a set of maximally localized WFs. Maximal localization can always be performed, once the initial choice of Bloch states is smooth.

The necessity for smoothness of the initial Bloch states is most easily seen from the general procedure used to construct a set of WFs, as described in Ref. [26]. Following this procedure, to construct a set of N WFs from N isolated (that is, separated by energy gaps from the rest of the spectrum) Bloch bands obtained from numerical diagonalization of the Hamiltonian, a set of N localized trial states $|\tau_i\rangle$ is chosen. For each momentum \mathbf{k} all of the trial orbitals are projected on the Bloch states to get a set of N Bloch-like states [26]

$$|\Upsilon_{i\mathbf{k}}\rangle = \hat{P}_{\mathbf{k}} |\tau_i\rangle = \sum_{n=1}^N |\psi_{n\mathbf{k}}\rangle \langle \psi_{n\mathbf{k}} | \tau_i \rangle, \quad (8)$$

which are not orthonormal.

For the construction of WFs these states need to be orthonormalized, and it is the orthonormalization procedure, where the smoothness of the gauge becomes crucial. To see this, apply Löwdin orthonormalization procedure to the states,

which is commonly used to get a set of orthonormalized Bloch states, to orthonormalize the states $|\Upsilon_{i\mathbf{k}}\rangle$,

$$|\tilde{\psi}_{n\mathbf{k}}\rangle = \sum_m [S(\mathbf{k})^{-1/2}]_{mn} |\Upsilon_{m\mathbf{k}}\rangle, \quad (9)$$

where

$$S_{mn}(\mathbf{k}) = \langle \Upsilon_{m\mathbf{k}} | \Upsilon_{n\mathbf{k}} \rangle \quad (10)$$

is the overlap matrix. While $|\tilde{\psi}_{n\mathbf{k}}\rangle$ are not the eigenstates of the single particle Hamiltonian, they span the same space as the usual Bloch eigenstates, and therefore describe the same ground state. For the trial states embodying a reasonable assumption about the character of the bands described, the $|\tilde{\psi}_{n\mathbf{k}}\rangle$ will be smooth functions of \mathbf{k} . In this case the WFs constructed from them by means of Eq. (3) are expected to be exponentially localized, and the degree of localization is further increased by doing maximal localization.

However, the orthonormalization procedure breaks down if at some \mathbf{k} the determinant of the overlap matrix vanishes, that is $\det S(\mathbf{k}) = 0$. This is guaranteed to happen if one runs into a topological obstruction, for example, by projecting onto a set of trial states $|\tau_i\rangle$ that respect the topology-protecting symmetry, as shown below.

B. Topological obstruction via hybrid Wannier functions

Hybrid WFs (HWFs) [50] differ from usual WFs in that the Wannier decomposition is done in one direction only

$$|R_y k_x n\rangle = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} dk_y e^{-iR_y k_y} |\tilde{\psi}_{n\mathbf{k}}\rangle. \quad (11)$$

This wave function is localized in momentum in the x direction and in position in the y direction, being a hybrid of Bloch- and Wannier-like functions. HWFs proved to be useful for classifying topologies of band structures [38,51], providing an intuitive approach to topological invariants.

The band structure topology can be obtained by tracking the charge centers of HWFs defined as

$$\bar{y}_n(k_x) = \langle 0k_x n | \hat{y} | 0k_x n \rangle, \quad (12)$$

where the HWFs are located within the home unit cell. Computation of the charge centers does not require an explicit construction of HWFs, and can be done by means of the parallel transport procedure, as described in Ref. [38]. The HWF charge centers obtained in this gauge [52] are the eigenvalues of the projected position operator [53]. However, for an isolated set of N bands the gauge can be chosen differently, resulting in different values of $\bar{y}_n(k_x)$, and it is only the sum of all the N centers that is gauge invariant (modulo a lattice vector in the y direction) at each k_x [38].

The parallel transport gauge respects the symmetries of the system. For example, in the presence of TR symmetry the HWF charge centers come in Kramers pairs $y_n(k_x) = y_m(-k_x)$ being doubly degenerate at the TR-invariant momenta $-k_x^* = k_x^* + G_x$, where G_x is a reciprocal lattice vector in the x direction. This is illustrated in Figs. 1(a) and 1(b) for a model with two occupied bands. The two centers are degenerate at $k_x = 0$ and $k_x = \pm\pi/a$. Both centers evolve smoothly in between $\pm\pi/a$. In the case of Fig. 1(a) they return to the original value at the BZ boundary and the band structure is topologically trivial. In

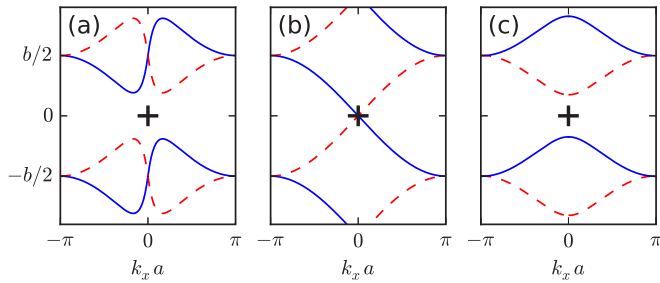


FIG. 1. Possible flows of the hybrid Wannier charge centers in a 2D TR-symmetric system. (a) Topologically trivial insulator. (b) and (c) Quantum spin Hall insulator. A TR-symmetric gauge is used in (a) and (b), while a TR-breaking gauge is used in (c).

the case of Fig. 1(b) they interchange, which is generally the case in a quantum spin Hall insulator [37].

This interchange results in a discontinuity of the hybrid Wannier charge center lines as a function k_x : the two centers are continuous functions of momentum for $k_x \in [-\pi/a, \pi/a]$, but the periodicity constraint is not satisfied, meaning that the center position does not necessarily return to the original value after a $2\pi/a$ change in momentum. This is equivalent to placing the topology-dictated gauge discontinuity on the boundary of the BZ. In this gauge choice, the Wannier centers each correspond to a particular, *individual* Chern number, since the shift of the Wannier center at the boundary is still an integer number of unit cells. Hence, each of the hybrid Wannier functions can be understood as a well-defined function, in a sense that a Chern number can be assigned to each of them, despite the possible presence of degeneracies in the energy spectrum.

For example, the two bands in the Kramers pair are degenerate at the TR-symmetric momenta, but the corresponding hybrid Wannier centers result in a splitting of the pair into two subspaces with well-defined Chern numbers. For a trivial insulator, illustrated in Fig. 1(a), the two centers correspond to zero Chern numbers. For the case of a quantum spin Hall insulator the hybrid Wannier centers shown in Fig. 1(b) indicate that in the chosen gauge the occupied space is split into two subspaces with Chern numbers equal ± 1 [54]. In both cases the two subspaces are TR images of each other, but the nonzero Chern numbers of the two states in the topological phase signal that the TR-symmetric gauge is not smooth on the whole BZ torus. The above argument is equally valid for more than two occupied bands, where each of the two subspace contains several bands and Chern numbers are then assigned to each subspace.

The Kramers degeneracy of the Wannier centers persists for any gauge, in which TR maps one state at \mathbf{k} onto the other state at $-\mathbf{k}$ [45]. For this reason TR symmetry in the gauge represents an obstruction for it to be smooth and periodic in the BZ. Tracking the presence of this obstruction is the basis for the methods of computing topological invariants [37–39].

The above analysis, in accord with other methods [37,42,44], leads to the conclusion that in the \mathbb{Z}_2 TIs a smooth gauge, and hence ELWFs, have to break TR symmetry. Breaking the symmetry in the gauge (but not in the Hamiltonian) lifts the Kramers degeneracy of Wannier

centers at the TR invariant momenta, and can result in a smooth gauge [43] like the one illustrated in Fig. 1(c). The Wannier centers are smooth and single valued throughout the BZ. This means that the occupied subspace was split [45] into two subspaces with zero Chern numbers, and hence the construction of the ELWFs becomes possible in this gauge.

The above line of reasoning can be easily generalized to band structures, where the nontrivial topology is protected by a symmetry different from TR, for instance a crystalline symmetry [17,18,36,55]. A particular example can be that of the crystalline TI SnTe [9,19], where the mirror Chern numbers [47] of ± 2 can be defined on the $\{110\}$ mirror plane in the BZ. Hence, the mirror symmetry should be broken in the gauge to obtain ELWFs for this material. Again, not any symmetry-breaking gauge would work, but only specific ones that provide a decomposition of the occupied space into states that are smooth in the BZ.

III. AVOIDING TOPOLOGICAL OBSTRUCTION

Finding suitable initial projections, for which $\det S(\mathbf{k}) \neq 0$ in the whole BZ, is particularly difficult for TIs due to the band inversion and hence, band character change, present in these materials. According to the discussion above and also the previous findings [43], the initial projections must be chosen such that they break the topology-protecting symmetries. However, as mentioned above, this requirement is necessary, but not sufficient. Given that with an increasing number of bands the number of possible projections increases exponentially, a brute force search for suitable projections is complicated and inefficient.

Here we describe a technique for obtaining a smooth gauge and ELWFs for topological bands that avoids the need of finding suitable initial projections. Instead, we construct an adiabatic path that connects a topologically trivial Hamiltonian to the topological one, such that the system remains gapped along the path, as illustrated in Fig. 2. To keep the system gapped, all the symmetries that protect the topology need to be broken along the path. The initial topologically trivial Hamiltonian can be chosen in different ways. For example, in calculation for a material, where the nontrivial topology is often driven by spin-orbit coupling (SOC), the Hamiltonian

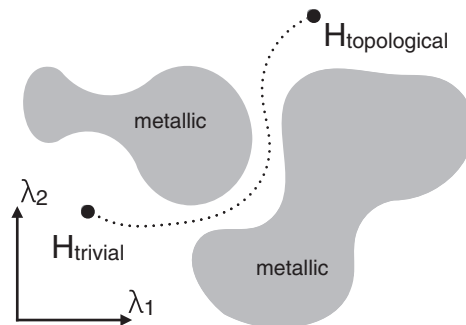


FIG. 2. Schematic path in a parameter space of some model Hamiltonian. The adiabatic path is chosen to connect the topologically trivial and nontrivial phases avoiding gap closures. The topology-protecting symmetries is broken along the path.

with SOC artificially set to zero is topologically trivial and can be used as a starting point, if it is gapped.

Finding suitable trial functions $|\tau_i\rangle$ for topologically trivial bands is usually straightforward, since they respect the symmetries of the Hamiltonian [26]. Moreover, an automated procedure for finding optimal projections for the trivial bands was proposed recently in Ref. [56]. Thus, the ELWFs at the initial step of the path can always be found.

At least two parameters are required to parametrize the adiabatic path. One controls the topological phase transition by tuning the strength of symmetry breaking along the path [57]. Care should be taken when breaking the symmetries, since *double symmetry protection* can occur, meaning that there is more than one symmetry protecting the nontrivial topology. The issue of double symmetry protection is discussed with more detail in Sec. III A.

Once the adiabatic path is constructed, the smooth Bloch states for the initial, topologically trivial, Hamiltonian H_0 , are found and used to construct ELWFs. After this is done the path is discretized into L steps, so that H_L is the Hamiltonian of the TI in question. At step 1 the Hamiltonian H_1 is diagonalized and the ELWFs obtained at the initial step are used as the trial orbitals for its occupied states. For dense enough discretization, the corresponding projection is guaranteed to have $\det S(\mathbf{k}) \neq 0$ throughout the BZ, since H_1 and H_0 are only slightly different. The resultant states are used to construct ELWFs for H_1 . The procedure continues by projecting the ELWFs obtained at the step ℓ onto the occupied state of the Hamiltonian $H_{\ell+1}$, until the final point of the path is reached. Since at each step ELWFs were found, one ends up with ELWF representation of the occupied topologically nontrivial space, which solves the problem of finding a smooth gauge [58].

We emphasize that this method is general, being applicable to any isolated set of bands and in any dimension, provided that the net Chern number of this set is zero.

Double symmetry protection

Here we discuss the double symmetry protection of the topological phase using two examples: \mathbb{Z}_2 TR-symmetric insulators with mirror or inversion symmetries.

We start by considering the case of coexisting mirror and TR symmetries. When a TR-symmetric plane in the BZ of the insulator is invariant under mirror symmetry, the nontrivial 2D \mathbb{Z}_2 invariant of this plane suggests that both TR and mirror symmetries need to be broken in the smooth gauge. The mirror symmetry breaking in the gauge means that the Bloch states $|\tilde{\psi}_{n\mathbf{k}}\rangle$ of Eq. (4) obtained in the smooth gauge and used to construct ELWFs are not eigenfunctions of the mirror operator on the mirror plane. This can be seen by noting that the two states in a mirror-symmetric Kramers pair have opposite mirror eigenvalues $\pm i$ on this plane, so the occupied space of an insulator on this plane can be split into two subspaces according to the value of the mirror eigenvalue. In the \mathbb{Z}_2 -odd phase $+i$ and $-i$ have opposite odd Chern numbers, and thus breaking TR symmetry only is not sufficient to remove the topological obstruction, since the two mirror-labeled subspaces remain nontrivial. Hence, the mirror symmetry has to be broken as well, to construct Bloch states that are smooth in the entire BZ for the \mathbb{Z}_2 -odd phase.

Another example of double protection is that of an inversion- and TR-symmetric TI. The smooth gauge in this case also has to break both symmetries, as we argue below. Several works [36,46] discussed the possibility of a topological classification in the presence of inversion-symmetry only. In particular, the work of Ref. [46] presented a study of the HWF centers in inversion symmetric insulators, obtained by a special construction (called a Wilson loop [39]). In that construction the projector onto all the occupied states $\hat{P}_{\mathbf{k}}^o$ is defined at each k point, and the HWF centers are obtained by taking the log of the eigenvalues of the product of operators $\prod_{\mathbf{k}} \hat{P}_{\mathbf{k}}^o$ taken over discretized values of k on a closed loop in momentum space [39]. For inversion symmetric systems the inversion symmetry \hat{I} puts the following constraint on the projector $\hat{P}_{\mathbf{k}}^o = \hat{I} \hat{P}_{-\mathbf{k}}^o \hat{I}$.

The resultant HWFs are inversion symmetric in a sense that for each center $\bar{y}_n(k_x)$ there exists an inversion symmetric partner $-\bar{y}_n(-k_x)$, where n is not necessarily equal to m . Thus, in the presence of inversion symmetry the sum of all the HWF centers at the inversion-invariant momenta $k_x = \{0, \pi/a\}$ can only be 0 or $b/2$, which are the inversion-symmetric values assuming the inversion center coincides with the center of the unit cell.

However, as discussed above, when constructing ELWFs, one needs a representation of the occupied subspace in terms of Bloch states that are smooth and periodic in \mathbf{k} space, meaning that the projector on each of these Bloch states is smooth. Thus, in a smooth gauge the net projector onto the occupied states is decomposed into a set of projectors $\hat{P}_{\mathbf{k}}^o = \sum \hat{P}_{\mathbf{k}}^n$, each of which is smooth. However, imposing inversion symmetry constraint of the form $\hat{P}_{\mathbf{k}}^n = \hat{I} \hat{P}_{-\mathbf{k}}^n \hat{I}$ on each of the individual projectors, restricts the corresponding HWF centers (not their sum, but each of them separately) to take on inversion-symmetric values at $k_x = \{0, \pi/a\}$, that is either 0 or $b/2$. We call a gauge, in which each of the projectors respects inversion symmetry in this sense an inversion-symmetric gauge.

In an inversion-symmetric gauge each Bloch state $|\tilde{\psi}_{n\mathbf{k}}\rangle$ respects inversion symmetry $e^{i\phi} |\tilde{\psi}_{n\mathbf{k}}\rangle = \hat{I} |\tilde{\psi}_{n-\mathbf{k}}\rangle$, and hence the corresponding WF centers are subject to the condition $\langle \mathbf{r} \rangle_n = -\langle \mathbf{r} \rangle_n \bmod \mathbf{R}$. Consequently, the HWF centers fulfill $\bar{y}_n(k_x) = -\bar{y}_n(-k_x)$ modulo b and are restricted to the values $\bar{y}_n(k_x) = \{0, b/2\}$ at $k_x = \{0, \pi/a\}$. Note that the HWFs are smooth in momentum in the interior of the BZ, and thus the index n refers to the same center on both sides of this equation. If for $k_x^* = \{0, \pi/a\}$, $|\tilde{\psi}_{n\mathbf{k}}\rangle$ has the same parity at $k_y = 0$ and π/b then $\bar{y}_n(k_x^*) = 0$, while if the parities are opposite $\bar{y}_n(k_x^*) = b/2$.

This means that if an inversion-symmetric Bloch state has a nonzero Chern number on some 2D BZ (2D cut of a 3D BZ), thus being not smooth, it is impossible to make it smooth (that is, change the Chern number to zero) if the inversion symmetry in the gauge is preserved. This is shown in Fig. 1. The illustrated centers are obtained in an inversion-symmetric gauge, and by construction are smooth in the interior of the BZ $k_x \in [-\pi/a, \pi/a]$. The Chern numbers of the centers shown as solid blue and dashed red lines in Fig. 1(b) are given by the number of unit cells traversed by the HWF center when going from one edge of the BZ to the other, and are equal to ± 1 . They cannot be changed to zero by breaking TR symmetry alone, while preserving inversion symmetry in the gauge in

the above sense, since that would require moving the center position away from the value fixed by inversion symmetry at some high-symmetry momentum. In the inversion-symmetric gauge, where each Bloch state separately is taken to be the eigenstate of an inversion operator, this argument holds for any number of Kramers pairs, since the argument applies for each individual HWF center.

To state it more rigorously, the TR and inversion invariant \mathbb{Z}_2 topological insulating state in 2D is characterized by the total number of negative parity eigenvalues of occupied states at the four TR invariant momenta [46,59]: if this number divided by two is odd the system is a topological insulator [60]. Let n_1^- be the difference in the number of negative parity eigenvalues between points $(0,0)$ and $(0,\pi/b)$, and n_2^- the difference between points $(\pi/a,0)$ and $(\pi/a,\pi/b)$. If n_1^- and n_2^- are different (which is the case in a TI) then the number of HWF centers for which $\bar{y}_n(k_x) = b/2$ is also different at $k_x = 0$ and $k_x = \pi/a$. Since the HWF centers are smooth in the BZ interior $k_x \in [-\pi/a, \pi/a]$, at least $|n_1^- - n_2^-|$ of the hybrid centers $\bar{y}_n(k_x)$ correspond to a nonzero Chern number. These Chern numbers can only be removed by breaking inversion symmetry of individual $|\psi_{n\mathbf{k}}\rangle$, similar to the TR symmetric case discussed in Sec. II B. Therefore, the obstruction to smoothness persists in the inversion symmetric gauge. Note that also in a \mathbb{Z}_2 -even case inversion symmetry can protect additional topologies as has been pointed out in Refs. [36,46].

Thus, the adiabatic connection of an inversion-symmetric topological insulator to a trivial one should be found by breaking the inversion symmetry along the path.

IV. APPLICATION TO KANE-MELE MODEL

We first illustrate our technique by applying it to the Kane-Mele (KM) model that describes a 2D quantum spin Hall (\mathbb{Z}_2 -odd) insulator in some of its parameter space [16]. A smooth gauge and the corresponding WFs were obtained previously for this model by other methods [43,45], and it is instructive to validate our method versus known results before going to more complicated cases.

A. Kane-Mele model

The KM is a tight-binding model on a honeycomb lattice with one spinor orbital per site. The primitive hexagonal lattice vectors are $\mathbf{a}_1 = a\hat{x}$ and $\mathbf{a}_2 = \frac{a}{2}(\hat{x} + \frac{\sqrt{3}}{2}\hat{y})$ with the atoms A and B located at sites $\mathbf{t}_A = \frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2)$ and $\mathbf{t}_B = \frac{2}{3}(\mathbf{a}_1 + \mathbf{a}_2)$. The KM Hamiltonian is given by

$$H = \sum_{\langle ij \rangle} t_{ij} c_i^\dagger c_j + i\lambda_{SO} \sum_{\langle\langle ij \rangle\rangle} v_{ij} c_i^\dagger \sigma_z c_j + i\lambda_R \sum_{\langle ij \rangle} c_i^\dagger (\mathbf{s} \times \hat{\mathbf{d}}_{ij})_z c_j + \lambda_v \sum_i \xi_i c_i^\dagger c_i, \quad (13)$$

where summation over the suppressed spin indices is assumed. The summation $\langle ij \rangle$ runs over all nearest neighbors, and the sum over $\langle\langle ij \rangle\rangle$ runs over all second-nearest neighbors. $v_{ij} = (2/\sqrt{3})(\hat{\mathbf{d}}_1 \times \hat{\mathbf{d}}_2) = \pm 1$, with $\hat{\mathbf{d}}_1$ and $\hat{\mathbf{d}}_2$ being the first-neighbor bond vectors encountered by an electron hopping from j to i , \mathbf{s} is a spin-1/2 operator. λ_{SO} and λ_R are parameters defining

the spin-orbit coupling and λ_v is a staggered on-site potential. In what follows we fix $t = 1$, $\lambda_{SO} = 0.6$, and $\lambda_R = 0.5$.

B. Constructing an obstruction-free path for the Kane-Mele model

Following our method, first an adiabatic path connecting the normal insulator (NI) to the TI phase of the KM model needs to be found. Let us first tune λ_v while keeping the spin-orbit coupling parameters λ_{SO} and λ_R fixed. This path corresponds to the vertical axis in Fig. 3. A gap closure occurs at the two Dirac points at K and K' in the BZ (illustrated in the inset of Fig. 3). We thus need to find a symmetry breaking field that would prevent the gap closure at both Dirac points simultaneously.

We thus introduce a hopping anisotropy δ to the nearest-neighbor hopping t_{ij} . For the reasons that are made clear below, it is chosen such that the hopping is $t(1 + \delta)$ in the $(1,1)$ direction and t in the other directions, and it breaks the C_3 -rotational symmetry of the KM model (see the Appendix A for a thorough discussion of the symmetries in the KM model). By tuning δ and moving along the $\lambda_v = 0$ line in Fig. 3 the gap closure can be moved to the M point. This path is depicted by the red arrows in Fig. 3.

The degeneracy at the M point is easily lifted by an appropriate weak TR-symmetry breaking field. We found the optimal field to be a staggered magnetic field in the $(-1,1)$ direction of the form [61]

$$H_{(2)}^{SB} = \left(-\frac{\sqrt{3}}{4}\sigma_x + \frac{3}{4}\sigma_y + \frac{1}{2}\sigma_z \right) \tau_z, \quad (14)$$

where σ and τ are the Pauli matrices acting in the spin and sublattice subspaces correspondingly. Apart from the TR symmetry, this field breaks the mirror and C_2 -rotational symmetries of the KM model (see Appendix A) and the resultant WFs coincide with the ones reported in Ref. [43].

The final path, shown by the red line in Fig. 3, is thus made of two parts: the line $\overline{P1P2}$, connecting two points with the

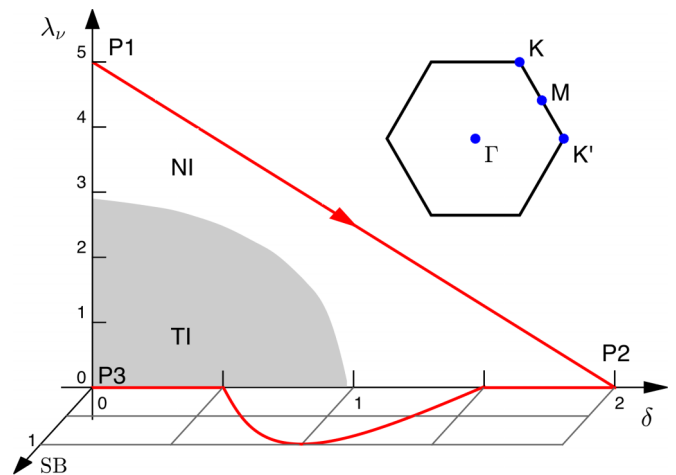


FIG. 3. Phase diagram of the KM model for $\lambda_{SO} = 0.6$ and $\lambda_R = 0.5$. The red arrows indicate the path used to construct an obstruction-free gapped path between the trivial and topological phases. The symmetry breaking is indicated in the third, out-of-plane, axis. The BZ of a honeycomb lattice is shown in the inset.

same \mathbb{Z}_2 invariant, and the line $\overline{P2P3}$, where the \mathbb{Z}_2 invariant changes from even to odd, and along which the TR symmetry is broken. Note that it is also possible to start the path directly at P2, which corresponds to the NI phase. However, for reasons made clear below, we find it illustrative to add $\overline{P1P2}$ to the discussion.

C. Wannier functions of the KM model

Now that the gapped path connecting points P1 and P3 is found, the above outlined method is implemented and the WFs are calculated along this path. The position space density [62] of the WFs is shown in Fig. 4 corresponding to the points P1, P2, and P3 on the adiabatic path of Fig. 3.

At P1 (NI phase) both WFs are localized on the lower energy sites. Each of them is mapped onto the other by TR, forming a Kramers pair of WFs with opposite spins.

The sum of the Wannier centers $\mathbf{P} = \mathbf{\bar{r}}_1 + \mathbf{\bar{r}}_2$, being proportional to the electronic polarization [27], is illustrated with a red dot in the left panel in Fig. 4. The C_3 symmetry of the KM model (see Appendix A) constrains the possible values of this sum, to the values that are invariant under this symmetry modulo a lattice vector. Consistent with C_3 are the following values: $(P_x, P_y) = (0,0), (1/3, 1/3), (2/3, 2/3)$. A change between distinct values cannot occur continuously, unless the C_3 symmetry is broken. The value of \mathbf{P} corresponding to the point P1 is $P_x = 1/3$ and $P_y = 1/3$. This is yet another example of additional topological protection. If there is a symmetry that

quantizes the values of electronic polarization [27], and if the values in the topologically trivial and nontrivial phases are different, then the symmetry must be broken so that the polarization can be changed continuously between the two discrete values [63].

When going from P1 to P2, $\delta \neq 0$ and $\lambda_v \neq 0$ and both the C_3 and C_2 symmetries are broken. This allows us to continuously change the electronic polarization without closing the band gap. At P2 (NI phase) the two lattice sites have equal energies, since $\lambda_v = 0$. Accordingly, the two WFs are equally distributed between both sites, again forming a Kramers pair. Note that $P_x = P_y = 0$ at this point.

At P3 we are in the desired \mathbb{Z}_2 -odd phase. The field of Eq. (14) is introduced to get from P2 to P3 along the gapped path. The two WFs are localized on different sites, having opposite spins, in accord with the earlier study of Ref. [43]. Clearly this configuration does not conserve the TR symmetry of the system. The sum of the Wannier centers remains at $P_x = P_y = 0$.

V. APPLICATION TO Bi_2Se_3

The standard example [6,64] of a TI is Bi_2Se_3 . It has a rhombohedral lattice structure with the space group D_{3d}^5 . The material is layered with hexagonal quintuple layers, consisting of three Se (two equivalent Se1 and one Se2) and two equivalent Bi atoms (shown as **A**, **B**, and **C** in Fig. 6) bounded together by van der Waals interaction. The structure is inversion symmetric, with an inversion center at the central Se0 atom marked with the cross in Fig. 6.

Without SOC the band structure of Bi_2Se_3 is topologically trivial, turning SOC on with a parameter λ_{SO} drives it into the TI phase. At $\lambda_{\text{SO}} = 0$ Bi_2Se_3 is a direct small band-gap semiconductor (NI phase), whereas at the full experimental SOC strength $\lambda_{\text{SO}} = 1$ it is a TI. At the topological phase transition there is an intermediate semimetallic state with a gap closure. This gap closure can be avoided by applying a suitable symmetry breaking field.

A. Constructing a model Hamiltonian for Bi_2Se_3

A fully self-consistent density functional theory calculation was carried out for Bi_2Se_3 without spin-orbit coupling (SOC) using the Vienna *ab initio* simulation package (VASP) [65,66] with the projector augmented-wave method [67], using generalized gradient approximation of the Perdew-Burke-Ernzerhof for the exchange-correlation potential [68]. The Wannier90 [29] package was then used to first disentangle an isolated group of bands and then to project the band structure onto the atomic *p* orbitals of all Bi and Se atoms. No further iterative minimization of the WF spread was done.

The resultant band structures, obtained from the projected Hamiltonian and the *ab initio* calculations are illustrated in the left panel of Fig. 5. SOC is added afterwards to this Hamiltonian by adding a local Hamiltonian H^{SOC} in the basis of the atomic *p* orbitals [69,70]. The case of full SOC is illustrated in the right panel of Fig. 5. The correct topological phase of the projected Hamiltonian was confirmed using the method of Ref. [38].

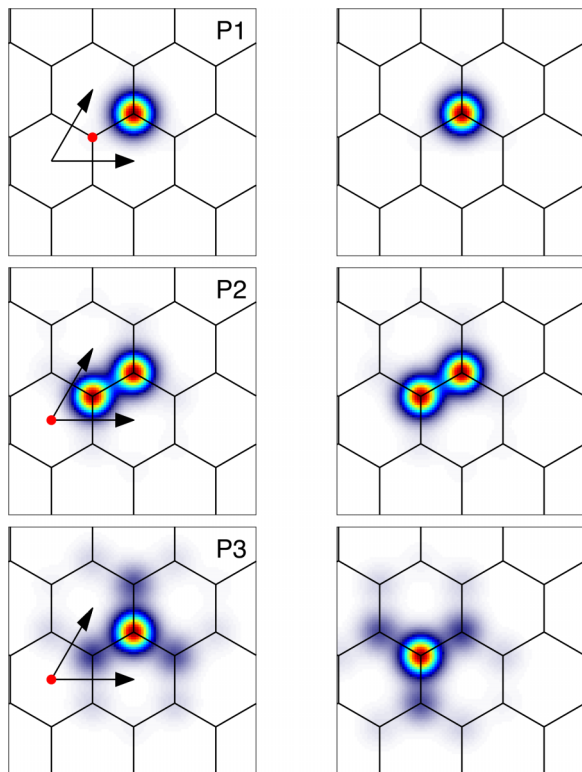


FIG. 4. Charge density distribution corresponding to the WFs obtained at points P1, P2, and P3 of Fig. 3. The left and right panels show the two WFs of the KM model at each of the points. The red point marks the sum of the corresponding Wannier centers (modulo a lattice vector).

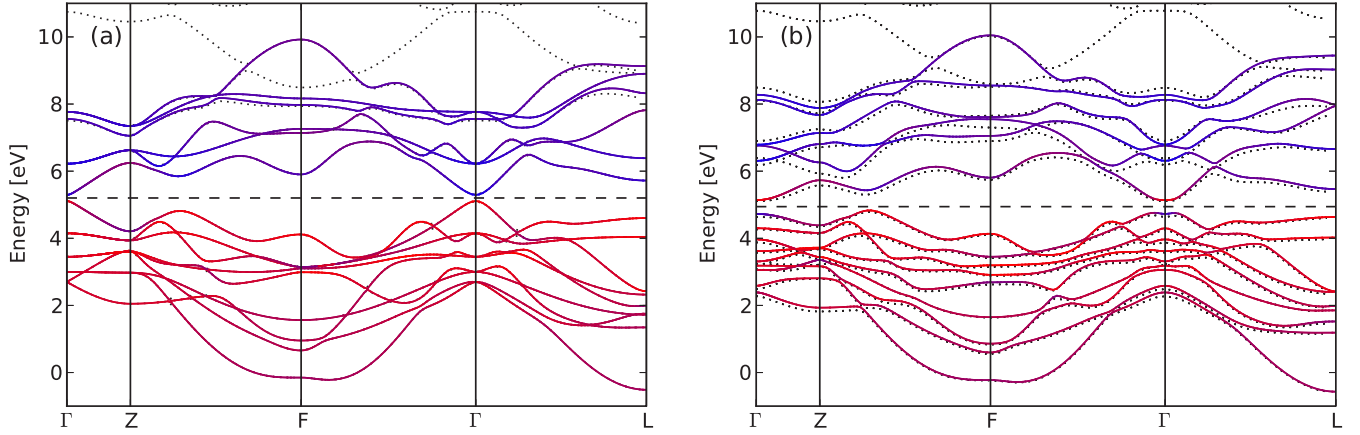


FIG. 5. Projected band structure [Se (Bi) character shown in red (blue)] and the *ab initio* band structures (black dotted lines) of Bi_2Se_3 . Left panel: no spin-orbit coupling. Right panel: spin-orbit coupling included.

In the following we work with the projected Hamiltonian, which allows for the easy control of SOC and a symmetry breaking field in the total Hamiltonian

$$H^{\text{tot}} = H^0 + \lambda_{\text{SO}} H^{\text{SOC}} + \alpha H^{\text{SB}}. \quad (15)$$

The term H^{SB} is introduced below. Note that both H^{SOC} and H^{SB} are local. The two parameters λ_{SO} and α are found to be sufficient to construct an adiabatic path between the NI and TI phase in Bi_2Se_3 .

B. Constructing an obstruction-free path for Bi_2Se_3

The SOC strength λ_{SO} is tuned from $\lambda_{\text{SO}} = 0$ (NI phase) to $\lambda_{\text{SO}} = 1$ (TI phase with full experimental SOC strength) [71], passing through a topological phase transition at $\lambda_{\text{SO}} \approx 0.47$. At this phase transition the gap closes and a 3D Dirac cone is formed at the Γ point. To obtain a gapped adiabatic path connecting NI and TI phases, a suitable symmetry breaking field H^{SB} that gaps the Dirac cone needs to be found by considering the symmetry breaking requirements. Due to the double protection described in Sec. III, the field must break both TR and inversion symmetries. This requirement is satisfied by a staggered magnetic field, analogous to that of Eq. (14) used in the case of the KM model. The optimal direction of the staggered field was found to be in the plane of hexagonal layers (see Fig. 6).

Examples of two such staggered fields are illustrated in Fig. 6. The first one, $H_{\text{Bi}}^{\text{SB}}$, shown with black arrows, acts on

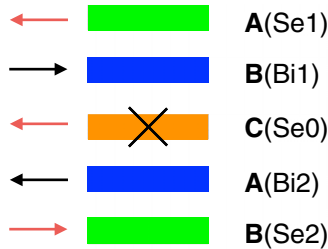


FIG. 6. The layered structure of Bi_2Se_3 . The cross denotes the inversion center at the Se0 atoms. The black arrows indicate the staggered magnetic field on the Bi sites corresponding to $H_{\text{Bi}}^{\text{SB}}$. The red arrows indicate an additional field corresponding to $H_{\text{BiSe}}^{\text{SB}}$.

Bi sites only. While both TR and inversion symmetries are broken by this field, the combination of the two symmetries survives. As a consequence, the band structure remains doubly degenerate upon the inclusion of this field.

The other symmetry breaking field, $H_{\text{BiSe}}^{\text{SB}}$, breaks this compound symmetry as well by applying, in addition to $H_{\text{Bi}}^{\text{SB}}$, a field on all Se sites, as marked by the red arrows in Fig. 6. Application of this field lifts the double degeneracy of the bands. Both fields prevent the gap closure and allow for the adiabatic connection between the NI and the TI.

The band gap as a function of λ_{SO} is shown in Fig. 7. The parameter α of the TR symmetry breaking field is tuned as $\alpha(\lambda_{\text{SO}}) = \sin(\lambda_{\text{SO}}\pi)$. We interpolate the path between $\lambda_{\text{SO}} = 0$ and $\lambda_{\text{SO}} = 1$ using nine intermediate equidistant steps. In what follows the BZ of the Hamiltonian 15 is discretized into a $16 \times 16 \times 16 \mathbf{k}$ mesh.

C. Wannier functions of Bi_2Se_3

We first need to find WFs for the initial step of the path, that is for $\lambda_{\text{SO}} = 0$. Suitable trial states $|\tau_i\rangle$ for the topologically trivial band structure can be guessed from the occupations of

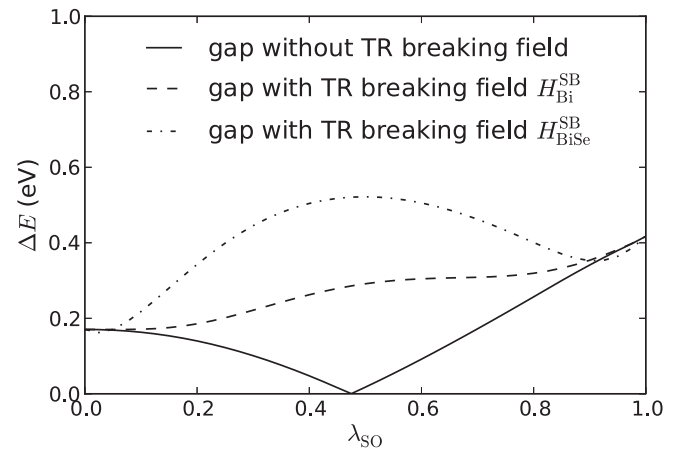


FIG. 7. The bulk energy gap as a function of λ_{SO} . The symmetry breaking is tuned as $\sin(\lambda_{\text{SO}}\pi)$. The field $H_{\text{Bi}}^{\text{SB}}$ preserves the double degeneracy of bands. The gap is larger for the $H_{\text{BiSe}}^{\text{SB}}$ field, which lifts this degeneracy.

TABLE I. The spread of the resultant WFs after minimization.

	$\lambda_{SO} = 0$		$\lambda_{SO} = 1$	
	Ω_I	$\tilde{\Omega}$	Ω_I	$\tilde{\Omega}$
H_{Bi}^{SB}	11.01	0.46	11.04	0.70
H_{BiSe}^{SB}	11.01	0.46	11.04	0.73

atoms shown with color in Fig. 5(a). The occupied subspace consists mainly of the Se p states. Therefore, the Se p orbitals (both up and down spin) are chosen to be the initial trial states.

Throughout the path we monitor the minimum of $\det S(\mathbf{k})$ and the Wannier spreads. Apart from the first projection, where we start with local projections, the determinant always stays reasonably close to 1. For this reason, the minimization of the Wannier spreads done at each step of the path, after projecting onto the WFs obtained at the previous step, results in only small improvement in localization.

The Wannier spreads in the NI and TI phase are given in Table I. As expected we find them to be larger in the TI than in the NI phase [43]. Both symmetry breaking fields lead to nearly identical WFs, although, the resulting spread is found to be slightly larger when the field H_{BiSe}^{SB} is used to break the symmetry. The smoothness of the resulting gauge is visible in Fig. 8 (see Fig. 1), where we show the flow of the hybrid Wannier centers in the $k_x = 0$ plane. All centers are smooth and periodic in the BZ.

Now that a smooth gauge for Bi_2Se_3 is obtained, the corresponding Bloch states can be used as a set of \mathbf{k} -dependent trial states $\tau_i(\mathbf{k})$. These states are useful to quickly obtain a smooth gauge and ELWFs on a different k mesh without repeating the above procedure. In the spirit of Wannier interpolation [25], the new $\tau_i(\mathbf{k})$ are obtained by Fourier transforming the ELWFs on the new \mathbf{k} mesh.

The obtained ELWFs also allowed us to find the local trial states that work for Bi_2Se_3 that result in well localized WFs for different \mathbf{k} meshes [72]. They are listed in Table II, where the site location (see Fig. 6), the orbital character, and the spin direction are provided for each of the trial states. Apart from the two exclusions, the states come in time-reversal pairs listed in the same row of the table. The trial states have mostly Se character, apart from the exclusions, which are listed in rows 4 and 7 of the table, and reflect the band inversion that occurs

TABLE II. The 18 trial states for the topological material Bi_2Se_3 . Left and right column show potential time reversed partners.

$ Se0, p_x, \uparrow_z\rangle$	$ Se0, p_x, \downarrow_z\rangle$
$ Se0, p_y, \uparrow_z\rangle$	$ Se0, p_y, \downarrow_z\rangle$
$ Se0, p_z, \uparrow_z\rangle$	$ Se0, p_z, \downarrow_z\rangle$
$\frac{1}{\sqrt{2}}(Se1, p_x, \uparrow_x\rangle + Bi1, p_x, \uparrow_x\rangle)$	$ Se1, p_x, \downarrow_x\rangle$
$ Se1, p_y, \uparrow_x\rangle$	$ Se1, p_y, \downarrow_x\rangle$
$ Se1, p_z, \uparrow_x\rangle$	$ Se1, p_z, \downarrow_x\rangle$
$ Se2, p_x, \uparrow_x\rangle$	$\frac{1}{\sqrt{2}}(Se2, p_x, \downarrow_x\rangle + Bi2, p_x, \downarrow_x\rangle)$
$ Se2, p_y, \uparrow_x\rangle$	$ Se2, p_y, \downarrow_x\rangle$
$ Se2, p_z, \uparrow_x\rangle$	$ Se2, p_z, \downarrow_x\rangle$

between the p_x states of Se and Bi in Bi_2Se_3 at the Γ point, which can be seen in Fig. 5(b).

As discussed above, Bi_2Se_3 is an example of a double symmetry protection of the band topology, where both inversion and TR are broken in the smooth gauge. This breaking of the two symmetries is clearly seen in the flow of hybrid Wannier centers shown in Fig. 8. Figure 8(a) shows hybrid Wannier centers constructed from a symmetry-preserving gauge. The topological obstruction is evident. In contrast, Fig. 8(b) illustrates the case of the smooth gauge obtained above. The smooth gauge Wannier centers are neither inversion, nor TR symmetric, but are smooth and periodic in the BZ, corresponding to zero individual Chern numbers.

D. Evaluation of the Chern-Simons magnetoelectric polarizability θ_{CS}

We now proceed to calculating the geometrical contribution to the magnetoelectric effect. The numerical evaluation of the Chern-Simons magnetoelectric polarizability θ_{CS} introduced in Sec. I is a tedious task [30]. While a direct simulation of a material's response to electromagnetic fields can potentially be used to evaluate this term, such a calculation requires the use of large supercells, which makes it inefficient and computationally expensive. Several methods to compute θ_{CS} from the bulk wave functions exist [30,33,34,73,74], but only one of them (Ref. [30]) was applied to materials within an *ab initio* framework. Our calculation is based on the formalism developed in the latter work, but the actual implementation of the formulas is also done directly in position space, which significantly improves convergence with respect to the \mathbf{k} mesh, compared to the \mathbf{k} -space implementations used before [30].

To calculate θ_{CS} the Bloch wave functions obtained from the bulk calculation are used. The Berry connection matrix for these states is defined as

$$\mathcal{A}_{mn,j}(\mathbf{k}) = \langle u_{m\mathbf{k}} | i \frac{\partial}{\partial k_j} | u_{n\mathbf{k}} \rangle, \quad (16)$$

which is computed using the smooth gauge, obtained above [75]. Then θ_{CS} is evaluated by integrating the Chern-Simons 3-form over the entire Brillouin zone [3,33]

$$\theta_{CS} = -\frac{1}{4\pi} \int_{BZ} d^3k \epsilon_{ijk} \text{tr} \left[\mathcal{A}_i \partial_j \mathcal{A}_k - \frac{2i}{3} \mathcal{A}_i \mathcal{A}_j \mathcal{A}_k \right], \quad (17)$$

where the summation over band and Cartesian indices is assumed. While this integral is gauge invariant modulo 2π , a

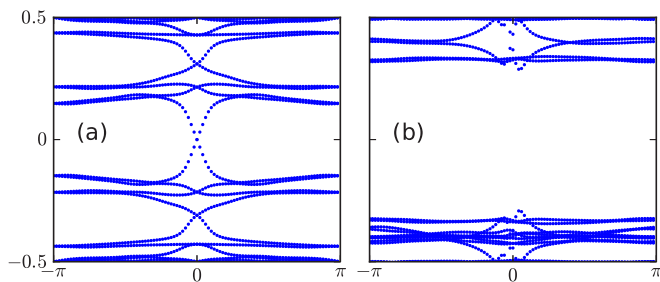


FIG. 8. Flow of the hybrid Wannier centers in the $k_x = 0$ plane in Bi_2Se_3 . (a) The gauge respects TR and inversion symmetries. (b) The smooth gauge used to obtain the ELWFs.

smooth gauge is assumed in the derivation of Eq. (17) [3,34]. Thus a smooth gauge is required for a meaningful evaluation of the integral. This is when the above technique for finding a smooth gauge becomes important.

Equation (17) can be rewritten in terms of matrix elements of position operators evaluated with WFs in position space. Using

$$A_{mn,j}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \langle \mathbf{0}m | r_j | \mathbf{R}n \rangle, \quad (18)$$

the result reported previously [30] can be obtained

$$\theta_{\text{CS}} = \frac{1}{4\pi} \frac{2\pi^3}{\Omega} \epsilon_{ijk} \text{Im} \left(\sum_{\mathbf{R}} \langle \mathbf{0}m | r_i | \mathbf{R}n \rangle \langle \mathbf{R}n | r_j | \mathbf{0}m \rangle R_k - \frac{2}{3} \sum_{\mathbf{R}\mathbf{P}} \langle \mathbf{0}l | r_i | \mathbf{R}m \rangle \langle \mathbf{R}m | r_j | \mathbf{P}n \rangle \langle \mathbf{P}n | r_k | \mathbf{0}l \rangle \right). \quad (19)$$

This equation still assumes a smooth gauge, since when the WFs are not exponentially localized, surface terms should be included to account for the slow decay of the Wannier matrix elements [30].

We now use the ELWFs obtained for Bi_2Se_3 above, to compute the θ_{CS} term in this material. While both formulas Eqs. (17) and (19) agree in the limit of an infinitely dense \mathbf{k} mesh, the use of the position space formula of Eq. (19) results in faster convergence, compared to the \mathbf{k} -space formulation of Eq. (17) [76].

We used the trial states of Table II to get localized WFs for a variety of \mathbf{k} meshes up to $80 \times 80 \times 80$. The scaling in the TI phase is shown in Fig. 9. Because of the slow convergence in \mathbf{k} -mesh density, an extrapolation to the infinitely dense mesh is required. A linear extrapolation in $(\Delta k)^2$ is done using the

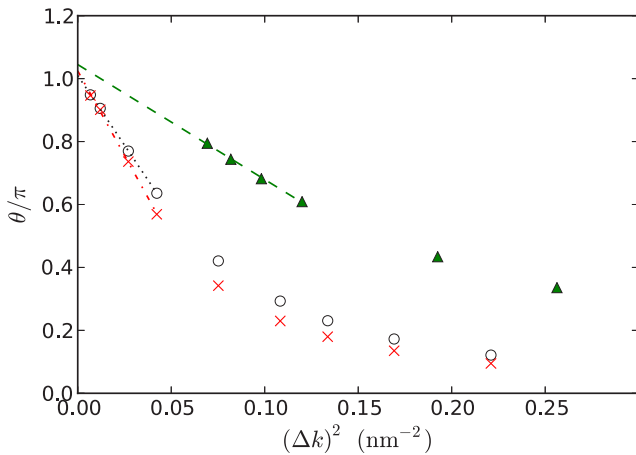


FIG. 9. Convergence of θ_{CS} in the TI for varying densities of \mathbf{k} meshes. Δk is the nearest-neighbor spacing on the grid. Convergence to $\theta_{\text{CS}} = \pi$ can be reached for mesh densities of order $80 \times 80 \times 80$ [$(\Delta k)^2 = 0.0068 \text{ nm}^{-2}$]. The black circles (red crosses) correspond to the results obtained by evaluating the Berry connection in \mathbf{k} space without (with) maximal localization of the WFs resultant from the projections of Table II. The green triangles correspond to the results obtained by evaluating $\langle \mathbf{0}m | r_j | \mathbf{R}n \rangle$ in position space. A linear extrapolation to the infinitely dense \mathbf{k} mesh was done using the last data points.

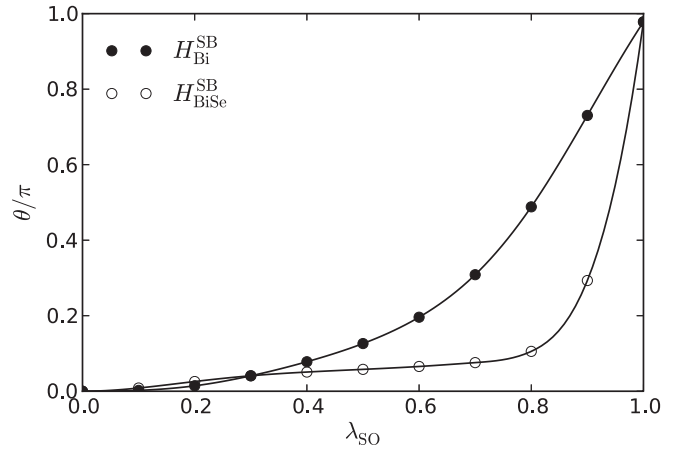


FIG. 10. θ_{CS} calculated along the adiabatic path connecting NI and TI phases. The full (empty) circles correspond to the symmetry breaking field $H_{\text{Bi}}^{\text{SB}}$ ($H_{\text{BiSe}}^{\text{SB}}$) [60].

last data points, getting in all cases very close to the expected value of $\theta_{\text{CS}} = \pi$. This extrapolation choice is dictated by the error in numerical evaluation of the \mathcal{A}_j matrices and the Wannier matrix elements $\langle \mathbf{0}m | r_j | \mathbf{R}n \rangle$, which is of order $O[(\Delta k)^2]$ [77].

Exponential convergence can be achieved by directly evaluating $\langle \mathbf{0}m | r_j | \mathbf{R}n \rangle$ in position space with the ELWFs [78]. This is, however, computationally more expensive and was done only for relatively coarse \mathbf{k} meshes up to $25 \times 25 \times 25$. The green triangles in Fig. 9 illustrate the results obtained from this position space approach. Even with this coarse \mathbf{k} mesh a reliable extrapolation to the expected value $\theta_{\text{CS}} = \pi$ can be done within a few percent accuracy.

The θ_{CS} term was also computed along the adiabatic path connecting NI to TI for the two different symmetry breaking fields ($H_{\text{Bi}}^{\text{SB}}$ and $H_{\text{BiSe}}^{\text{SB}}$). The results are shown in Fig. 10. The θ_{CS} term is found to be larger for the $H_{\text{Bi}}^{\text{SB}}$ field, which does not break the product symmetry of TR and inversion.

VI. CONCLUSION

In this paper we established a general procedure for constructing ELWFs and smooth Bloch states to describe a group of bands with nontrivial topology. This is done by connecting the topologically nontrivial Hamiltonian to some trivial one by a gapped adiabatic path. Our technique works for all symmetry-protected topologies, provided that the net Chern number of the bands is zero. It was illustrated for \mathbb{Z}_2 TIs in two dimensions using the example of the Kane-Mele tight-binding model, and the real three-dimensional topological insulator material Bi_2Se_3 . We also introduced the concept of double symmetry protection of nontrivial topology, which describes (ubiquitous) situations when there is more than one symmetry that presents the obstruction for choosing smooth Bloch states. We expect that this technique can also be generalized to the case of Chern insulators to obtain numerically a smooth lattice representation for all the occupied Bloch states, apart from one that carries the net Chern number of the system, and hence cannot be made smooth.

Finally, we described how the proposed scheme of constructing a smooth gauge allows for the calculation of the Chern-Simons θ_{CS} term that captures the geometric contribution to the orbital magnetoelectric response of materials. A detailed discussion of the numerical implementation of this calculation was provided, showing how to efficiently implement our technique to materials, where θ_{CS} is not quantized. This technique is especially useful in the presence of band topologies that do not result in a quantized value of the θ_{CS} term. More generally, the numerical construction of the smooth gauge for lattice models presented here can have broader applications in evaluation of various (band) geometric effects that do not have an immediate gauge-invariant formulation.

ACKNOWLEDGMENTS

We would like to thank David Vanderbilt and Xiao-Liang Qi for useful discussions. This work was supported by Microsoft Research, the European Research Council through ERC Advanced Grant SIMCOFE, and the Swiss National Science Foundation through the National Competence Centers in Research MARVEL and QSIT.

APPENDIX: SYMMETRIES OF THE KANE-MELE MODEL

Here the symmetries of the Kane-Mele model discussed in the main text are derived explicitly. Using the basis set ($|A \uparrow\rangle, |B \uparrow\rangle, |A \downarrow\rangle, |B \downarrow\rangle$) the Hamiltonian Eq. (13) can be written in \mathbf{k} space

$$H(\mathbf{k}) = t \sigma_0 [\tau_x \epsilon_{t1}(\mathbf{k}) + \tau_y \epsilon_{t2}(\mathbf{k})] + \lambda_{\text{SO}} \sigma_z \tau_z \epsilon_{\text{SO}}(\mathbf{k}) + \lambda_R \{ \sigma_x [\tau_x \epsilon_{R1}(\mathbf{k}) + \tau_y \epsilon_{R2}(\mathbf{k})] + \sigma_y [\tau_y \epsilon_{R3}(\mathbf{k}) + \tau_x \epsilon_{R4}(\mathbf{k})] \} - \lambda_v \sigma_0 \tau_z, \quad (\text{A1})$$

with σ and τ matrices acting in spin and lattice subspaces; σ_0 being the unit matrix, and $\sigma_{x,y,z}$ ($\tau_{x,y,z}$) the Pauli matrices.

The \mathbf{k} -dependent coefficients in the Hamiltonian are

$$\begin{aligned} \epsilon_{t1}(\mathbf{k}) &= (1 + \delta) f_1(\mathbf{k}) + f_2(\mathbf{k}) + f_3(\mathbf{k}), \\ \epsilon_{t2}(\mathbf{k}) &= (1 + \delta) g_1(\mathbf{k}) + g_2(\mathbf{k}) + g_3(\mathbf{k}), \\ \epsilon_{\text{SO}}(\mathbf{k}) &= 2[-\sin(k_1) + \sin(k_2) + \sin(k_1 - k_2)], \\ \epsilon_{R1} &= -\frac{1}{2} g_1(\mathbf{k}) - \frac{1}{2} g_2(\mathbf{k}) + g_3(\mathbf{k}), \\ \epsilon_{R2} &= -\frac{1}{2} f_1(\mathbf{k}) - \frac{1}{2} f_2(\mathbf{k}) + f_3(\mathbf{k}), \\ \epsilon_{R3} &= \frac{\sqrt{3}}{2} f_1(\mathbf{k}) - \frac{\sqrt{3}}{2} f_2(\mathbf{k}), \\ \epsilon_{R4} &= \frac{\sqrt{3}}{2} g_1(\mathbf{k}) - \frac{\sqrt{3}}{2} g_2(\mathbf{k}), \end{aligned} \quad (\text{A2})$$

where

$$\begin{aligned} f_1(\mathbf{k}) &= \cos\left(\frac{k_1}{3} + \frac{k_2}{3}\right), & g_1(\mathbf{k}) &= \sin\left(\frac{k_1}{3} + \frac{k_2}{3}\right), \\ f_2(\mathbf{k}) &= \cos\left(-2\frac{k_1}{3} + \frac{k_2}{3}\right), & g_2(\mathbf{k}) &= \sin\left(-2\frac{k_1}{3} + \frac{k_2}{3}\right), \\ f_3(\mathbf{k}) &= \cos\left(\frac{k_1}{3} - 2\frac{k_2}{3}\right), & g_3(\mathbf{k}) &= \sin\left(\frac{k_1}{3} - 2\frac{k_2}{3}\right), \end{aligned}$$

TABLE III. Point symmetries (relative to the origin) of the KM model for $\lambda_v = \delta = 0$. Both momentum transformations and matrix representations are given.

Symmetry	Representation \hat{R}	$k_1 \rightarrow \dots$	$k_2 \rightarrow \dots$
$\sigma_v^{(11)}$	$i\left(-\frac{\sigma_x}{2} + \frac{\sqrt{3}}{2}\sigma_y\right)\tau_0$	k_2	k_1
$\sigma_v^{(1\bar{1})}$	$i\left(\frac{\sqrt{3}}{2}\sigma_x + \frac{\sigma_y}{2}\right)\tau_x$	$-k_2$	$-k_1$
C_2	$i\sigma_z\tau_x$	$-k_1$	$-k_2$
C_3	$\left(\frac{\sigma_0}{2} + i\frac{\sqrt{3}}{2}\sigma_z\right)\tau_0$	$k_2 - k_1$	$-k_1$
C_6	$\left(\frac{\sqrt{3}}{2}\sigma_0 + i\frac{\sigma_z}{2}\right)\tau_x$	k_2	$k_2 - k_1$

with k_1 and k_2 being the reduced coordinates in terms of the reciprocal lattice vectors $\mathbf{b}_1 = \frac{2\pi}{a}(\hat{\mathbf{k}}_x - \frac{1}{\sqrt{3}}\hat{\mathbf{k}}_y)$ and $\mathbf{b}_2 = \frac{2\pi}{a}\frac{2}{\sqrt{3}}\hat{\mathbf{k}}_y$.

We first consider the case of $\lambda_v = \delta = 0$, corresponding to the point P3 in Fig. 3. In this case the 2D space group (wallpaper group) of the model is $p6m$ in IUC notation (*632 in orbifold notation) [79], and the corresponding 2D point group is D_6 . The specific matrix representation of all the point group symmetries for this case is provided in Table III. Each of the symmetries acts according to

$$H(S(k_1, k_2)) = \hat{R}_S H(k_1, k_2) \hat{R}_S^\dagger, \quad (\text{A3})$$

where S is the symmetry operation and \hat{R}_S is its matrix representation.

The consistency of the matrix representation of Table III can be checked by noting that the C_2 symmetry is the product of the two mirrors $\sigma_v^{(11)}$ and $\sigma_v^{(1\bar{1})}$, and the C_3 symmetry is the C_6 rotation applied twice, as expected. Other reflection symmetries can be obtained by appropriate combinations of these operations. The relations expected for spin-1/2 systems, namely $\sigma^2 = C_2^2 = C_3^3 = C_6^6 = -1$ are satisfied. Without the Rashba term, that is for $\lambda_R = 0$, there would be an additional inversion symmetry $\sigma_0\tau_z$.

The symmetries corresponding to the different points in the path of Fig. 3 are summarized in Table IV. When both $\lambda_v \neq 0$ and $\delta \neq 0$, which is the case on the line connecting P1 and P2 ($\bar{P}1\bar{P}2$ in Fig. 3), the sum of the Wannier centers \mathbf{P} is allowed to change continuously along the lines $P_y = \frac{1}{\sqrt{3}}P_x + \frac{m}{2}$, where m is an integer.

The TR symmetry in the model is given by $i\sigma_y\tau_0\mathcal{K}$, with \mathcal{K} being complex conjugation, and its action in \mathbf{k} space given by $k_1 \rightarrow -k_1$ and $k_2 \rightarrow -k_2$. Two TR-breaking fields $H_{(1)}^{\text{SB}}$ and $H_{(2)}^{\text{SB}}$ were considered in this work:

$$\begin{aligned} H_{(1)}^{\text{SB}} &= \left(-\frac{1}{2}\sigma_x + \frac{\sqrt{3}}{2}\sigma_y\right)\tau_z, \\ H_{(2)}^{\text{SB}} &= \left(-\frac{\sqrt{3}}{4}\sigma_x + \frac{3}{4}\sigma_y + \frac{1}{2}\sigma_z\right)\tau_z. \end{aligned}$$

Of them, $H_{(1)}^{\text{SB}}$ commutes with both mirror symmetries $\sigma_v^{(11)}$, $\sigma_v^{(1\bar{1})}$. The field is applied along the line $\bar{P}_2\bar{P}_3$, where, according to Table IV, both these mirrors are preserved. Therefore, the

TABLE IV. Symmetries at the different points of the adiabatic path of Fig. 3. The wallpaper group is given in IUC and orbifold notation.

	$\lambda_v = \delta = 0$	$\lambda_v \neq 0, \delta = 0$	$\lambda_v = 0, \delta \neq 0$	$\lambda_v \neq 0, \delta \neq 0$
See Fig. 3	P3	P1	P2 P3	P1 P2
Wallpaper group	$p6m, *632$	$p3m1, *333$	$cmm, 2*22$	$cm, *x$
Point group	D_6	D_3	D_2	D_1
Conserved symmetries	$\sigma_v^{(11)}, \sigma_v^{(1\bar{1})}, C_2, C_3, C_6$	$\sigma_v^{(11)}, C_3$	$\sigma_v^{(11)}, \sigma_v^{(1\bar{1})}, C_2$	$\sigma_v^{(11)}$
P quantized	Yes	Yes	Yes	No

resulting WFs reflect this symmetry and the minimization of their spread gets stuck in a local minimum. By adding the σ_z

component, leading to $H_{(2)}^{\text{SB}}$, both mirrors are broken and we obtain the maximally localized WFs also found in Ref. [43].

-
- [1] M. Z. Hasan and C. L. Kane, *Rev. Mod. Phys.* **82**, 3045 (2010).
[2] X.-L. Qi and S.-C. Zhang, *Rev. Mod. Phys.* **83**, 1057 (2011).
[3] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, *Phys. Rev. B* **78**, 195424 (2008).
[4] L. Fu and C. L. Kane, *Phys. Rev. Lett.* **100**, 096407 (2008).
[5] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, *Science* **318**, 766 (2007).
[6] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Nat. Phys.* **5**, 398 (2009).
[7] D. Hsieh, Y. Xia, D. Qian, L. Wray, F. Meier, J. H. Dil, J. Osterwalder, L. Patthey, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Phys. Rev. Lett.* **103**, 146401 (2009).
[8] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *Nature (London)* **452**, 970 (2008).
[9] Y. Tanaka, Z. Ren, T. Sato, K. Nakayama, S. Souma, T. Takahashi, K. Segawa, and Y. Ando, *Nat. Phys.* **8**, 800 (2012).
[10] X. Chen, Z.-X. Liu, and X.-G. Wen, *Phys. Rev. B* **84**, 235141 (2011).
[11] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, *Phys. Rev. B* **87**, 155114 (2013).
[12] F. D. M. Haldane, *Phys. Rev. Lett.* **61**, 2015 (1988).
[13] G. Volovik, *Phys. Lett. A* **128**, 277 (1988).
[14] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Phys. Rev. Lett.* **49**, 405 (1982).
[15] X.-L. Qi, Y.-S. Wu, and S.-C. Zhang, *Phys. Rev. B* **74**, 045125 (2006).
[16] C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
[17] L. Fu, *Phys. Rev. Lett.* **106**, 106802 (2011).
[18] T. L. Hughes, E. Prodan, and B. A. Bernevig, *Phys. Rev. B* **83**, 245132 (2011).
[19] T. H. Hsieh, H. Lin, J. Liu, W. Duan, A. Bansil, and L. Fu, *Nat. Commun.* **3**, 982 (2012).
[20] Y. Ando and L. Fu, *Annu. Rev. Condens. Matter Phys.* **6**, 361 (2015).
[21] M. V. Berry, *Proc. R. Soc. London Ser. A* **392**, 45 (1984).
[22] J. Zak, *Phys. Rev. Lett.* **62**, 2747 (1989).
[23] G. H. Wannier, *Phys. Rev.* **52**, 191 (1937).
[24] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, *Rev. Mod. Phys.* **84**, 1419 (2012).
[25] I. Souza, N. Marzari, and D. Vanderbilt, *Phys. Rev. B* **65**, 035109 (2001).
[26] N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847 (1997).
[27] R. D. King-Smith and D. Vanderbilt, *Phys. Rev. B* **47**, 1651 (1993).
[28] Y.-S. Lee, M. Buongiorno Nardelli, and N. Marzari, *Phys. Rev. Lett.* **95**, 076804 (2005).
[29] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, *Comput. Phys. Commun.* **185**, 2309 (2014).
[30] S. Coh, D. Vanderbilt, A. Malashevich, and I. Souza, *Phys. Rev. B* **83**, 085108 (2011).
[31] J. Íñiguez, *Phys. Rev. Lett.* **101**, 117201 (2008).
[32] E. Bousquet, N. A. Spaldin, and K. T. Delaney, *Phys. Rev. Lett.* **106**, 107202 (2011).
[33] A. M. Essin, J. E. Moore, and D. Vanderbilt, *Phys. Rev. Lett.* **102**, 146805 (2009).
[34] A. M. Essin, A. M. Turner, J. E. Moore, and D. Vanderbilt, *Phys. Rev. B* **81**, 205104 (2010).
[35] F. Wilczek, *Phys. Rev. Lett.* **58**, 1799 (1987).
[36] A. M. Turner, Y. Zhang, R. S. K. Mong, and A. Vishwanath, *Phys. Rev. B* **85**, 165120 (2012).
[37] L. Fu and C. L. Kane, *Phys. Rev. B* **74**, 195312 (2006).
[38] A. A. Soluyanov and D. Vanderbilt, *Phys. Rev. B* **83**, 235401 (2011).
[39] R. Yu, X. L. Qi, A. Bernevig, Z. Fang, and X. Dai, *Phys. Rev. B* **84**, 075119 (2011).
[40] D. J. Thouless, *J. Phys. C* **17**, L325 (1984).
[41] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, *Phys. Rev. Lett.* **98**, 046402 (2007).
[42] T. A. Loring and M. B. Hastings, *Europhys. Lett.* **92**, 67004 (2010).
[43] A. A. Soluyanov and D. Vanderbilt, *Phys. Rev. B* **83**, 035108 (2011).
[44] R. Roy, *Phys. Rev. B* **79**, 195321 (2009).
[45] A. A. Soluyanov and D. Vanderbilt, *Phys. Rev. B* **85**, 115415 (2012).
[46] A. Alexandradinata, X. Dai, and B. A. Bernevig, *Phys. Rev. B* **89**, 155114 (2014).
[47] J. C. Y. Teo, L. Fu, and C. L. Kane, *Phys. Rev. B* **78**, 045426 (2008).
[48] C.-K. Chiu, H. Yao, and S. Ryu, *Phys. Rev. B* **88**, 075142 (2013).
[49] C. Fang, M. J. Gilbert, and B. A. Bernevig, *Phys. Rev. B* **86**, 115112 (2012).

- [50] C. Sgierovello, M. Peressi, and R. Resta, *Phys. Rev. B* **64**, 115202 (2001).
- [51] D. Gresch, A. A. Soluyanov, G. Autés, D. Vanderbilt, B. A. Bernevig, O. Yazyev, and M. Troyer (unpublished).
- [52] A discussion why parallel transport gives a valid gauge can be found in Sec. II B of Ref. [38].
- [53] S. Kivelson, *Phys. Rev. B* **26**, 4269 (1982).
- [54] How this split may be actually facilitated on a lattice is described in Ref. [45].
- [55] A. Alexandradinata, C. Fang, M. J. Gilbert, and B. A. Bernevig, *Phys. Rev. Lett.* **113**, 116403 (2014).
- [56] J. I. Mustafa, S. Coh, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **92**, 165134 (2015).
- [57] The procedure is facilitated by choosing a path, along which the gap remains large on the scale of the topological gap. The larger the gap, the coarser can be the discretization of the path.
- [58] We also find it beneficial to run maximal localization of the WFs at each step of the procedure, prior to proceeding to the next one.
- [59] L. Fu and C. L. Kane, *Phys. Rev. B* **76**, 045302 (2007).
- [60] This argument is easily generalized to weak and strong 3D TIs, taking into account that these phases have certain ordering of trivial and nontrivial 2D TR-symmetric planes in the BZ [80].
- [61] We found that the field of the form $H_{(1)}^{SB} = (-\frac{1}{2}\sigma_x + \frac{\sqrt{3}}{2}\sigma_y)\tau_z$ works as well, but it breaks less symmetries than $H_{(2)}^{SB}$, and the resultant WFs are less localized compared to those obtained with $H_{(2)}^{SB}$, as explained in the Appendix A.
- [62] Gaussian-shaped orbitals were used for the graphical representation.
- [63] P. Jadaun, D. Xiao, Q. Niu, and S. K. Banerjee, *Phys. Rev. B* **88**, 085110 (2013).
- [64] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, *Nat. Phys.* **5**, 438 (2009).
- [65] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [66] G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- [67] We used 340 eV for the cut-off energy and a $10 \times 10 \times 10$ \mathbf{k} -point mesh in the BZ. The lattice constants were taken from Ref. [81].
- [68] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [69] W. Zhang, R. Yu, H.-J. Zhang, X. Dai, and Z. Fang, *New J. Phys.* **12**, 065013 (2010).
- [70] D. J. Chadi, *Phys. Rev. B* **16**, 790 (1977).
- [71] K. Wittel and R. Manne, *Theor. Chim. Acta* **33**, 347 (1974).
- [72] A naive choice for the $|\tau_i\rangle$ to be just the ELWFs truncated outside the home unit cell is equivalent to averaging the smooth Bloch states over the BZ. This can cause a problem during projection, since the volume of the region in the BZ, where the band ordering is inverted, is often much smaller than that of the region with normal ordering. Therefore, this choice may under represent the band character corresponding to the inverted band ordering in the trial states, so $\det S(\mathbf{k})$ will become small when projecting onto them. We found that this problem can easily be solved by manually adding a contribution of the dominating character $|\rho\rangle$ of the smooth Bloch states at the region of band inversion to the trial states, so that the actual trial states become $|\bar{\tau}_i\rangle = 1/\sqrt{2}(|\tau_i\rangle + |\rho\rangle)$. Table II lists the trial states obtained with this procedure.
- [73] M. Taherinejad and D. Vanderbilt, *Phys. Rev. Lett.* **114**, 096401 (2015).
- [74] J. Liu and D. Vanderbilt, *Phys. Rev. B* **92**, 245138 (2015).
- [75] This is done using the smooth Bloch functions obtained for the projected Hamiltonian as $\mathcal{A}_{mn,j}(\mathbf{k}) \simeq i \sum_{\mathbf{b}} w_{\mathbf{b}} b_j (\langle u_{m\mathbf{k}} | u_{n\mathbf{k}+\mathbf{b}} \rangle - \delta_{mn})$, where \mathbf{b} are the vectors connecting \mathbf{k} to its nearest neighbors on the \mathbf{k} mesh. The overlap matrices $\langle u_{m\mathbf{k}} | u_{n\mathbf{k}+\mathbf{b}} \rangle$ are calculated assuming the basis WF, used to construct the projected Hamiltonian, to be the eigenstates of the position operator. Certain (small) corrections that account for this approximation, and arise only when the TR symmetry is broken, are omitted here. Their calculation, possible in principle [77], goes beyond the scope of the present paper.
- [76] The first term of Eq. (17) contains \mathbf{k} -derivatives $\partial_j \mathcal{A}_k$ that cause slow convergence. Using Eq. (18) the derivatives can be treated analytically as it is done in Eq. (19). The convergence of the first term with respect to the \mathbf{k} mesh is improved. The convergence of the second term in Eqs. (17)–(19) is found to be identical in both real and momentum space formulations. However, \mathbf{k} -space formulation of Eq. (17) was used for the evaluation of this term being more efficient numerically.
- [77] X. Wang, J. R. Yates, I. Souza, and D. Vanderbilt, *Phys. Rev. B* **74**, 195118 (2006).
- [78] M. Stengel and N. A. Spaldin, *Phys. Rev. B* **73**, 075121 (2006).
- [79] J. Conway, in *Groups, Combinatorics & Geometry*, edited by M. W. Liebeck and J. Saxl (Cambridge University Press, Cambridge, 1992), pp. 438–447.
- [80] L. Fu, C. L. Kane, and E. J. Mele, *Phys. Rev. Lett.* **98**, 106803 (2007).
- [81] R. W. Wyckoff, in *Crystal Structures* (Wiley, New York, 1964), Vol. 2.