Symmetry, topology, and geometry: The many faces of the topological magnetoelectric effect

Perry T. Mahon ●,[*] Chao Lei ●, and Allan H. MacDonald
Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA

A delicate tension complicates the relationship between the topological magnetoelectric effect (TME) in three-dimensional (3D) $\mathbb{Z}_2$ topological insulators (TIs) and time-reversal symmetry (TRS). TRS underlies a particular $\mathbb{Z}_2$ topological classification of the electronic ground state of crystalline band insulators and the associated quantization of the magnetoelectric response coefficient calculated using bulk linear-response theory but, according to standard symmetry arguments, simultaneously forbids a nonzero magnetoelectric coefficient in any physical finite-size system. This contrast between theories of magnetoelectric response in formal bulk models and in real finite-sized materials originates from the distinct approaches required to introduce notions of (electronic) polarization and orbital magnetization in these fundamentally different environments. In this work we argue for a modified interpretation of the bulk linear-response calculations in nonmagnetic $\mathbb{Z}_2$ TIs that is more plainly consistent with TRS and use this interpretation to discuss the effect's observation—still absent over a decade after its prediction. Our analysis is reinforced by microscopic bulk and thin-film calculations carried out using a simplified but still realistic effective model for the well established V_2VI_3 [V = (Sb, Bi) and VI = (Se, Te)] family of nonmagnetic $\mathbb{Z}_2$ TIs. When a uniform dc magnetic field is included in this model, the anomalous $n = 0$ Landau levels (LLs) play the central role, both in thin films and in bulk. In the former case, only the $n = 0$ LL eigenfunctions can support a dipole moment, which vanishes if there are no magnetic surface dopants and is quantized in the thick-film limit if magnetic dopants at the top and bottom surfaces have opposite orientation. In the latter case, the Hamiltonian projected into the $n = 0$ LL subspace is a one-dimensional Su-Schrieffer-Heeger model with ground-state polarization that is quantized in accordance with the bulk linear-response coefficient calculated for (a lattice regularization of) the starting 3D model. Motivated by analytical results, we conjecture a type of microscopic bulk-boundary correspondence: a bulk insulator with (generalized) TRS supports a magnetoelectric coefficient that is purely itinerant (which is generically related to the geometry of the ground state) if and only if magnetic surface dopants are required for the TME to manifest in finite samples thereof. We conclude that in nonmagnetic $\mathbb{Z}_2$ TIs the TME is activated by magnetic surface dopants, that the charge-density response to a uniform dc magnetic field is localized at the surface and specified by the configuration of those dopants, and that the TME is qualitatively less robust against disorder than the integer quantum Hall effect.

I. INTRODUCTION

One of Peierls' surprises in theoretical physics [1] is that the orbital magnetization of metals can be correctly calculated using infinite lattice models that neglect the surfaces of realistic material samples, even though that magnetization can be understood to arise from bound currents at those surfaces. This mysterious success is now often taken for granted. In recent times, strong interest in the multitude of topological insulators (TIs) and the linear response thereof has highlighted similar issues, in particular when disentangling the roles played by the topology ascribed to the bulk electronic ground state and by topologically protected electronic surface states. The integer quantum Hall effect (IQHE) provides a familiar example. In a two-dimensional (2D) crystalline band insulator the IQHE occurs when the vector bundle of occupied electronic Bloch states over the Brillouin zone (BZ) torus is characterized by a nonzero Chern invariant [2–4]. Meanwhile, experiments are often interpreted in terms of the gapless chiral edge states [5–9] whose presence is more immediately related to charge conduction. In the IQHE, it is generally accepted that this bulk topology of a band insulator implies a quantized bulk Hall conductivity but also implies the existence of chiral edge states that yield a consistent conductance in finite-size samples thereof, unifying the two interpretations. Notably, breaking time-reversal symmetry (TRS) is necessary for a nonzero Chern invariant and, by the usual symmetry arguments, for a nonzero Hall conductivity in finite-sized systems. The topological magnetoelectric effect (TME) presents an even more stark puzzle. In three-dimensional (3D) crystalline band insulators, TRS leads to a $\mathbb{Z}_2$ topological classification of the electronic ground state and in $\mathbb{Z}_2$-odd phases to a magnetoelectric linear-response coefficient that is argued [10,11] to equal $(n + 1/2)e^2/hc$ for $n \in \mathbb{Z}$. On the other hand, in
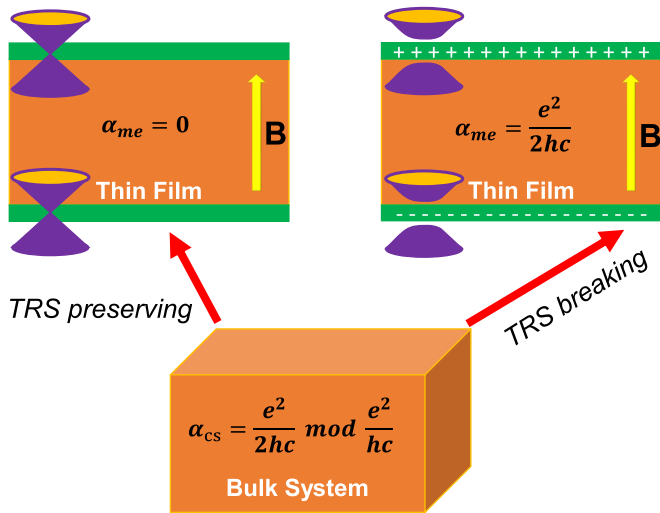
---

*perry.mahon@austin.utexas.edu

FIG. 1. Relationship between α_{CS} in bulk insulators that exhibit time-reversal symmetry (TRS) and α_{me} in thin films thereof. In the bulk, TRS leads to a $\mathbb{Z}_2$ topological classification of the electronic ground state, to which α_{CS} is sensitive. In three-dimensional $\mathbb{Z}_2$-odd phases, α_{CS} is quantized at $e^2/2hc$ up to an integer multiple of e^2/hc. In nonmagnetic thin films, however, the physically realized magnetoelectric response α_{me} is nonzero only when the energy dispersion of the topologically protected surface states is everywhere gapped by magnetic surface dopants which break TRS locally.

any finite-sized system with TRS, the usual symmetry arguments dictate that the magnetoelectric coefficient must vanish. Indeed, the requirement of magnetic surface dopants for realization of the TME has been noted previously [10–14]. There is no physical bulk response directly related to the 3D $\mathbb{Z}_2$ invariant, but a nontrivial $\mathbb{Z}_2$ invariant implies the existence of surface states that can then be gapped to activate the TME. This more convoluted connection has implications for the experimental robustness of the effect, making it qualitatively less robust against disorder compared with the IQHE.

It seems therefore that the magnetoelectric response of nonmagnetic $\mathbb{Z}_2$ TIs is a rare violation of the Peierls principle referred to above, i.e., that evaluating a quantity using a band theoretic description does not always correctly produce the value of that quantity in a large finite-sized sample thereof (see Fig. 1). In this paper we explicitly address the relationship between the magnetoelectric linear-response coefficient of a bulk insulator that exhibits TRS and that of a finite-size sample counterpart. As always, the bulk crystal is merely a convenient theoretical construct. Our interest is in understanding when it yields the physically correct magnetoelectric response. In Sec. II we describe this conundrum in more detail and highlight the main results of the calculations that follow. We argue for a slightly different interpretation of the bulk magnetoelectric coefficient derived in previous work [15], which is informed by the relationship (see Fig. 1) between particular linear-response coefficients in finite-sized samples and bulk that we explicitly establish in Secs. III and IV. Since the magnetoelectric coefficient lacks physical significance in macroscopically uniform bulk insulators that exhibit TRS, the bulk interpretation that we propose implicitly involves the

consideration of surfaces, a seemingly unavoidable feature. In this interpretation the magnetoelectric coefficient adheres to the Peierls principle and we reach a conclusion that is manifestly consistent with TRS; in nonmagnetic $\mathbb{Z}_2$ TIs, magnetoelectric response occurs only when the surface states are gapped by magnetic dopants, the magnetoelectric coefficient is quantized only when the surface magnetization configuration satisfies stringent conditions, and currents proportional to the magnetic field can flow through the material when the surface magnetization profile is changed.

In Sec. III we extract the magnetoelectric coefficient for thin-film and bulk samples of nonmagnetic $\mathbb{Z}_2$ TIs in the V_2VI_3 family of materials by directly including a uniform dc magnetic field in a coupled-Dirac cone model of the low-energy electronic states, then computing the polarization of the occupied energy eigenfunctions. In the case of thin films, we find that when the Dirac cones associated with the top and bottom surfaces are gapped by magnetic dopants of opposite magnetization, the magnetoelectric coefficient α_{me} is nonzero and approaches a quantized value as the film thickness is increased. The induced polarization is sensitive to the magnetic dopant configuration and α_{me} moves continuously through zero to $-\alpha_{me}$ as the component of magnetization perpendicular to the surface is continuously taken to its opposite value. When there is no surface magnetization, there is no magnetoelectric response. We show that the quantization of α_{me} in the thick film limit can be understood in terms of the topological properties of a Su-Schrieffer-Heeger model that arises in the bulk description.

In Sec. IV we employ the 3D tight-binding model that we obtain from a lattice regularization of the coupled-Dirac cone model used in Sec. III and calculate α_{CS}, the purported magnetoelectric coefficient in crystalline band insulators that exhibit TRS, semi-analytically using the well-known bulk linear-response expression; we reproduce the expected quantization. Calculating α_{CS} is technically challenging because it must be evaluated with respect to a smooth global gauge of the vector bundle of occupied Bloch states over the 3D BZ. Generically, Bloch energy eigenvectors that are smooth over the entire BZ do not exist in a $\mathbb{Z}_2$ TI [16] and a Wannierization-like procedure [17] is required to obtain an adequate gauge choice.[1] Fortunately, the model we employ exhibits a fermionic time-reversal symmetry and an inversion symmetry, thus each energy band is at least twofold degenerate over the BZ. In fact, the energy bands are everywhere pairwise isolated. Thus there exists a smooth global Hamiltonian gauge, and with respect to such a gauge we make our calculations. In Sec. V we analytically demonstrate that, in a particular smooth global Hamiltonian gauge, α_{CS} is entirely attributed to the itinerant portion of the electric-field-induced orbital magnetization, and speculate that this is a universal feature unique to nonmagnetic $\mathbb{Z}_2$ TIs. Our calculations explicitly demonstrate that for $\mathbb{Z}_2$ TIs, the bulk theory of magnetoelectric response does not always capture the properties of large finite-size samples.

---

[1]See, e.g., Essin *et al.* [11].

II. LINEAR-RESPONSE THEORY

The interaction between macroscopic electromagnetic fields and the charged constituents of material media is often described at the level of linear response by introducing phenomenological susceptibility tensors that are nonlocal in space and time, and relate changes in the (macroscopic) charge $\varrho(\boldsymbol{r},t)$ and current $\boldsymbol{J}(\boldsymbol{r},t)$ densities of the material to the applied electric $\boldsymbol{E}(\boldsymbol{r},t)$ and magnetic $\boldsymbol{B}(\boldsymbol{r},t)$ Maxwell fields that induce those changes. The charge and current densities and the Maxwell fields are understood to be coarse grained[2] over a length scale intermediate between the atomic spacing and the wavelength of light; indeed local field corrections, which can be important, are not of interest here and are neglected. It can often be assumed that a bulk material is spatially uniform at that intermediate length scale, in which case the susceptibility tensors are translationally invariant. Then, for example, the linear response of $\boldsymbol{J}(\boldsymbol{r},t)$ is of the form

$$J^{i(1)}(\boldsymbol{q},\omega) = \sigma^{il}(\boldsymbol{q},\omega)E^l(\boldsymbol{q},\omega), \tag{1}$$

where $\sigma^{il}(\boldsymbol{q},\omega)$ is the effective conductivity tensor, $\boldsymbol{q}$ and ω are respectively the wave vectors and frequencies that arise in the Fourier transforms of $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$, and superscript indices here and below identify Cartesian components that are summed over when repeated. If $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$ only involve wavelengths in the optical regime, then $\sigma^{il}(\boldsymbol{q},\omega)$ can be expanded in powers of $\boldsymbol{q}$,

$$\sigma^{il}(\boldsymbol{q},\omega) = \sigma^{il}(\omega) + \sigma^{ilj}(\omega)q^j + \cdots, \tag{2}$$

where $\sigma^{il}(\omega) \equiv \sigma^{il}(\boldsymbol{0},\omega)$ and $\sigma^{ilj}(\omega) \equiv (\partial\sigma^{il}(\boldsymbol{q},\omega)/\partial q^j)|_{\boldsymbol{q}=\boldsymbol{0}}$. Using Eq. (2) in (1) and performing a Fourier transform to coordinate space yields

$$J^{i(1)}(\boldsymbol{r},\omega) = \sigma^{il}(\omega)E^l(\boldsymbol{r},\omega) - i\sigma^{ilj}(\omega)\frac{\partial E^l(\boldsymbol{r},\omega)}{\partial r^j} + \cdots. \tag{3}$$

The first term in Eq. (3) is the familiar long-wavelength frequency-dependent conductivity tensor. Using Faraday's law, the second term can be shown to include contributions that involve $\boldsymbol{B}(\boldsymbol{r},\omega)$ and symmetrized spatial derivatives of $\boldsymbol{E}(\boldsymbol{r},\omega)$. Any susceptibility in a material that is uniform at the course grained length scale can be analyzed in this way.

Physical insight into the distribution and dynamics of the charged (quasi-)particles that constitute a material medium can be gleaned by identifying bound and free contributions to $\varrho(\boldsymbol{r},t)$ and $\boldsymbol{J}(\boldsymbol{r},t)$, and associating the former with macroscopic polarization $\mathscr{P}(\boldsymbol{r},t)$ and (orbital) magnetization $\mathscr{M}(\boldsymbol{r},t)$ fields[3] such that

$$\varrho(\boldsymbol{r},t) = -\boldsymbol{\nabla}\cdot\mathscr{P}(\boldsymbol{r},t) + \varrho_F(\boldsymbol{r},t),$$

$$\boldsymbol{J}(\boldsymbol{r},t) = \frac{\partial\mathscr{P}(\boldsymbol{r},t)}{\partial t} + c\boldsymbol{\nabla}\times\mathscr{M}(\boldsymbol{r},t) + \boldsymbol{J}_F(\boldsymbol{r},t). \tag{4}$$

For a given $\varrho(\boldsymbol{r},t)$ and $\boldsymbol{J}(\boldsymbol{r},t)$ there is always ambiguity in defining $\mathscr{P}(\boldsymbol{r},t)$, $\mathscr{M}(\boldsymbol{r},t)$, $\varrho_F(\boldsymbol{r},t)$, $\boldsymbol{J}_F(\boldsymbol{r},t)$ that satisfy Eq. (4). Nevertheless, let us assume that definitions for

$\mathscr{P}(\boldsymbol{r},t)$ and $\mathscr{M}(\boldsymbol{r},t)$ in a bulk material have been made and that the applied Maxwell fields are in the linear-response regime so that expansions of the electric and magnetic dipole moments[4] as sums of spontaneous and field-dependent contributions,

$$P^i(\boldsymbol{r},\omega) = P^{i(0)} + \chi_E^{il}(\omega)E^l(\boldsymbol{r},\omega) + \alpha_P^{il}(\omega)B^l(\boldsymbol{r},\omega) + \cdots,$$

$$M^i(\boldsymbol{r},\omega) = M^{i(0)} + \alpha_M^{il}(\omega)E^l(\boldsymbol{r},\omega) + \cdots, \tag{5}$$

are justified.[5]

At low frequencies ω, $\alpha_P^{il}(\omega)$, and $\alpha_M^{il}(\omega)$ are well approximated by their static $\omega = 0$ values, which are in fact related by a thermodynamic Maxwell relation

$$\alpha_P^{il}(\omega=0) = \alpha_M^{li}(\omega=0) \equiv \alpha^{il}. \tag{6}$$

In this case, using Eq. (5) in (4) and comparing with Eq. (2) yields

$$\sigma^{ilj} = -ic\alpha^{ik}\epsilon^{kjl} + ic\epsilon^{ijb}\alpha^{bl} + \cdots, \tag{7}$$

where "$\cdots$" denotes purely electric contributions to σ^{ilj}.

Explicit forms for phenomenological susceptibilities can be obtained from quantum linear-response theory. The semiclassical Hamiltonian of the coupled light-matter system consists of electron and Maxwell contributions, and an interaction term that results from minimal coupling. The latter involves the electromagnetic scalar and vector potentials multiplied by charge and current-density operators that are obtained from components of the Noether 4-current of the electron theory. In a finite sized material those charge and current densities are nonzero only in a localized region of space, in which case there exists well-motivated definitions for polarization and magnetization fields developed by Power, Zinau, and Wooley (PZW),[6] Healy,[7] and others. This approach begins with a unitary transformation of the minimal-coupling Hamiltonian (differential operator) to yield the physically equivalent PZW Hamiltonian. After defining the polarization and magnetization fields, the interaction of the electronic degrees of freedom with the electromagnetic field as described in the PZW Hamiltonian involves terms in which the polarization and magnetization fields are multiplied by $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$, respectively;[8] notably this formulation does not

---

[2] See, e.g., Chap. 2 of Swiecicki [18] and references therein.

[3] See, e.g., Chap. 6.7 of Jackson [19] or Chap. 2 of Swiecicki [18].

---

[4] In writing Eq. (4), $\mathscr{P}(\boldsymbol{r},t)$ and $\mathscr{M}(\boldsymbol{r},t)$ can be understood as an infinite sum of electric and magnetic multipole moments (see, e.g., Chap. 6.7 of Jackson [19] or Chap. 2 of Swiecicki [18]). It is often taken implicitly that it is the spontaneous electric and magnetic dipole moments and their linear response that is of primary interest. Of course, this is not always true. In fact, the electric-quadrupole response contributes to the linear response of charge and current densities at first-order in $\boldsymbol{q}$, just like the magnetoelectric response. Nevertheless, in this work we will not explicitly consider response beyond that of the dipole moments.

[5] Here "$\cdots$" denote other contributions to the linear response of the electric and magnetic dipole moments. For example, the linear response of $\boldsymbol{P}$ to the symmetrized spatial derivative of $\boldsymbol{E}$ and of $\boldsymbol{M}$ to $\boldsymbol{B}$ are contained in "$\cdots$."

[6] See, e.g., Ref. [20] and references therein.

[7] See, e.g., Ref. [21] and references therein.

[8] For an overview of this procedure, see, e.g., Chap. 2 of Swiecicki [18].

require an explicit choice of the electromagnetic gauge. When a multipole expansion of those polarization and magnetization fields is made, the PZW Hamiltonian takes the classically anticipated form of a multipole Hamiltonian.[9] Thus, the PZW definitions for polarization and magnetization fields are physically reasonable, and from them the (minimally coupled) electronic charge and current-density expectation values, and thus the finite-size sample analog of the susceptibility tensors mentioned above (which are position dependent), can be rigorously obtained. Moreover, that multipole expansion results in Hermitian operators in the electronic Hilbert space; that is, the electric and magnetic multipole moments are genuine physical observables. For example, the ground-state expectation value of the electronic contribution to the PZW electric dipole moment per unit volume (which is equal to the interior polarization of the material) is

$$\boldsymbol{P}_{\text{el}}^{(0)} = -\frac{e}{\Omega} \sum_{E < E_F} \int \boldsymbol{r} |\Psi_E(\boldsymbol{r})|^2 d\boldsymbol{r}, \tag{8}$$

where $e > 0$ is the elementary charge, Ω is the finite volume of the sample, E_F is the Fermi energy, and $\Psi_E(\boldsymbol{r})$ are the electronic energy eigenfunctions of the unperturbed Hamiltonian.[10]

Unfortunately the PZW approach cannot be directly applied to crystalline solids since, even in the absence of electromagnetic fields, the electronic charge and current-density expectation values are expressed in terms of Bloch energy eigenfunctions, the support of which is all of space.[11] Indeed, this is related to the property that the usual position operator is not well-defined in the Hilbert space of Bloch functions [22]. There are a variety of different approaches [22–27] that can be used to extend the notions of polarization and magnetization to bulk crystals. Inspired by the PZW approach, Sipe *et al.* have developed a formalism [27–29] applicable to general extended systems, crystalline or otherwise, which has been employed to account for spatial variation of electromagnetic fields in crystal insulators [30,31], and more [32–36]. However, none of these bulk crystal approaches are able to define electric and magnetic multipole moments as genuine physical observables in the PZW sense.

The most commonly used approach to sidestep difficulties in defining the electric and magnetic dipole moments in crystalline insulators is the so-called modern theories of polarization [23,37] and (orbital) magnetization [24,25]. Focusing on the former, one aims to deduce an electric dipole moment from calculation of the current density, rather than to propose a general definition for it. To do so, the macroscopic free current density [appearing in the second of Eq. (4)] is assumed to vanish at linear response in typical insulators, which

seems to be a sensible assignment, thus the current density is related only to the polarization and magnetization fields. It is less obvious that this assignment is sensible in Chern insulators, but materials of that type are not considered here. Next, focus is restricted to describing the influence of electric and magnetic fields that are spatially uniform, in which case all of the macroscopic densities appearing in Eq. (4) are also expected to be uniform. Thus the polarization and magnetization fields are entirely described by their dipole moments, and $\boldsymbol{J}(t) = \partial \boldsymbol{P}(t)/\partial t$.

The magnetoelectric susceptibility α^{il} can generally be written as the sum of an isotropic $\delta^{il}\alpha_\theta$ and a traceless contribution [15]. Examining Eq. (7) makes it evident that in a static and macroscopically uniform bulk material, the σ^{ilj} tensor is insensitive to the former. That is, α_θ cannot be determined by calculating the wave-vector-dependent optical conductivity. In other words, the physical implications of α_θ are not equivalent to those of a wave-vector dependent conductivity.

If the isotropic magnetoelectric response cannot be calculated from the bulk current response to nonuniform electric fields, how can it be found? Two independent strategies that can be used to obtain α_θ have been identified in the literature: (i) Essin *et al.* [15] consider an insulating state of a Bloch Hamiltonian that is adiabatically varied in time, such that α^{il} becomes time-dependent and there is an additional contribution to Eq. (7) (originating from $\partial \boldsymbol{P}/\partial t$) proportional to $(\partial \alpha^{il}/\partial t)B^l$; (ii) Malashevich *et al.* [38] consider a finite sized material, in which case α^{il} becomes position-dependent and there is an additional contribution to Eq. (7) (originating from $\boldsymbol{\nabla} \times \boldsymbol{M}$) that is proportional to $(\epsilon^{iab}\partial \alpha^{bl}/\partial r^a)E^l$. Both approaches reach the same conclusion, which is that in a crystalline insulator initially occupying its zero-temperature electronic ground state

$$\alpha^{il} = \alpha_{\text{G}}^{il} + \delta^{il}\alpha_{\text{CS}}, \tag{9}$$

where the explicit form of α_{G}^{il} resembles that of a conventional linear-response tensor, involving cross-gap matrix elements between Bloch states that are initially occupied and unoccupied, and α_{CS} (the Chern-Simons contribution) can be expressed in terms of occupied Bloch states alone. If the bulk insulator exhibits TRS then $\alpha^{il} = \delta^{il}\alpha_{\text{CS}}$ and is quantized since α_{CS} evaluates to an element of a discrete lattice of values [10], which does not include zero in the case of $\mathbb{Z}_2$ TIs. It is at the step that the conundrum mentioned in the introduction is introduced, since this is where the conclusion is reached that $\alpha^{ii} \bmod e^2/hc$[12] can be nonzero in insulators with TRS.

The strategy (i) of Essin *et al.* [15] considers a 3D crystalline insulator described by a Bloch Hamiltonian that varies in time via adiabatic variation of the crystal Hamiltonian. They show that in the presence of a static and uniform magnetic field $\boldsymbol{B}$, adiabatic linear response to slow changes in the crystal Hamiltonian induces a spatially uniform macroscopic (minimal-coupling) electronic current density of the form $\boldsymbol{J}^{(B)}(t) = \boldsymbol{J}_{\text{G}}^{(B)}(t) + \boldsymbol{J}_{\text{CS}}^{(B)}(t)$.[13] Unlike above, where the

---

[9]Compare, e.g., Eqs. (4.24) and (5.72) of Jackson [19] with Eq. (2.86) of Swiecicki [18].

[10]In Sec. III we do not consider the static magnetic field as a perturbation and therefore $\boldsymbol{P}_{\text{el}}^{(0)}$ will there involve $\boldsymbol{B}$.

[11]In deriving the PZW Hamiltonian, it is assumed that the charge and current densities do not extend to infinity so integrals involving these quantities and spatial derivatives may be integrated by parts and the surface terms can be taken to vanish.

---

[12]For a discussion on the physical basis of the discrete ambiguity that is inherent to α^{ii}, see, e.g., Chap. 6.2 of Vanderbilt [39].

[13]Importantly, $\boldsymbol{J}^{(B)}(t)$ is not the magnetic-field-induced electronic current in a material with unperturbed electrons governed by the

time dependence of $J^{(1)}(t)$ resulted from that of the dynamical electromagnetic field, $\boldsymbol{B}$ is here taken to be static and the t-dependence results from that of the Bloch Hamiltonian itself; t-dependence manifests implicitly in the Bloch energy eigenvectors and eigenvalues. Although their strategy applies more generally, we focus on the special case in which the Bloch Hamiltonian exhibits TRS at every instant t. Then $J_{\mathrm{G}}^{(B)}(t) = \boldsymbol{0}$ and [15,40]

$$\boldsymbol{J}^{(B)}(t) = \boldsymbol{J}_{\mathrm{CS}}^{(B)}(t) = -\frac{e^2 \boldsymbol{B}}{2hc} \int_{\mathrm{BZ}} \left(\frac{\partial}{\partial t} Q_3(\boldsymbol{k};t) \right) d^3k. \quad (10)$$

The explicit expression for $Q_3(\boldsymbol{k};t)$, the Chern-Simons 3-form[14] at instant t, is discussed in Sec. IV. It is important to note that $Q_3(\boldsymbol{k};t)$ explicitly involves only the Berry connection and curvature that are defined on the vector bundle of occupied Bloch states over the Brillouin zone torus at instant t. That is, at every t, $Q_3(\boldsymbol{k};t)$ is a purely geometrical quantity. In Ref. [15] the derivative with respect to time in Eq. (10) is taken outside of the Brillouin zone integral to obtain

$$\boldsymbol{J}^{(B)}(t) = \frac{\partial \boldsymbol{P}^{(B)}(t)}{\partial t} = \boldsymbol{B} \frac{\partial}{\partial t} \alpha_{\mathrm{CS}}(t), \quad (11)$$

where

$$\alpha_{\mathrm{CS}}(t) \equiv -\frac{e^2}{2hc} \int_{\mathrm{BZ}} Q_3(\boldsymbol{k};t) d^3k. \quad (12)$$

At each instant t,

$$\int_{\mathrm{BZ}} Q_3(\boldsymbol{k};t) d^3k \in \mathbb{Z} \quad (13)$$

is determined (modulo 2) by the TRS-induced $\mathbb{Z}_2$ topological classification [10,42,43] of the electronic ground state[15] and the value of $\alpha_{\mathrm{CS}}(t)$ is therefore fixed (modulo e^2/hc) by that topology; if the ground state is $\mathbb{Z}_2$-even then α_{CS} evaluates to 0 mod e^2/hc and if it is $\mathbb{Z}_2$-odd then α_{CS} evaluates to $e^2/2hc$ mod e^2/hc [10,11]. Importantly, $Q_3(\boldsymbol{k};t)$ and indeed $\alpha_{\mathrm{CS}}(t)$ are gauge dependent—they are both sensitive to the smooth global frame of the vector bundle of occupied Bloch states over the Brillouin zone torus that is used for their evaluation—which leads to the above described discrete ambiguity. Therefore, even when considering instants t_1 and t_2 between which a band insulating ground state remains in the same topological class, $\alpha_{\mathrm{CS}}(t_1)$ need not equal $\alpha_{\mathrm{CS}}(t_2)$ since at

---

Bloch Hamiltonian evaluated at t; this would be given by $J_{\mathrm{G}}^{(B)}(t)$. Instead, $\boldsymbol{J}^{(B)}(t)$ also involves Streda-type transport, which manifests through the time-dependence of the ground-state density matrix and leads to $J_{\mathrm{CS}}^{(B)}(t)$ [see Eqs. (35b) and (35c) of Ref. [15]].

[14]See, e.g., Chap. 11 of Nakahara [41].

[15]More precisely, TRS induces a topological classification of the vector bundle (for reference, see, e.g., Chap. 10 of Lee [44] or Chap. 9.3 of Nakahara [41]) that is naturally constructed using the Bloch states that are occupied in the ground state of a band insulator [45,46]. We adopt the common terminology that ascribes the topology of that vector bundle to the ground state itself. In this context, the $\mathbb{Z}_2$ topological classification of the ground state is most intuitively understood in terms of the types of gauge choices (frames) of that bundle. If there exists a smooth global TRS frame of that bundle then the ground state is called $\mathbb{Z}_2$-even, otherwise it is called $\mathbb{Z}_2$-odd.

each t one has the freedom to choose one of many gauges and with respect to each of those gauge choices a different value of $\alpha_{\mathrm{CS}}(t) \in (e^2/2hc)\mathbb{Z}$ can result; in this scenario it is only guaranteed that $\alpha_{\mathrm{CS}}(t_1)$ mod e^2/hc equals $\alpha_{\mathrm{CS}}(t_2)$ mod e^2/hc.

To obtain Eq. (10) it must be assumed that the gauge choice used to express $Q_3(\boldsymbol{k};t)$ is smooth in $\boldsymbol{k} \in$ BZ for each t. And in order to move from Eq. (10) to (11), that gauge choice must also be continuous in t.[16] The latter condition implies that $\alpha_{\mathrm{CS}}(t)$ is continuous in t. If there is TRS at each t then $\alpha_{\mathrm{CS}}(t) \in (e^2/2hc)\mathbb{Z}$ and thus $\alpha_{\mathrm{CS}}(t)$ must be constant in t. Indeed our explicit calculations in Sec. IV demonstrate that once a gauge choice is made at any t, the value of $\alpha_{\mathrm{CS}}(t)$ is fixed for all t if that gauge is continuous in t. Then under any TRS preserving adiabatic variation of the Bloch Hamiltonian in time, the right-hand side of Eq. (11) evaluates to zero; no bulk currents flow when the parameters of the Hamiltonian change within the space of a given topological phase and the adiabatic approximation is valid. Indeed this result should be expected since Eq. (11) is consistent with TRS only if nonzero values for $\partial \alpha_{\mathrm{CS}}(t)/\partial t$ are activated by some time-reversal symmetry-breaking perturbation.

If a Bloch Hamiltonian is adiabatically varied in t without necessarily exhibiting TRS at each t, then $\alpha_{\mathrm{CS}}(t)$ need not evaluate to an element of $(e^2/2hc)\mathbb{Z}$ for each t [10] and therefore $\alpha_{\mathrm{CS}}(t)$ is not necessarily constant in t. For example, if there is TRS at some initial t_i and final t_f times, then $\alpha_{\mathrm{CS}}(t_i)$, $\alpha_{\mathrm{CS}}(t_f) \in (e^2/2hc)\mathbb{Z}$ and if the electronic ground state at t_i (t_f) is $\mathbb{Z}_2$-even ($\mathbb{Z}_2$-odd), then $\int_{t_i}^{t_f} \boldsymbol{J}^{(B)}(t)dt$ mod $e^2\boldsymbol{B}/hc$ equals $e^2\boldsymbol{B}/2hc$. Crucially, this ambiguity in $\int_{t_i}^{t_f} \boldsymbol{J}^{(B)}(t)dt$ is not a consequence of the gauge dependence of α_{CS}, but rather a manifestation of the fact that the current pumped through the material depends on how the model parameters are changed in time. Indeed if the Bloch Hamiltonian evaluated at t_i and at t_f equal, then one can define a vector bundle of occupied electronic states over BZ $\times S^1$ and the topology of this structure is characterized by a second Chern invariant which determines $\int_{t_i}^{t_f} \boldsymbol{J}^{(B)}(t)dt$ [10,47]. A similar situation arises in the case of the unperturbed bulk 1D polarization, where a certain first Chern invariant determines the bound charge [39,48].

The linear-response result (11) for crystalline insulators that exhibit TRS has been interpreted [15] by identifying the magnetic-field-induced bulk polarization $\boldsymbol{P}^{(B)}(t)$ with $\alpha_{\mathrm{CS}}(t)\boldsymbol{B}$. For a static insulator of interest, its bulk $\boldsymbol{P}^{(B)}$ is then obtained by evaluating $\alpha_{\mathrm{CS}}(t)\boldsymbol{B}$ at any time t at which the time-dependent Bloch Hamiltonian describes that insulator. If we accept this identification, then $\boldsymbol{P}^{(B)}$ can differ from the physically realized electric dipole response in a large but finite sample [49]. For example, consider a finite crystallite of a nonmagnetic $\mathbb{Z}_2$ TI that has TRS everywhere. TRS implies that there is no electric dipole induced by $\boldsymbol{B}$ in linear response. The bulk expression for the polarization is evidently

---

[16]For there to exist a smooth global gauge of the vector bundle of occupied Bloch states over BZ at instant t implies that the bulk bandgap does not vanish at that t. Therefore, this assumption can only be satisfied if the bulk bandgap is nonzero throughout the duration of time during which the crystalline parameters are varied.

misleading in the absence of magnetic dopants somewhere. Because this identification of $\boldsymbol{P}^{(B)}$ does not imply induced bulk charge or current densities it is still technically acceptable as a bulk quantity, but associating physical significance with it requires additional insight. Indeed the necessity of magnetic surface dopants for the topological magnetoelectric effect to manifest in nonmagnetic TIs has been noted [10,11,49,50] previously. A ubiquitous aspect of bulk theories of polarization and orbital magnetization is the importance of interpretation [39]. We believe it is useful to formulate an alternative, physically equivalent, perspective to that of Ref. [49]. By combining the known topological constraints in bulk with the known symmetry constraints in finite-sized systems, we produce below a physically meaningful notion of $\boldsymbol{P}^{(B)}$ in bulk insulators that exhibit TRS, one that adheres to the Peierls principle and implicitly accounts for additional surface-related criteria. To that end, the V_2VI_3 family of TIs are an ideal test bed on which to develop that notion.

We have argued above that even in the strategy (i) of Ref. [15], wherein the Bloch Hamiltonian for a bulk insulator involves crystal parameters that depend on time, $\alpha_{\text{CS}}(t)\boldsymbol{B}$ lacks physical consequences if that Hamiltonian exhibits TRS at each t. Thus, in this setting any physically significant identification of $\boldsymbol{P}^{(B)}(t)$ must involve the consideration of surfaces. As a result, any such identification will be *ad hoc* in that one has to artificially include the role played by the surface in a proposed bulk quantity. This is an aesthetically unappealing, but a seemingly unavoidable, feature of *any* physically meaningful identification of $\boldsymbol{P}^{(B)}(t)$ in bulk insulators with TRS.

One such identification of $\boldsymbol{P}^{(B)}(t)$ could be obtained in a manner similar to that of Essin *et al.* [15]—by considering the magnetic-field-dependent current $\boldsymbol{J}^{(B)}(t)$ that flows in a material as its properties are changed—but, rather than considering a time-dependent Bloch Hamiltonian, consider a Hamiltonian that describes a macroscopically large but finite-sized material with time-dependence only at its surfaces, where magnetic moments order to break TRS and open gaps in the Dirac-like energy dispersion of the surface states. (TRS is maintained in the interior region.) If the effect of interactions that break TRS is local [51], then the charge-density response to $\boldsymbol{B}$ can only occur near the surfaces. Our interpretation for nonmagnetic $\mathbb{Z}_2$ TIs,[17] which is supported by later calculations in this paper, imagines replacing the bulk quantity $\alpha_{\text{CS}}(t)$ by a new function $\alpha_{\text{expt}}(t)$ that compares more directly to experiment[18] by implicitly accounting for the role of surface magnetic dopants. The (physically meaningful) bulk polarization in these materials is

$$\boldsymbol{P}^{(B)}(t) = \alpha_{\text{expt}}(t)\boldsymbol{B}, \tag{14}$$

where $\alpha_{\text{expt}}(t)$ must have the following properties:

---

[17]In other bulk insulators that exhibit an effective TRS, such as antiferromagnetic $\mathbb{Z}_2$ TIs [52], that symmetry is typically broken merely by the existence of a surface. In that case the following is unnecessary since there is no conundrum related to TRS and magnetoelectric response, and the consideration of magnetic surface dopants seems unnecessary.

[18]In principle, experiments only probe polarization differences (see, e.g., Chap. 4.4 of Vanderbilt [39]).
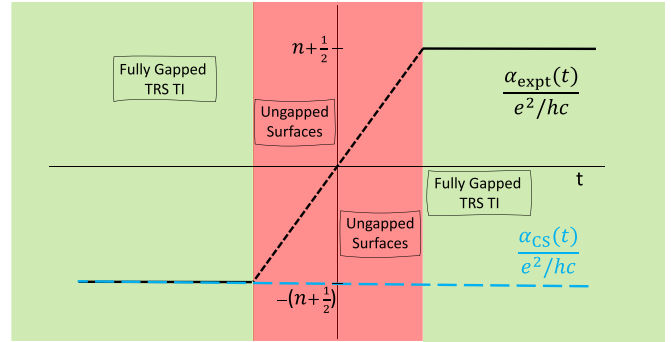

FIG. 2. Relationship between the bulk magnetoelectric coefficient $\alpha_{\text{expt}}(t)$ (black line) that we propose and $\alpha_{\text{CS}}(t)$ (blue line) as obtained from a simple interpretation of the bulk current that is adiabatically induced by slow changes of a Bloch Hamiltonian describing a nonmagnetic bulk $\mathbb{Z}_2$ TI. The illustrated time interval is imagined to be one over which the magnetization describing a configuration of surface magnetic dopants that initially gaps the energy dispersion of the surface states is changed to its opposite. The red region, where this magnetization vanishes, is given finite width for illustrative purposes. We imagine making a particular gauge choice at some initial time such that $\alpha_{\text{CS}}(t)$ exactly agrees with $\alpha_{\text{expt}}(t)$ until a surface gap closes. A purely bulk linear-response calculation [15] uses the adiabatic approximation to show that $\alpha_{\text{CS}}(t)$ is time-independent and the (magnetic field dependent) bulk current $\boldsymbol{J}^{(B)}(t)$ vanishes (no currents flow) when an instantaneous band insulating state is varied in a manner that maintains TRS. We argue that in finite samples $\int \boldsymbol{J}^{(B)}(t)dt$ is insensitive to gapped surfaces and is therefore given correctly by the bulk linear-response calculation at early and late times within the illustrated time interval. The adiabatic approximation fails to describe finite samples when the surface gap vanishes, in which case the physically realized $\boldsymbol{J}^{(B)}(t)$ is nonzero and the polarization can change. In practice, the subinterval within which $\boldsymbol{J}^{(B)}(t) \neq \boldsymbol{0}$ will always be finite because of disorder at the surface. We argue that the interior polarization at a particular point along the surface is a local quantity [10] that changes sign with the local orientation of magnetic dopants, in agreement with TRS.

(1) In the absence of magnetic surface dopants $\boldsymbol{J}^{(B)}(t)$ vanishes in finite samples, which is consistent with the bulk prediction (11). In this case, TRS dictates that $\alpha_{\text{expt}}(t) \equiv 0$ and the identification of $\boldsymbol{P}^{(B)}(t)$ with $\alpha_{\text{CS}}(t)\boldsymbol{B}$ for $\mathbb{Z}_2$ TI materials is incorrect.

(2) (Green regions in Fig. 2). During time intervals over which the magnetic dopant configuration on top and bottom surfaces is opposite and such that the energy dispersion of the surface states remains gapped, the physically realized response is that of the bulk; the prediction (11) that $\boldsymbol{J}^{(B)}(t)$ vanishes and the identification of $\boldsymbol{P}^{(B)}(t)$ with $\alpha_{\text{CS}}(t)\boldsymbol{B}$ are correct (the latter modulo $e^2\boldsymbol{B}/hc$), i.e., $\alpha_{\text{expt}}(t) = \alpha_{\text{CS}}(t)$. $\boldsymbol{P}^{(B)}(t)$ is constant in t as long as the surface states remain gapped.

(3) (Red region in Fig. 2). During time intervals over which the magnetic dopant configuration is such that there is a finite surface density of states, a current proportional to $\boldsymbol{B}$ flows through the material. Indeed, the bulk prediction (11) using the adiabatic approximation fails to describe these finite-size samples and $\alpha_{\text{expt}}(t) \neq \alpha_{\text{CS}}(t)$. That current, and the pattern of polarization that develops, is sensitive to the magnetic configuration of the surface dopants. For realistic

samples with potential and magnetic disorder, the surface gaps will remain closed over a finite range of dopant configurations. Under surface magnetization reversal, currents proportional to $\boldsymbol{B}$ will flow through the material during the time interval over which there is no surface energy gap. $\alpha_{\mathrm{expt}}(t)$ now has a contribution that is sensitive to the surface magnetization, which allows the polarization to interpolate between the topologically allowed values of gapped states.

Our interpretation is that during time intervals $t \in [t_i, t_f]$ over which the surface magnetization is reversed, the value of $\boldsymbol{P}^{(B)}(t)$ changes and these changes are associated with the flow of currents through the bulk of the material. Varying the configuration of magnetic surface dopants in such a way that the total magnetization always vanishes but that surface gaps close then reopen induces bulk currents in much the same way that varying crystal parameters such that at intermediary times TRS is broken in the bulk[19] does; notably, neither mechanism yields a bulk charge density. The change in polarization may be obtained by integrating $[\partial \alpha_{\mathrm{expt}}(t)/\partial t]\boldsymbol{B}$ over time, which, due to TRS in the bulk, will integrate to the difference[20] between two quantized values of $\alpha_{\mathrm{CS}}(t)$ that differ in sign. Since in our interpretation we assume that the TRS-induced bulk topology is unaffected by the manipulation of the magnetic surface dopants, $\alpha_{\mathrm{expt}}(t_f) - \alpha_{\mathrm{expt}}(t_i) \in (e^2/hc)\mathbb{Z}$. If the magnetization of the dopants at t_f is opposite to that at t_i, we require $\alpha_{\mathrm{expt}}(t_f) = -\alpha_{\mathrm{expt}}(t_i)$.

Under this identification, the magnetoelectric response in nonmagnetic $\mathbb{Z}_2$ TIs is nonzero only when magnetic surface dopants are present, solving the TRS conundrum. If there is no surface magnetization, then there will never be a surface energy gap everywhere and no bulk current flows in linear response to a magnetic field. In the following sections of this paper we describe electronic polarization (orbital magnetization) calculations in model nonmagnetic $\mathbb{Z}_2$ TIs subject to a uniform dc magnetic (electric) field, which support this interpretation. In particular, in sufficiently thick films the polarization changes and interior currents flow only when the surface-normal magnetization component is varied through zero.

III. MAGNETOELECTRIC RESPONSE IN $\mathbb{Z}_2$ TOPOLOGICAL INSULATOR THIN FILMS

In this section we employ a simplified but realistic model of the electronic states near the Fermi energy in thin films

of V_2VI_3-type nonmagnetic $\mathbb{Z}_2$ TIs in which it is possible to account for the presence or absence of magnetic surface dopants. We calculate the magnetoelectric coefficient α_{me} by introducing a magnetic field that is perpendicular to the thin film and accounting for the Landau quantization it produces. We begin in Sec. III A by considering a toy version of this model in which the only electronic states retained are the topologically protected surface states modeled by two coupled two-dimensional (2D) Dirac cones, one associated with each surface. The surface magnetic dopants couple to these Dirac-like states separately via an exchange mass m, and for finite-thickness films those states are hybridized by an intersurface tunneling parameter Δ. This model is amenable to analytic analysis, from which it is found that in the limit $m \to 0 \; \alpha_{\mathrm{me}} \to 0$, as expected for finite-sized material samples with TRS, and in the limit $\Delta \to 0 \; \alpha_{\mathrm{me}} \to e^2/2hc$, as expected for bulk $\mathbb{Z}_2$ TIs. In Sec. III B we consider a more realistic model, which involves many coupled Dirac cones, and reach the same conclusions. We show that the quantization of α_{me} in the thick film limit can be understood in terms of the topological properties of a Su-Schrieffer-Heeger model that arises in the bulk description.

A. Simplified toy model of the surface states

The analytic calculations in this section illustrate the essential microscopic physics of magnetic-field-induced charge redistribution in finite-thickness thin-films of nonmagnetic $\mathbb{Z}_2$ TIs. In materials of this type, the electronic states that are nearest to the Fermi energy occur about the $\boldsymbol{k}_\Gamma$ point $(k_{\Gamma,x}, k_{\Gamma,y}) = (0, 0)$ and are localized at the surface of the sample [53,54]. Indeed the bulk topology of a $\mathbb{Z}_2$ TI implies the existence of an odd number of Dirac points at each surface of a finite sample thereof [55,56]. A model of a thin film can therefore be constructed using two copies of a 2D $\boldsymbol{k} \cdot \boldsymbol{p}$ Dirac Hamiltonian, one associated with each surface. We account for the (finite) film thickness in the surface-normal direction (taken to be $\boldsymbol{z}$) by hand by endowing the basis states of each copy of the $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian with z-dependence to encode its association with a particular surface.

We therefore consider two Dirac Hamiltonians written in a basis of states $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}=\boldsymbol{k}_\Gamma}\rangle$ that are orthogonal linear combinations of the $\boldsymbol{k} \cdot \boldsymbol{p}$ basis states and are associated with the bottom ($\alpha = 0$) and top ($\alpha = 1$) surfaces of the film. At $\boldsymbol{k}_\Gamma$, the $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian will commute with s_z, thus the corresponding spin-up and spin-down eigenvalues $\sigma = \pm \hbar/2$ ($\equiv \uparrow, \downarrow$) can be used to identify Hilbert space basis states at $\boldsymbol{k}_\Gamma$ (and therefore label basis states of any $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian about $\boldsymbol{k}_\Gamma$). We include an exchange coupling m_α associated with each surface, here describing the effect of magnetic dopants at that surface, that couples to the s_z spin components of the states at the surface α; this interaction breaks TRS. We take the exchange mass to be of opposite[21] value at the top and bottom surfaces, $m_0 = -m_1 \equiv m$ [50,58]. We also include an intersurface hybridization parameter Δ. In the basis $(|\bar{u}_{(0,\uparrow),\boldsymbol{k}_\Gamma}\rangle, |\bar{u}_{(0,\downarrow),\boldsymbol{k}_\Gamma}\rangle, |\bar{u}_{(1,\uparrow),\boldsymbol{k}_\Gamma}\rangle, |\bar{u}_{(1,\downarrow),\boldsymbol{k}_\Gamma}\rangle)$, the Hamiltonian

---

[19]See, e.g., the discussion at the beginning of Sec. 3 of Ref. [15] or Chap. 6.4.2 of Ref. [39]. If the crystalline parameters are adiabatically varied such that the infinite material remains macroscopically uniform, then $\partial\alpha^{ii}/\partial t$ could be nonzero throughout the material. Indeed, as described above, if there is TRS at t_i and t_f, and if the adiabatic approximation applies, then $\partial\alpha^{ii}/\partial t$ equals $\partial\alpha_{\mathrm{CS}}/\partial t$, which can only be nonzero if TRS is broken at intermediary times.

[20]This assumes nonpathological variation of the magnetization profile in time such that the jumps of $|\boldsymbol{P}^{(B)}(t)|$ do not resemble, for example, a Cantor ternary function. If it did, then even though $\boldsymbol{P}^{(B)}(t)$ is continuous almost everywhere (except possibly on a set of measure zero) the integral of the derivative of $\boldsymbol{P}^{(B)}(t)$ is not equal to the difference of $\boldsymbol{P}^{(B)}(t)$ evaluated at the endpoints.

---

[21]When the two exchange masses have the same sign, the model describes a quantum anomalous Hall insulator [57].

is specified by [59]

$$\mathcal{H}_{\text{eff}}(\boldsymbol{k}) = \begin{pmatrix} m & i\hbar v_{\text{D}} k_- & \Delta & 0 \\ -i\hbar v_{\text{D}} k_+ & -m & 0 & \Delta \\ \Delta & 0 & -m & -i\hbar v_{\text{D}} k_- \\ 0 & \Delta & i\hbar v_{\text{D}} k_+ & m \end{pmatrix},$$
(15)

where $k_\pm \equiv k_x \pm i k_y$.

To account for the presence of a uniform dc magnetic field that is parallel to the surface-normal of the thin film, we first implement the usual prescription[22] to obtain an envelope function approximated (EFA) Hamiltonian that is related to a given $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian. Following this, we minimally couple the electronic degrees of freedom to the corresponding vector potential. That is, first take $\hbar\boldsymbol{k} \to \mathfrak{p}(\boldsymbol{r}) \equiv -i\hbar\boldsymbol{\nabla}$ followed by $\mathfrak{p}(\boldsymbol{r}) \to \mathfrak{p}_{\text{mc}}(\boldsymbol{r}) = -i\hbar\boldsymbol{\nabla} + \frac{e}{c}\boldsymbol{A}(\boldsymbol{r})$, where $-e$ is the electronic charge. We take $\boldsymbol{B} = (0, 0, -B)$ and in the Landau gauge $\boldsymbol{A}(\boldsymbol{r}) = (0, -Bx, 0)$. The 2D EFA Hamiltonian that is obtained by employing this prescription to Eq. (15) is invariant under translations along $\boldsymbol{y}$ within the film. Thus, we seek energy eigenfunctions of the form

$$\Psi_{E,q_y}(x, y) = \frac{e^{iq_y y}\Phi_{E,q_y}(x)}{\sqrt{L_y}}.$$
(16)

Introducing the usual differential ladder operators $a \equiv \frac{1}{\sqrt{2}}(\tilde{x} + \frac{\partial}{\partial\tilde{x}})$ and $a^\dagger \equiv \frac{1}{\sqrt{2}}(\tilde{x} - \frac{\partial}{\partial\tilde{x}})$,[23] where $\tilde{x} \equiv l_B q_y - x/l_B$, $l_B^2 \equiv \hbar c/eB$, and $[a, a^\dagger] = 1$, the EFA Hamiltonian that is related to (15) acts on the $\Phi_{E,q_y}(x)$ via

$$H_{\text{EFA}}\left(x, \frac{\partial}{\partial x}; q_y\right) = \hbar\omega_c \tau_z \otimes \begin{pmatrix} 0 & a^\dagger \\ a & 0 \end{pmatrix}$$
$$+ m\tau_z \otimes \sigma_z + \Delta\tau_x \otimes \sigma_0,$$
(17)

where $\omega_c \equiv \sqrt{2}v_{\text{D}}/l_B$, τ_a are the Pauli matrices acting on the surface label (i.e., orbital type) degree of freedom, and σ_a are the Pauli matrices on spin degree of freedom.

A general eigenfunction of Eq. (17) has the form $\Phi_{E_n,q_y}(x) = \Phi_n(\tilde{x})$, where

$$\Phi_n(\tilde{x}) = \Big(C_{n,(\text{b},\uparrow)}\chi_n(\tilde{x}),\ C_{n,(\text{b},\downarrow)}\chi_{n-1}(\tilde{x}),$$
$$C_{n,(\text{t},\uparrow)}\chi_n(\tilde{x}),\ C_{n,(\text{t},\downarrow)}\chi_{n-1}(\tilde{x})\Big)$$
(18)

for integers $n > 0$ and

$$\Phi_0(\tilde{x}) = \big(C_{0,(\text{b},\uparrow)}\chi_0(\tilde{x}),\ 0,\ C_{0,(\text{t},\uparrow)}\chi_0(\tilde{x}),\ 0\big)$$
(19)

for $n = 0$, where $\chi_{n,q_y}(x) = \chi_n(\tilde{x})$ are normalized eigenfunctions of $a^\dagger a$. The corresponding eigenvalues are

$$E_n^\pm = \pm\sqrt{\hbar^2\omega_c^2 n + m^2 + \Delta^2},$$
(20)

with the $n = 0$ eigenvalues nondegenerate and the $n > 0$ eigenvalues each twofold degenerate. We adopt the notation $\Phi_0^\pm(\tilde{x})$ and $\Phi_n^{\pm,1}(\tilde{x})$, $\Phi_n^{\pm,2}(\tilde{x})$ for the corresponding eigenfunctions.

We endow the $\Psi_{n,q_y}^\pm(x, y)$ with z dependence by multiplying the components of (18) and (19) that are associated with the top (bottom) surface of the thin film by $\delta(z - z_{\text{t(b)}})$, where $z_{\text{t}} = a/2$ and $z_{\text{b}} = -a/2$. Then, taking $E_F = 0$ such that the ($\Psi_{n,q_y}^+(\boldsymbol{r})$) $\Psi_{n,q_y}^-(\boldsymbol{r})$ are (un)occupied, and using $\sum_{q_y}(1) = \sum_{q_y}(\int |\chi_n(\tilde{x})|^2 dx) = eBL_xL_y/hc$, the number of q_y per Landau level n, employing Eq. (8) yields

$$P_{\text{el}}^z = -\frac{e^2B}{2hc}\sum_{\substack{n\geqslant 0 \\ \sigma\in\{\uparrow,\downarrow\}}} \big(|C_{n,(\text{t},\sigma)}^-|^2 - |C_{n,(\text{b},\sigma)}^-|^2\big),$$
(21)

from which we can identify $\alpha_{\text{me}} \equiv \alpha^{zz}$ via $P_{\text{el}}^z = \alpha^{zz}B^z$.

The $n = 0$ anomalous Landau levels $\Phi_0^\pm(\tilde{x})$ are spin polarized, and the action of (17) on (19) simplifies such that $(C_{0,(\text{b},\uparrow)}^\pm, C_{0,(\text{t},\uparrow)}^\pm)$ is obtained by diagonalizing $m\tau_z + \Delta\tau_x$. At half filling, the occupied $n = 0$ eigenfunction is

$$\Phi_0^-(\tilde{x}) = \mathcal{N}\Big(-\Delta\chi_0(\tilde{x}),\ 0,\ (m + \sqrt{m^2 + \Delta^2})\chi_0(\tilde{x}),\ 0\Big),$$

where

$$\mathcal{N} = 1/\sqrt{2(m^2 + \Delta^2 + m\sqrt{m^2 + \Delta^2})}.$$

Using (21), the contribution to the magnetoelectric coefficient α_{me} from $\Phi_0^-(\tilde{x})$ is then

$$\alpha_{\text{me}}^{n=0} = \frac{e^2}{2hc}\frac{m}{\sqrt{m^2 + \Delta^2}}.$$
(22)

In the bulk limit of this toy model, the top and bottom surface states decouple and $\Delta \to 0$, in which case Eq. (22) yields the quantized magnetoelectric coefficient $e^2/2hc$ independent of m. Meanwhile in thin films with TRS, which implies $m = 0$, Eq. (22) gives that $\alpha_{\text{me}}^{n=0}$ vanishes.

Of course, the occupied $n > 0$ Landau levels may also contribute to (21). Unfortunately, the eigenvector-eigenvalue equations are not as simple as the $n = 0$ case. Moreover, due to the twofold degeneracy of each eigenvalue, the energy eigenvectors are nonunique. One convenient non-orthogonal pair is[24]

$$\big(C_{n,(\alpha,\sigma)}^{\pm,1}\big)_{\alpha,\sigma} = \mathcal{N}_n^{\pm,1}\Big(\frac{\Delta \pm \hbar\omega_c\sqrt{n}}{E_n^\pm - m}, \pm 1, 1, \frac{\pm\Delta - \hbar\omega_c\sqrt{n}}{E_n^\pm - m}\Big),$$

$$\big(C_{n,(\alpha,\sigma)}^{\pm,2}\big)_{\alpha,\sigma} = \mathcal{N}_n^{\pm,2}\Big(\pm 1, \frac{\Delta \pm \hbar\omega_c\sqrt{n}}{E_n^\pm + m}, \frac{\pm\Delta - \hbar\omega_c\sqrt{n}}{E_n^\pm + m}, 1\Big),$$

where $\mathcal{N}_n^{\pm,1}$ and $\mathcal{N}_n^{\pm,2}$ are normalization factors. These eigenvectors can be orthogonalized and their contributions to the induced polarization calculated. In the limit $\Delta \to 0$, which is of primary interest, these eigenvectors are indeed orthogonal and are moreover equally weighted combinations of states that are associated with the top and bottom surfaces. Thus, at half filling and in the limit $\Delta \to 0$, it is only the $\Phi_0^-(\tilde{x})$ anomalous Landau level that contributes to the magnetic-field-induced polarization.

_____

[22]See, e.g., Chap. 2 of Winkler [60].

[23]See, e.g., pp. 89–94 of Sakurai [61].

_____

[24]Here we denote $(C_{n,(\text{b},\uparrow)}, C_{n,(\text{b},\downarrow)}, C_{n,(\text{t},\uparrow)}, C_{n,(\text{t},\downarrow)})$ by $(C_{n,(\alpha,\sigma)})_{\alpha,\sigma}$.

B. Coupled Dirac cone model of layered thin films

The results for α_{me} obtained using the simple model described in Sec. III A can be confirmed by considering a more realistic model for finite thickness films of V_2VI_3-type TIs. Materials in this family consist of many stacked 2D layers. They are more accurately modeled by introducing a pair of Dirac cones for each layer, one associated with its top and bottom surface, then coupling the Dirac-like states within the same layer via Δ_S and coupling those associated with the closest surfaces of the nearest-neighbor layers via Δ_D [57,62]. This model has the merits that it is readily solved numerically for films that contain many van der Waals layers and that it is readily solved in the presence of a perpendicular external magnetic field, allowing the magnetic-field dependence of the electronic polarization to be evaluated explicitly.

The influence of a static and uniform magnetic field that is parallel to the surface-normal of the thin film is accounted for layer by layer; a 2D EFA Hamiltonian that is minimally coupled to $\boldsymbol{B}$ is assumed to describe the electronic dynamics within an isolated layer, and is taken as Eq. (17) with $\Delta \to \Delta_S$ and $m \to M_{l_z}$ for $l_z \in \{1, \dots, N\}$ a layer index. We then couple N copies (each indexed by an l_z) of this 2D isolated layer Hamiltonian via Δ_D as described above. We take $M_{l_z} \neq 0$ only if $l_z = 1$ or N to account for the presence of magnetic dopants only in the outermost top and bottom layer. As in Sec. III A, we take the orientation of the magnetization describing the configuration of dopants at the top and bottom surface to be opposite ($M_1 = -M_N \equiv m$).[25]

This model can be solved numerically for a varying number $N \geqslant 2$ of layers, which we always take to be even, and a general energy eigenfunction is again of the form [63] $\Psi_{n,q_y}^r(x, y) = e^{iq_y y}\Phi_n^r(\tilde{x})/\sqrt{L_y}$, where

$$\Phi_n^r(\tilde{x}) = \prod_{j \in \{0,\dots,2N-1\}} \left(C_{n,(j,\uparrow)}^r \chi_n(\tilde{x}), \; C_{n,(j,\downarrow)}^r \chi_{n-1}(\tilde{x}) \right) \quad (23)$$

for $n > 0$ and

$$\Phi_0^r(\tilde{x}) = \prod_{j \in \{0,\dots,2N-1\}} \left(C_{0,(j,\uparrow)}^r \chi_0(\tilde{x}), \; 0 \right) \quad (24)$$

for $n = 0$, and even (odd) values of j identify components that are associated with the bottom (top) surface of layer number $j/2$ [$(j-1)/2$] enumerated in ascending order along the stacking axis z. We again endow the $\Psi_{n,q_y}^r(x, y)$ with z dependence by taking $C_{n,(j,\uparrow)}^r \to C_{n,(j,\uparrow)}^r \delta(z - z_j)$; choosing the origin to be at the center of the material, for even values of j we have $z_j = a(j - N)/2$ and $z_{j+1} = z_j + a$ (see Fig. 1 of Lei *et al.* [57]). The electronic polarization (8) can then be written as

$$P_{\mathrm{el}}^z = -\frac{e^2 B}{aNhc} \sum_{j \in \{0,\dots,2N-1\}} \sum_{\substack{E_{n,r} < 0 \\ \sigma \in \{\uparrow,\downarrow\}}} z_j |C_{n,(j,\sigma)}^r|^2. \quad (25)$$

The results are shown in Fig. 3. In particular, as found in Sec. III A, if $m = 0$ then $\alpha_{\mathrm{me}} = 0$ any N, whereas if $m \neq 0$ then α_{me} approaches the expected quantized bulk value for

---

[25]In future studies of antiferromagnetic materials we consider $M_{l_z} = M_{l_z+1}$ for $l_z \in \{1, \dots, N-1\}$.
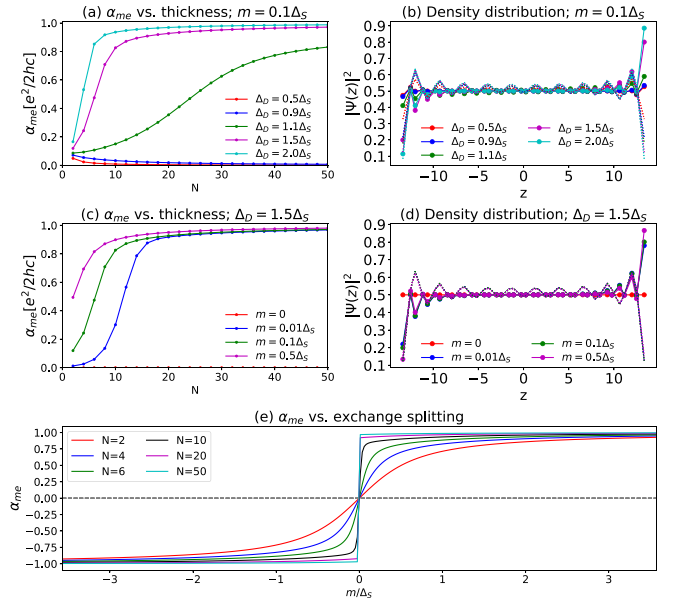

FIG. 3. Magnetoelectric coefficient α_{me} vs film thickness and the distribution of the electronic charge density in the stacking direction (z). (a) α_{me} vs film thickness for various Δ_D. Here the exchange splitting m of the Dirac cones associated with the top and bottom layers (resulting from the presence of surface magnetic dopants) is taken to be $0.1\Delta_S$ and taken to vanish for all other layers. (b) Distribution of the electronic charge density along the stacking direction for the same values of Δ_D used in panel (a), with the film thickness taken to be 20 layers and $z = 0$ taken to coincide with the middle of the center layer. The dotted (solid) lines identify the contribution of the $n = 1$ ($n = 0$) Landau levels to this distribution, which is (anti-)symmetric in z. (c) α_{me} vs film thickness for various exchange splittings m, with Δ_D taken as $1.5\Delta_S$. (d) Distribution of the electronic charge density in the stacking direction for the same values of m used in panel (c) for a 20-layer thin film. (e) α_{me} vs exchange splitting m due to magnetic surface dopants for films of varying thickness.

$N \to \infty$. That is, $\alpha_{\mathrm{me}} \to e^2/2hc$ for $|\Delta_D| > |\Delta_S|$ (in the $\mathbb{Z}_2$-odd regime of the model) and $\alpha_{\mathrm{me}} \to 0$ for $|\Delta_D| < |\Delta_S|$ (in the $\mathbb{Z}_2$-even regime). Moreover, numerical sums of the $n > 0$ ($n = 0$) energy eigenfunction distributions below the Fermi energy at half filling show that the charge density is (anti-)symmetric across the thin film, leading to a (non)vanishing contribution to the ground-state dipole moment and thus to α_{me}. This is shown for $n = 1$ ($n = 0$) by the dotted (solid) curves in Figs. 3(b) and 3(d).

The quantization of α_{me} in thick films can be understood by formally examining the model for an infinite number layers. In this case the analog of Eq. (17) is

$$H_{\mathrm{EFA}}\left(x, \frac{\partial}{\partial x}; q_y, k_z\right)$$
$$= \begin{pmatrix} 0 & \hbar\omega_c a^\dagger & \Delta_{k_z}^* & 0 \\ \hbar\omega_c a & 0 & 0 & \Delta_{k_z}^* \\ \Delta_{k_z} & 0 & 0 & -\hbar\omega_c a^\dagger \\ 0 & \Delta_{k_z} & -\hbar\omega_c a & 0 \end{pmatrix}, \quad (26)$$

where $\Delta_{k_z} \equiv \Delta_S + e^{ik_z a}\Delta_D$. To obtain Eq. (26) we have taken the basis states $|\bar{u}_{(\alpha,\uparrow),(k_x,k_y)}\rangle \to |\bar{u}_{(\alpha,\uparrow),(k_x,k_y,R)}\rangle$ for $R \in a\mathbb{Z}$

as hybrid WFs [17] (in $\mathbb{R}^3$) that are spatially localized in z about the surface α in the unit cell at Rz. In the nonmagnetic bulk there is one layer per unit cell and therefore $\alpha \in \{0, 1\}$. Again projecting into the $n = 0$ subspace, Eq. (26) acts on $(C^r_{0,(b,\uparrow)}, C^r_{0,(t,\uparrow)})$ via

$$H_{n=0}(k_z) = \begin{pmatrix} 0 & \Delta^*_{k_z} \\ \Delta_{k_z} & 0 \end{pmatrix}, \tag{27}$$

which is a Su-Schrieffer-Heeger (SSH) model with Δ_S and Δ_D playing the role of the hopping parameters. This model has particle-hole symmetry: $U^\dagger_{\mathcal{C}} H_{n=0}(-k_z)^* U_{\mathcal{C}} = -H_{n=0}(k_z)$ for $U_{\mathcal{C}} = \tau_z$ and therefore at half filling supports topologically distinct ground states characterized by a $\mathbb{Z}_2$-valued invariant,[26] which happens to be obtained by performing the BZ_{1D}-integral in the expression for the one-dimensional (1D) bulk electronic ground-state polarization [23,48]. Indeed, the $n = 0$ contribution to the bulk polarization P^z of Eq. (26) (which is related to adiabatically induced electronic currents in z[27]) is proportional to that 1D polarization and, in-line with the well-known results of the usual SSH model, at half filling we find

$$P^z_{n=0} \bmod \frac{e^2}{hc} B^z = \begin{cases} 0 & \text{if } |\Delta_S| > |\Delta_D| \\ \frac{e^2}{2hc} B^z & \text{if } |\Delta_S| < |\Delta_D|. \end{cases} \tag{28}$$

In contrast, the projected Hamiltonian in the $n > 0$ subspace is symmetric under 1D center-of-inversion (about the layer center) times spin-flip transformation and therefore does not support an electronic polarization [63].

IV. SEMI-ANALYTIC CALCULATION OF α_{CS} IN A 3D TIGHT-BINDING MODEL

In this section our primary aim is to calculate the Chern-Simons coefficient α_{CS} in bulk 3D nonmagnetic V_2VI_3-type insulators. Effective models for the electronic states near the Fermi energy in these materials have been developed [54,57,62], but since effective models are not generally defined over the entire BZ, they are inadequate for this purpose. The reason for this can be understood by noting that the topology of the vector bundle of occupied Bloch states over the Brillouin zone torus constructed for the true Hamiltonian $H(\mathbf{x}, \mathfrak{p}(\mathbf{x}))$ of a crystalline insulator manifests as an obstruction to the existence of a smooth *global* gauge thereof [16,45], but smooth local gauges always exist.[28] In particular, as explained previously, the linear-response expression for

α_{CS} as a BZ-integral of the Chern-Simons 3-form is valid *only* in a smooth global gauge of that bundle. We emphasize this issue in Appendix A, where we explicitly show that in a low-energy effective model an incorrect half quantization of α_{CS} can occur, which is reminiscent of the situation that arises when attempting to calculate the first Chern invariant of a 2D $\mathbf{k} \cdot \mathbf{p}$ model. Thus, lattice regularization of the previously developed effective models is required, and that is where we begin.

A. Construction of a 3D tight-binding model

In this subsection we construct a lattice regularization of a previously presented [62] effective model for the low-energy electronic states in bulk 3D V_2VI_3-type band insulators. Since a formally similar effective model describes antiferromagnetic $\mathbb{Z}_2$ TIs including $Mn(Sb_xBi_{1-x})_2X_4$ (where $X =$ Se, Te) [57], which will be the focus of future work, we in fact construct a sufficiently general lattice regularization scheme that can be applied to both families of materials. Notably, both families of materials can be viewed as layered compounds, with each layer having threefold rotational symmetry about the stacking (z) axis and twofold rotational symmetry about an axis perpendicular to the stacking axis [54,66]. The bulk crystal structure of the nonmagnetic (magnetic) family of materials consists of stacked five-atom (seven-atom) layers, called quintuple (septuple) layers, and has a center-of-inversion symmetry within each such layer [54,66]. That layered structure motivated the development of low-energy effective models [57,62] in which discrete 2D $\mathbf{k} \cdot \mathbf{p}$ continuum models in the planes perpendicular to the stacking axis are coupled to one another with interlayer hopping parameters Δ_D to yield a 1D tight-binding model along the stacking axis for each 2D momentum (see Appendix A); these models can be thought of as the 3D bulk limit of the quasi-2D Hamiltonian employed in Sec. III B. Since the electronic states within each layer are Dirac-like, effective Hamiltonians of this type are called coupled-Dirac cone models. We construct 3D tight-binding models to describe both families of materials by first developing a 2D lattice regularization of the $\mathbf{k} \cdot \mathbf{p}$ model of an isolated layer (obtained by taking $\Delta_D = 0$), then re-introduce Δ_D as in those past works.

1. A lattice regularized description of an isolated 2D layer

Two proposed low-energy effective models of 3D V_2 VI_3-type insulators include the above described coupled-Dirac cone model [57,62] and a traditional 3D $\mathbf{k} \cdot \mathbf{p}$ model [54]. When restricted to a single layer (i.e., the k_x-k_y plane), the 2D $\mathbf{k} \cdot \mathbf{p}$ models that result from these descriptions are unitarily equivalent to first order in k_x and k_y (when certain parameter values are taken to coincide), as one would expect. Furthermore, it has been observed [67] that the restriction of the 3D $\mathbf{k} \cdot \mathbf{p}$ model to the k_x-k_y plane is unitarily equivalent to the Bernevig-Hughes-Zhang (BHZ) model [68] for certain parameter values. Thus, when the coupled-Dirac cone model is employed for an isolated layer of a nonmagnetic material (i.e., $\Delta_D = 0$ and $M_{l_z} = 0$), it is unitarily equivalent to the BHZ model to first order in k_x and k_y. In fact, in this limit of the coupled-Dirac cone model, ex-

---

[26]See Sec. 2D of Ref. [10] and Ref. [64].

[27]See, e.g., Chap. 4 of Ref. [39] and Ref. [65].

[28]This follows immediately from the definition of the vector bundle of occupied states as a vector bundle [45,73].

act unitary equivalence with the BHZ model occurs if we generalize $\Delta_S \rightarrow \Delta_S - B(k_x^2 + k_y^2)$ and if Δ_S equals M of BHZ. Rather than generalizing $\Delta_S \rightarrow \Delta_S - B(k_x^2 + k_y^2)$, we could instead obtain exact unitary equivalence by restricting the BHZ model to first order in k_x and k_y by taking $B = 0$. However, doing so would result in the lattice regularization that we employ below to admit bandgap minima at points in the BZ_{2D} in addition to that at $(k_x, k_y) = (0, 0)$, in contrast with the known monolayer band structure [69]. Indeed, it will turn out that the parameter regime of immediate interest is $B/\Delta_S < 0$.

In their seminal work on 2D $\mathbb{Z}_2$ TIs [68], BHZ present a square lattice regularization of their $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian. Thus, we seek a square lattice regularization of Eq. (1) of Lei *et al.* [57] [generalized by taking $\Delta_S \rightarrow \Delta_S - B(k_x^2 + k_y^2)$] that, when restricted to the nonmagnetic case and applied to a single layer, is unitarily equivalent to Eq. (5) of BHZ [68]. In doing so, we generally consider Hamiltonian operators of the form

$$\hat{H}^{(d)} = \int_{BZ_d} \hat{c}_{\boldsymbol{k}}^{\dagger} \mathcal{H}^{(d)}(\boldsymbol{k}) \hat{c}_{\boldsymbol{k}} d^d k, \tag{29}$$

where d is the spatial dimension of the crystal, $\boldsymbol{k} \in BZ_d$ for BZ_d a d-dimensional first Brillouin zone of Γ_H^*—$\Gamma_H \subset \mathbb{R}^d$ is the Bravais lattice of the spatially periodic Bloch Hamiltonian under consideration and $\Gamma_H^* \subset \mathbb{R}^d$ its dual—and $\hat{c}_{\boldsymbol{k}} \equiv (\hat{c}_{(0,\uparrow),\boldsymbol{k}}, \hat{c}_{(0,\downarrow),\boldsymbol{k}}, \hat{c}_{(1,\uparrow),\boldsymbol{k}}, \hat{c}_{(1,\downarrow),\boldsymbol{k}})^{\mathrm{T}}$ and $\hat{c}_{\boldsymbol{k}}^{\dagger} \equiv (\hat{c}_{(0,\uparrow),\boldsymbol{k}}^{\dagger}, \hat{c}_{(0,\downarrow),\boldsymbol{k}}^{\dagger}, \hat{c}_{(1,\uparrow),\boldsymbol{k}}^{\dagger}, \hat{c}_{(1,\downarrow),\boldsymbol{k}}^{\dagger})$ are tuples of fermionic operators that act in the electronic Fock space. Products in Eq. (29) are the usual matrix multiplication. The fermionic operators involved in Eq. (29) are of the form $\hat{c}_{I,\boldsymbol{k}}, \hat{c}_{I,\boldsymbol{k}}^{\dagger}$, for I a general Wannier-type label. We require those operators to be such that the following is satisfied: $|\bar{\psi}_{I,\boldsymbol{k}}\rangle \equiv \hat{c}_{I,\boldsymbol{k}}^{\dagger}|\text{vac}\rangle$ are Bloch-type vectors[29] that are smooth over BZ_d, orthonormal such that $\langle \bar{\psi}_{I,\boldsymbol{k}} | \bar{\psi}_{J,\boldsymbol{k}'} \rangle = \delta_{I,J} \delta(\boldsymbol{k} - \boldsymbol{k}')$, and constitute a basis of the "relevant" electronic Hilbert space.[30] The penultimate and final criteria imply that $\{\hat{c}_{I,\boldsymbol{k}}, \hat{c}_{J,\boldsymbol{k}}^{\dagger}\} = \delta_{I,J} \delta(\boldsymbol{k} - \boldsymbol{k}')$, which

follows from the anticommutation of the electron field operators themselves. These are the usual criteria that are employed when constructing a tight-binding model and, for a given crystal, there are many sets of vectors $|\bar{\psi}_{I,\boldsymbol{k}}\rangle$ (and therefore operators $\hat{c}_{I,\boldsymbol{k}}, \hat{c}_{I,\boldsymbol{k}}^{\dagger}$) that satisfy them. An equivalent description can be obtained by noting that a set of Bloch-type vectors $|\bar{\psi}_{I,\boldsymbol{k}}\rangle$ that satisfy the above criteria bijectively maps to a set of exponentially localized Wannier functions (WFs) [17,46,70],

$$|W_{I,\boldsymbol{R}}\rangle = \sqrt{\frac{\Omega_{uc}}{(2\pi)^d}} \int_{BZ_d} d^d k \, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} |\bar{\psi}_{I,\boldsymbol{k}}\rangle, \tag{30}$$

where $\boldsymbol{R} \in \Gamma_H$. We emphasize that in order for the WFs $W_{I,\boldsymbol{R}}(\boldsymbol{r}) \equiv \langle \boldsymbol{r} | W_{I,\boldsymbol{R}} \rangle$ to be well localized, the $|\bar{\psi}_{I,\boldsymbol{k}}\rangle$ must be smooth over BZ_d [17]. Typically it is with respect to the fermionic operators $\hat{c}_{I,\boldsymbol{R}}, \hat{c}_{I,\boldsymbol{R}}^{\dagger}$ defined by $|W_{I,\boldsymbol{R}}\rangle \equiv \hat{c}_{I,\boldsymbol{R}}^{\dagger}|\text{vac}\rangle$ that a tight-binding Hamiltonian is specified.

In-line with typical tight-binding approaches, we further *assume* the existence of a Wannier-type label I of the form $I = (\alpha, \sigma)$, where $\alpha \in \{0, 1\}$ is an orbital index that identifies WFs of distinct spatial distribution and $\sigma \in \{\uparrow, \downarrow\}$ identifies the s_z spin component.[31] To coincide with those coupled-Dirac cone effective models [57,62] for which we develop this lattice regularization, we take $\alpha = 0$ (1) to label states that are associated with the bottom (top) surface of a layer; this identification will not explicitly enter calculations in this paper, but will provide physical motivation for the terms that appear in the 3D Hamiltonian of the following subsection. We further discuss (see preceding footnote) the topological implications related to the assumed existence of a Wannier-type label of the form $I = (\alpha, \sigma)$ in the following subsections, where we also address additional ubiquitous constraints that are imposed on the WFs (30).

A 2D low-energy effective model for the electronic states in a single layer of a MnV_2VI_4 magnetic (V_2VI_3 nonmag-

---

[29]We require that the position space representation $\bar{\psi}_{I,\boldsymbol{k}}(\boldsymbol{r}) \equiv \langle \boldsymbol{r} | \bar{\psi}_{I,\boldsymbol{k}} \rangle$ of each vector $|\bar{\psi}_{I,\boldsymbol{k}}\rangle$ to be of Bloch's form $\forall \boldsymbol{R} \in \Gamma_H$: $\bar{\psi}_{I,\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R}) = e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \bar{\psi}_{I,\boldsymbol{k}}(\boldsymbol{r})$. In general, any such $|\bar{\psi}_{I,\boldsymbol{k}}\rangle$ need not be an energy eigenvector.

[30]In general, all tight-binding constructions aim to model the electronic energy eigenvalues and eigenvectors within some "relevant" subspace of the full Hilbert space. This space can either be thought of as spanned by a set of "relevant" WFs or equivalently a set of "relevant" Bloch functions. Typically one aims to model those eigenvalues and vectors near the Fermi energy, usually starting with a representation in terms of the former (e.g., include only "nearest-neighbor hopping," etc.) then mapping to the latter before diagonalizing the model Hamiltonian.

[31]We simply include the spin degree of freedom as an additional component σ of the Wannier index I, and associate spin-up and spin-down components with the corresponding eigenvectors of s_z. In particular, we assume that the Hermitian bundle constructed using all relevant states can be decomposed as a product of Hermitian bundles corresponding with spin-up and spin-down sectors [71], and that each of these subbundles is topologically trivial [45,73] such that they each admit smooth global frames (see, e.g., Theorem 4.2.19 of Hamilton [72]). One instance of a smooth global frame for one such subbundle is our $(|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle)_{\alpha \in \{0,1\}}$ for a given σ. Therefore, the WFs $\{|W_{(\alpha,\sigma),\boldsymbol{R}}\rangle\}_{\alpha \in \{0,1\}}$ exist and can be obtained via Eq. (30). Since the eigenvectors of s_z are an orthonormal basis of the "spin space," the set of such WFs that result is orthonormal and complete.

netic) van der Waals semiconductor is obtained by taking $\Delta_D = 0$ (and $M_{l_z} = 0$) in Eq. (1) of Lei *et al.* [57]. A lattice

regularization thereof is motivated by BHZ [68], and is specified by taking $d = 2$ and the general $\mathcal{H}^{(d)}(\boldsymbol{k})$ of (29) to be

$$
\mathcal{H}_{\text{reg}}^{(2D)}(\boldsymbol{k}; l_z) = \begin{pmatrix} J_S M_{l_z} & iA(sk_x - isk_y) & \Delta_S(k_x, k_y) & 0 \\ -iA(sk_x + isk_y) & -J_S M_{l_z} & 0 & \Delta_S(k_x, k_y) \\ \Delta_S(k_x, k_y) & 0 & J_S M_{l_z} & -iA(sk_x - isk_y) \\ 0 & \Delta_S(k_x, k_y) & iA(sk_x + isk_y) & -J_S M_{l_z} \end{pmatrix}, \tag{31}
$$

where A is related to $\hbar v_D$ and fixed later by fitting the 2D band structure to the $\boldsymbol{k} \cdot \boldsymbol{p}$ dispersion near Γ and

$$
\Delta_S(k_x, k_y) \equiv \Delta_S - 2B(2 - ck_x - ck_y), \tag{32}
$$

where $sk_x \equiv \sin(k_x)$, $ck_x \equiv \cos(k_x)$, etc. We have also introduced a layer index $l_z \in \mathbb{Z}$ in anticipation of the generalization to 3D in the following subsection. This is a square lattice regularization since $\mathcal{H}_{\text{reg}}^{(2D)}(\boldsymbol{k} + \boldsymbol{G}; l_z) = \mathcal{H}_{\text{reg}}^{(2D)}(\boldsymbol{k}; l_z)$ for any $\boldsymbol{G} \in \Gamma_{2D}^* \equiv \text{span}_{\mathbb{Z}}(\{2\pi\boldsymbol{x}, 2\pi\boldsymbol{y}\})$, thus $\Gamma_{2D} = \text{span}_{\mathbb{Z}}(\{\boldsymbol{x}, \boldsymbol{y}\})$; the 2D lattice constant is taken to unity, and k_x, k_y are taken to be dimensionless. Of course, in the physical materials under consideration Γ_{2D} is a triangular lattice [54,66]. Thus, there likely exists a lattice regularization that better captures the lattice-scale physics, but for our purposes a square lattice regularization is sufficient.

In Eq. (31) the terms involving M_{l_z} act to generate an exchange mass within the layer l_z in the magnetic case, and arise from the exchange interactions of the dynamical electronic degrees of freedom deemed relevant with those well below the Fermi energy and thus approximated as static. We assume that this interaction is approximated by a Heisenberg interaction between dynamic spins and the quenched magnetic moments. We assume the direction of the static magnetization is parallel to the stacking axis (the $\boldsymbol{z}$ direction) [57]. The terms involving A and Δ_S can be understood as arising from spin-preserving electronic transitions between some relevant WFs of the layer that are associated with its top and bottom surfaces.

The eigenvalues of the single-layer Hamiltonian (31) demonstrate that it is indeed unitarily equivalent to that of BHZ [68] when our $M_{l_z} = 0$, their $C = D = 0$, and our Δ_S equals their M. If $M_{l_z} = 0$, then we similarly find that all of the energy bands of (31) are degenerate at Γ when $\Delta_S = 0$. Similar degeneracies appear at the other high-symmetry points $(k_x, k_y) = (\pi, 0)$ and $(0, \pi)$ when $\Delta_S - 4B = 0$, and at (π, π) when $\Delta_S - 8B = 0$. Indeed it was previously found [57] that the materials of immediate interest, for which the bandgap at half filling is about Γ, are described by $\Delta_S > 0$ and thus $B < 0$. At half filling, BHZ find that the zero-temperature ground state of their model is insulating and in a $\mathbb{Z}_2$-even (odd) phase for $M < 0$ $(0 < M/2B < 2)$ [68]; varying $M/2B$ across 0 (2) closes and re-opens the bandgap(s) at Γ (at $(\pi, 0)$ and $(0, \pi)$), and a topological phase transition occurs. However, this 2D topology is not directly related to that captured by α_{CS} in 3D, which will vanish in the limit of isolated layers independent of the parameters of a single-layer, even though the symmetries on which the classification is based are the same. Our focus in this paper is on the properties of the 3D crystals formed when these 2D layers are stacked and their states hybridized, not on the properties of individual layers.

2. A nonmagnetic 3D multi-layer Hamiltonian

We consider 3D models in which the electronic states associated with adjacent surfaces of nearest-neighbor layers are hybridized by a phenomenological interlayer hopping parameter Δ_D:

$$
\begin{aligned}
\hat{H}_{\text{reg}}^{(3D)} &= \sum_{l_z \in \mathbb{Z}} \hat{H}_{\text{reg}}^{(2D)}(l_z) \\
&+ \Delta_D \sum_{\substack{\boldsymbol{R} \in \Gamma_{2D} \times \{0\} \\ l_z \in \mathbb{Z} \\ \sigma \in \{\uparrow, \downarrow\}}} \left(\hat{c}_{(0,\sigma), \boldsymbol{R} + (l_z+1)a\boldsymbol{z}}^{\dagger} \hat{c}_{(1,\sigma), \boldsymbol{R} + l_z a\boldsymbol{z}} + \text{H.c.} \right),
\end{aligned} \tag{33}
$$

where we have assumed that all layers are identical apart from a layer-dependent M_{l_z}, that each layer is described by the 2D Hamiltonian (31), and that each pair of nearest-neighbor layers are separated by the same distance a in the $\boldsymbol{z}$ direction. In writing Eq. (33) we have assumed a nonmagnetic or ferromagnetic configuration of the static magnetic moments, in which case Δ_D is always identified with a hopping parameter between WFs associated with different unit cells. For general magnetic configurations there are two issues that need be addressed related to the precise definition of the WFs implicit in defining fermionic operators $\hat{c}_{(\alpha,\sigma), \boldsymbol{R} + l_z a\boldsymbol{z}}$, $\hat{c}_{(\alpha,\sigma), \boldsymbol{R} + l_z a\boldsymbol{z}}^{\dagger}$ that involve the layer label l_z. First, depending on $\{M_{l_z}\}_{l_z \in \mathbb{Z}}$, the group of translations Γ_H under which a single-particle 3D Bloch Hamiltonian $H(\boldsymbol{r}, \mathfrak{p}(\boldsymbol{r}))$—from which the tight-binding model might be obtained—is invariant may not be equal to the group of translations $\Gamma_{3D} \equiv \Gamma_{2D} \times a\mathbb{Z}$ under which the multilayer crystal lattice is invariant ($\Gamma_H \subseteq \Gamma_{3D}$). Then, although $\boldsymbol{R} + l_z a\boldsymbol{z} \in \Gamma_{3D}$ for all $l_z \in \mathbb{Z}$, generically $\boldsymbol{R} + l_z a\boldsymbol{z} \notin \Gamma_H$ and therefore *cannot* be used to label WFs or their corresponding operators. In the nonmagnetic ($M_{l_z} = 0$) and ferromagnetic ($M_{l_z} = M_{l_z'}$) cases, the magnetic and chemical unit cells coincide, $\Gamma_H = \Gamma_{3D}$, and $\boldsymbol{R} + l_z a\boldsymbol{z} \in \Gamma_H$ for all $l_z \in \mathbb{Z}$. However, in the antiferromagnetic case ($M_{l_z} = -M_{l_z \pm 1}$) this is not so and, strictly speaking, Eq. (33) requires modification. For completeness, we present the tight-binding Hamiltonian for the antiferromagnetic case in Appendix B and focus on the nonmagnetic case in the remainder of this paper. Second, a set of 3D WFs of a multilayer crystal [whose corresponding operators appear in Eq. (33)] is not generally equal to the set of all $l_z \boldsymbol{a}_3$-translated 2D WFs of an individual layer [whose corresponding operators appear in Eq. (31)] thereof. In principle, to construct an accurate 3D tight-binding model one indeed requires WFs of the 3D Bloch Hamiltonian $H(\boldsymbol{r}, \mathfrak{p}(\boldsymbol{r}))$. However, such details will not be relevant in our calculation of α_{CS} as we will not employ actual WFs but rather make simplifying approximations regarding their form. Consequently, this sec-
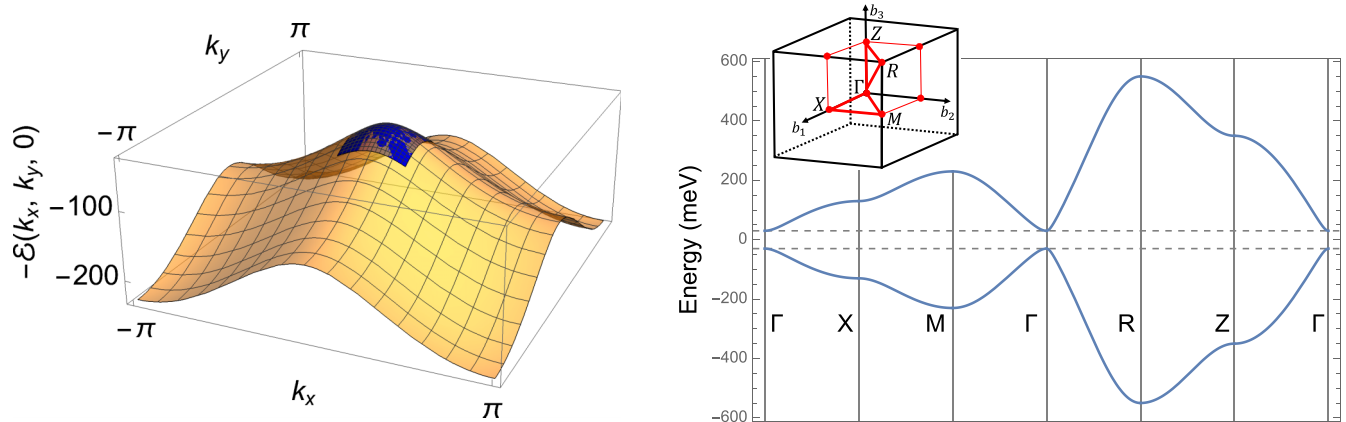
FIG. 4. Energy dispersion of the 3D tight-binding model of bulk V_2VI_3-type insulators for (somewhat arbitrary) parameters $A = 52.7$ meV, $B = -25$ meV, $\Delta_S = 190$ meV, $\Delta_D = -160$ meV. (a) The value of A was chosen to fit $-\varepsilon(\boldsymbol{k})$ (orange surface) to the in-plane dispersion of a single layer of Bi_2Te_3 near $(k_{\Gamma,x}, k_{\Gamma,y}) = (0, 0)$ (blue surface) when the regularizing lattice has a lattice constant of 1 nm. (b) Bulk 3D energy dispersion along the high-symmetry axes.

ond type of imprecision in writing Eq. (33) is inconsequential to our calculation. Meanwhile the line of reasoning leading to Eq. (33) yields a physically well-motivated form of $\hat{H}_{\text{reg}}^{(3D)}$.

Focusing on the nonmagnetic case ($M_{l_z} = 0$) such that Eq. (33) is valid, we use Eqs. (30) and (31) and

$$\frac{\Omega_{uc}}{(2\pi)^d} \sum_{\boldsymbol{R}\in\Gamma_H} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}} = \delta(\boldsymbol{k} - \boldsymbol{k}') \qquad (34)$$

for $d = 3$ and $\boldsymbol{k}, \boldsymbol{k}' \in BZ_{3D}$ to recast Eq. (33) as a BZ_{3D}-integral of the form (29). In this case the general $\mathcal{H}^{(d)}(\boldsymbol{k})$ is taken to be

$$\mathcal{H}_{\text{reg}}^{(3D)}(\boldsymbol{k}) = \begin{pmatrix} 0 & iA(sk_x - isk_y) & \Delta_{\boldsymbol{k}}^* & 0 \\ -iA(sk_x + isk_y) & 0 & 0 & \Delta_{\boldsymbol{k}}^* \\ \Delta_{\boldsymbol{k}} & 0 & 0 & -iA(sk_x - isk_y) \\ 0 & \Delta_{\boldsymbol{k}} & iA(sk_x + isk_y) & 0 \end{pmatrix}, \qquad (35)$$

and $\Delta_{\boldsymbol{k}} \equiv \Delta_S(k_x, k_y) + e^{iak_z}\Delta_D$. The eigenvalues of Eq. (35) are $E_{1,2}(\boldsymbol{k}) = -\varepsilon(\boldsymbol{k})$ and $E_{3,4}(\boldsymbol{k}) = \varepsilon(\boldsymbol{k})$ (see Fig. 4), where

$$\varepsilon(\boldsymbol{k}) \equiv \sqrt{\tfrac{1}{2}A^2(2 - c2k_x - c2k_y) + |\Delta_{\boldsymbol{k}}|^2}. \qquad (36)$$

Crucially, in this model there is a twofold degeneracy at each $\boldsymbol{k} \in BZ_{3D}$, which is a consequence[32] of a center-of-inversion symmetry and a (fermionic) time-reversal symmetry of (35); we explicitly demonstrate these symmetries in Appendix C. Consequently, eigenvectors of (35) are highly nonunique, in a more general sense than the usual $\boldsymbol{k}$-dependent phase ambiguity associated with Bloch's theorem. However, as described in Appendix A, the coupled-Dirac cone effective model that motivated (35) has a natural set of eigenvectors, and under the substitutions $\hbar v_D k_a \to Ask_a$ and $\Delta_S \to \Delta_S(k_x, k_y)$, those eigenvectors are related to a set of orthogonal eigenvectors of (35),

$$|\psi_{1,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(-\frac{\Delta_{\boldsymbol{k}}^*}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(0,\uparrow),\boldsymbol{k}}\rangle + |\bar{\psi}_{(1,\uparrow),\boldsymbol{k}}\rangle - \frac{iA(sk_x + isk_y)}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(1,\downarrow),\boldsymbol{k}}\rangle\right),$$

$$|\psi_{2,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(-\frac{iA(sk_x - isk_y)}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(0,\uparrow),\boldsymbol{k}}\rangle + |\bar{\psi}_{(0,\downarrow),\boldsymbol{k}}\rangle - \frac{\Delta_{\boldsymbol{k}}}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(1,\downarrow),\boldsymbol{k}}\rangle\right),$$

$$|\psi_{3,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(\frac{\Delta_{\boldsymbol{k}}^*}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(0,\uparrow),\boldsymbol{k}}\rangle + |\bar{\psi}_{(1,\uparrow),\boldsymbol{k}}\rangle + \frac{iA(sk_x + isk_y)}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(1,\downarrow),\boldsymbol{k}}\rangle\right),$$

$$|\psi_{4,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(\frac{iA(sk_x - isk_y)}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(0,\uparrow),\boldsymbol{k}}\rangle + |\bar{\psi}_{(0,\downarrow),\boldsymbol{k}}\rangle + \frac{\Delta_{\boldsymbol{k}}}{\varepsilon(\boldsymbol{k})}|\bar{\psi}_{(1,\downarrow),\boldsymbol{k}}\rangle\right), \qquad (37)$$

written here in the basis of Bloch-type vectors $|\bar{\psi}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ that are assumed smooth over BZ_{3D} [recall the discussion below Eq. (29)]. The energy eigenvectors in (37) are of

---

[32]See, e.g., Chap. 2 of Vanderbilt [39].

Bloch's form, satisfying $|\psi_{n,\mathbf{k}}\rangle = |\psi_{n,\mathbf{k}+\mathbf{G}}\rangle$ for any $\mathbf{G} \in \Gamma_{3D}^* = \text{span}_{\mathbb{Z}}(\{2\pi\mathbf{x}, 2\pi\mathbf{y}, \frac{2\pi}{a}\mathbf{z}\})$. When $\varepsilon(\mathbf{k}) > 0$ for all $\mathbf{k} \in \text{BZ}_{3D}$, i.e., when the model is a band insulator at half filling (the case of primary interest), these $|\psi_{n,\mathbf{k}}\rangle$ are smooth over BZ_{3D}. We can also identify from Eqs. (37) a Γ_{3D}^*-periodic unitary matrix $T(\mathbf{k})$ relating Wannier and energy Bloch-type vectors,

$$|\bar{\psi}_{(\alpha,\sigma),\mathbf{k}}\rangle = \sum_{n=1}^{4} |\psi_{n,\mathbf{k}}\rangle T_{n,(\alpha,\sigma)}(\mathbf{k}). \tag{38}$$

Since both $\bar{\psi}_{(\alpha,\sigma),\mathbf{k}}(\mathbf{r}) \equiv \langle\mathbf{r}|\bar{\psi}_{(\alpha,\sigma),\mathbf{k}}\rangle$ and $\psi_{n,\mathbf{k}}(\mathbf{r}) \equiv \langle\mathbf{r}|\psi_{n,\mathbf{k}}\rangle$ are of Bloch's form, they have associated Γ_{3D}-periodic functions $\bar{u}_{(\alpha,\sigma),\mathbf{k}}(\mathbf{r}) \equiv \langle\mathbf{r}|\bar{u}_{(\alpha,\sigma),\mathbf{k}}\rangle = (2\pi)^{3/2}e^{-i\mathbf{k}\cdot\mathbf{r}}\bar{\psi}_{(\alpha,\sigma),\mathbf{k}}(\mathbf{r})$ and $u_{n,\mathbf{k}}(\mathbf{r}) \equiv \langle\mathbf{r}|u_{n,\mathbf{k}}\rangle = (2\pi)^{3/2}e^{-i\mathbf{k}\cdot\mathbf{r}}\psi_{n,\mathbf{k}}(\mathbf{r})$, respectively, which are similarly related to one another by that $T(\mathbf{k})$ and are smooth (modulo $\exp(-i\mathbf{G}\cdot\mathbf{r})$ for all $\mathbf{G} \in \Gamma_H^*$) over BZ_{3D}.

The Bloch energy eigenvectors could alternately be chosen to simultaneously diagonalize the symmetry operators and the Hamiltonian itself. However, topological considerations forbid one from making that choice if the goal is to compute α_{CS} in a gauge defined by those energy eigenvectors, as we now describe.

Many of the surprising single-particle properties of the various types of topological insulators can be understood by studying the structure of a particular (complex) vector bundle [16,45,73], the Hermitian bundle of occupied Bloch states over the Brillouin-zone torus denoted $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \text{BZ}_d$. Band insulators are special in that the Hilbert space of occupied Bloch states at $\mathbf{k} \in \text{BZ}_d$, $\mathcal{V}_{\mathbf{k}} \equiv \text{span}_{\mathbb{C}}(\{|\psi_{n,\mathbf{k}'}\rangle : \mathbf{k}' = \mathbf{k} \text{ and } E_n(\mathbf{k}') < E_F\})$, has a dimension that is constant through BZ_d, thus $\mathcal{V}_{\mathbf{k}} \cong \mathcal{V}_{\mathbf{k}'}$ holds for all $\mathbf{k}, \mathbf{k}' \in \text{BZ}_d$. If there are N fully occupied energy bands then $\mathcal{V}_{\mathbf{k}} \cong \mathbb{C}^N$ for all $\mathbf{k} \in \text{BZ}_d$. The essential mathematical object is (isomorphic to) $\bigcup_{\mathbf{k}\in\text{BZ}_d}\{\mathbf{k}\} \times \mathcal{V}_{\mathbf{k}}$, which is the smooth manifold that is obtained by attaching to each $\mathbf{k} \in \text{BZ}_d$ the corresponding $\mathcal{V}_{\mathbf{k}}$ and equipping this set with a topology and smooth structure such that the natural projection map $(\mathbf{k}, |\psi_{\mathbf{k}}\rangle) \mapsto \mathbf{k}$ is smooth[33]. Although this construction is intuitive, it is convenient to instead consider an isomorphic bundle over BZ_d[34] with total space defined by

$$\mathcal{V} \equiv \left\{[\kappa, |u\rangle]_{\Gamma^*} \in (\mathbb{R}^d \times \mathcal{H})/\sim_{\Gamma^*} : |u\rangle \in \text{Ran}P_{\mathcal{V}}(\kappa)\right\},$$

where $\mathcal{H}$ is the Hilbert space of Γ-periodic functions, $P_{\mathcal{V}}(\kappa) = \sum_{n=1}^{N} |u_{n\kappa}\rangle \otimes \langle u_{n\kappa}|$ is the projector onto the space of Γ-periodic parts of occupied Bloch functions associated with $\kappa \in \mathbb{R}^d$, and the equivalence class $[\kappa, |u\rangle]_{\Gamma^*} \subset \mathbb{R}^d \times \mathcal{H}$ is defined by $(\kappa, |u\rangle) \sim_{\Gamma^*} (\kappa', |u'\rangle) : \iff \kappa' = \kappa + \mathbf{G}$ and $u'(\mathbf{r}) = e^{-i\mathbf{G}\cdot\mathbf{r}}u(\mathbf{r})$ for some $\mathbf{G} \in \Gamma^*$. The bundle projection map $\pi_{\mathcal{V}} : \mathcal{V} \to \text{BZ}_d$ is defined by $\pi_{\mathcal{V}}([\kappa, |u\rangle]_{\Gamma^*}) \equiv [\kappa]$. It has been proved [45,73] that in any band insulator $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \text{BZ}_d$ is a vector bundle (in particular, a Hilbert bundle) and is therefore locally trivial: for every $\mathbf{k} \in \text{BZ}_d$ there exists an open neighborhood $U \subseteq \text{BZ}_d$ of $\mathbf{k}$ over which $(\pi_{\mathcal{V}}^{-1}(U) \xrightarrow{\pi_{\mathcal{V}}} U) \cong (U \times \mathbb{C}^N \xrightarrow{\text{pr}_1} U)$ holds. In general, $\mathcal{V}$ need not be globally trivial; the above is not necessarily satisfied for $U = \text{BZ}_d$. However, it has been shown [45,46,73] that time-reversal symmetry implies $(\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \text{BZ}_d) \cong (\text{BZ}_d \times \mathbb{C}^N \xrightarrow{\text{pr}_1} \text{BZ}_d)$ is satisfied.[35]

In physics, the topology of $\mathcal{V}$ often manifests through the possible frames (also called gauge choices) thereof. A local frame of $\mathcal{V}$ over an open subset $U \subseteq \text{BZ}_d$ can be obtained from a collection of maps $\tilde{u}_i$ in $\mathscr{U} \subset \mathbb{R}^d$ (the complete set of representatives of U that is contained in the Wigner-Seitz cell of Γ^*) defined by $\tilde{u}_i(\kappa) \equiv (\kappa, |\tilde{u}_{i,\kappa}\rangle)$ that satisfies $\forall\kappa \in \mathscr{U} : (|\tilde{u}_{i,\kappa}\rangle)_{i\in\{1,...,N\}}$ is a basis of $\text{Ran}P_{\mathcal{V}}(\kappa)$ (i.e. the related Bloch-type functions $|\tilde{\psi}_{i,\mathbf{k}}\rangle$ are a basis of $\mathcal{V}_{\mathbf{k}}$ for $\mathbf{k} \in [\kappa]$); if the $|\tilde{u}_{i,\kappa}\rangle$ are energy eigenvectors then the gauge is called Hamiltonian. A general result[36] is that a vector bundle $E \xrightarrow{\pi} M$ is locally trivial over an open subset $U \subseteq M$ if and only if there exists a smooth frame of $\pi^{-1}(U) \xrightarrow{\pi} U$. Then, about every $\mathbf{k} \in \text{BZ}_d$ a smooth local frame of $\mathcal{V}$ always exists and if there is TRS then a smooth global frame of $\mathcal{V}$ exists. For example, if a 2D insulator is characterized by a nonzero first Chern invariant C, then $\mathcal{V}$ is not globally trivial; nonzero C acts as an obstruction to the existence of a smooth global gauge [45]. In this case, that $\mathcal{V}$ is not globally trivial can be understood to manifest in calculations though the fact that the Berry connection cannot be made smooth over BZ_{2D} and the integral expression for C then returns a nonzero value. The $\mathbb{Z}_2$ invariant relevant in this work is more subtle. TRS implies that $\mathcal{V}$ is globally trivial, but it has been shown [16,17,49,73] that a nontrivial $\mathbb{Z}_2$ invariant acts as an obstruction to the existence of a global gauge that is both smooth and time-reversal-symmetric. Thus, there need not exist a smooth global Hamiltonian frame of $\mathcal{V}$.[37] And since α_{CS} must be calculated with respect to a smooth global gauge of $\mathcal{V}$, in a $\mathbb{Z}_2$-odd insulator the components of that gauge are topologically forbidden to be energy eigenvectors that simultaneously diagonalize the time-reversal operator. Indeed, the generic nonexistence of a suitable Hamiltonian gauge

---

[33]This is a typical approach to construct a vector bundle given a collection of desired fibers, one associated with each point of the base manifold, and is employed in the textbook example of the tangent bundle (see, e.g., Chapter 3 of Lee [44]). However, in that prescription an additional choice is required, specifically a local frame (see *Local and Global Frames* in Chapter 10 of Lee [44]) that is desired to be smooth is identified and the result is that the natural projection map is smooth (Proposition 10.24 of Lee [44]). In the case of the tangent bundle a natural choice exists, but that is not so in our case. To make technical progress, researchers have found it convenient to identify the desired fibers via a family of projectors $P(\mathbf{k})$ that is associated with a particular set of energetically isolated bands since any such $P(\mathbf{k})$ is always smooth over the Brillouin zone torus.

[34]In this paper, as is typical in condensed matter physics, we use the symbol BZ_d to denote *both* the Wigner-Seitz cell of Γ^* and $\mathbb{R}^d/\sim_{\Gamma^*} \cong \mathbb{T}^d$ constructed using the equivalence relation for $\kappa, \kappa' \in \mathbb{R}^d$ that $\kappa \sim_{\Gamma^*} \kappa' : \iff \kappa' = \kappa + \mathbf{G}$ for some $\mathbf{G} \in \Gamma^*$.

[35]For example, when $d = 2$ this is equivalent to the vanishing of the first Chern invariant associated with $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \text{BZ}_{2D}$.

[36]See, e.g., Proposition 10.19 and Corollary 10.20 of Lee [44].

[37]If the energy bands are not degenerate anywhere in $\mathbf{k}$ then the energy eigenvectors necessarily diagonalize the symmetry operators and are not be continuous over BZ_d in a $\mathbb{Z}_2$-odd phase. And if there were degeneracies, but not over the entire BZ_d, then only linear combinations of the energy eigenvectors could be smooth in $\mathbf{k}$ [17], regardless of the $\mathbb{Z}_2$ classification of the ground state.

means that to compute α_{CS} requires [11] the employment of a Wannierization-like process [17] in order to construct Bloch-type functions $|\tilde{\psi}_{i,\boldsymbol{k}}\rangle$ that constitute a smooth global frame of $\mathcal{V}$. Within the model we employ, at half filling and in the case of a band insulator, time-reversal and inversion symmetry together imply that the two occupied Bloch states are globally degenerate. Any frame associated with these bands is therefore Hamiltonian; every smooth global gauge of this $\mathcal{V}$ is Hamiltonian.

3. Band number truncation in tight-binding models and crystal-momentum dependence of basis states

Before we can employ this model to compute α_{CS}, there is one more issue to address. Like most tight-binding models that appear in the literature, the Hamiltonian operator identified by Eq. (31) and (33) is not fully defined since the Bloch-type vectors $|\tilde{\psi}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ that correspond with the operators $\hat{c}_{(\alpha,\sigma),\boldsymbol{k}}$, $\hat{c}^\dagger_{(\alpha,\sigma),\boldsymbol{k}}$ are not completely specified.[38] This can be particularly problematic when calculating, for example, band-diagonal components of the Berry connection, in which case one must explicitly take $\boldsymbol{k}$-derivatives of the cell-periodic part of Bloch energy eigenvectors. (In $\boldsymbol{k} \cdot \boldsymbol{p}$ models this issue is avoided by construction, since these models are formulated with respect to a constant frame of the space of relevant Bloch states over the subset of BZ$_d$ for which the model is accurate [74,75].) In analytical calculations using tight-binding models it is almost always[39] assumed that the WFs (here the $|W_{(\alpha,\sigma),\boldsymbol{R}}\rangle$) with respect to which the model is specified are atomic-like, with negligible overlap between WFs that are associated with different unit cells. In this case, the variation of $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ with $\boldsymbol{k}$ is negligible (within the above defined equivalence classification), allowing it to remain unspecified. In this work we *assume* that the Hamiltonian operator (33) is written in a basis of Bloch-type functions $|\tilde{\psi}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ with corresponding $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ that are constant (modulo a phase) over BZ$_{3D}$, $\frac{\partial}{\partial k^a}|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle \equiv 0$ for $a = x$, y, or z ($\equiv 1$, 2, or 3).[40] While this may seem to be a drastic approximation, and in many cases it is indeed too drastic, we provide a physically motivated justification and argue that it is (at least) not mathematically inconsistent with the assumptions already inherent to our model.

A condition that one could potentially require of tight-binding models, in which the full electronic Hilbert space of a crystal is truncated to include only the states associated with a finite number of bands around the Fermi energy, is that the $\boldsymbol{k}$ dependence of the Bloch states be accurately captured

throughout the Brillouin zone. The local $\boldsymbol{k}$ dependence of the $|u_{n,\boldsymbol{k}}\rangle$ around some $\boldsymbol{k}_0$ can be calculated using $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, which yields

$$\frac{\partial}{\partial k^a}|u_{n,\boldsymbol{k}}\rangle = \frac{\hbar}{m_e} \sum_{m \neq n} \frac{\langle u_{n,\boldsymbol{k}_0}|p^a|u_{m,\boldsymbol{k}_0}\rangle}{E_{n,\boldsymbol{k}_0} - E_{m,\boldsymbol{k}_0}}|u_{m,\boldsymbol{k}_0}\rangle, \quad (39)$$

where m_e is the electron mass. It follows from Eq. (39) that a good criteria for the reliability of a tight-binding model is that all momentum matrix elements between the $|u_{m,\boldsymbol{k}_0}\rangle$ corresponding with included bands and neglected bands are small at all $\boldsymbol{k}_0$ of interest. When such a set of bands that are dissociated in this sense across the entire Brillouin zone can be identified,[41] we say that they satisfy a global isolation condition. When this global isolation condition is satisfied, we can always choose the finite-dimensional basis vectors of the included bands to be independent of $\boldsymbol{k}$, for example they could be the eigenstates at one particular $\boldsymbol{k}_0$. Our expectation is that global isolation holds only when it is implied by chemistry, i.e., the set of energy bands derives from a linear combination of atomic or molecular orbitals.

This global isolation condition is probably difficult to satisfy in practice and probably unnecessary for many physically relevant calculations, but it greatly simplifies calculations of Berry connections and related quantities. When we start from a phenomenological tight-binding model, as in this calculation, we know nothing about momentum matrix elements between included and neglected bands, so we have little choice but to assume the global isolation condition. It is probably common in *ab initio* DFT-derived models that global isolation is not satisfied. In particular, for $\mathbb{Z}_2$ TIs, in which level inversion at one point in the Brillouin zone plays the essential physical role, it will never be possible for the set of occupied bands to satisfy the global isolation condition.

To consider the mathematical implications of taking $\frac{\partial}{\partial k^a}|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle \equiv 0$ over BZ$_{3D}$, we return to the bundle-theoretic framework. In this paper we always consider the model at half filling. Thus, at any $\boldsymbol{k} \in$ BZ$_{3D}$, the $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ ($|\tilde{\psi}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$) need not be contained in Ran$P_\mathcal{V}(\boldsymbol{k})$ ($\mathcal{V}_{\boldsymbol{k}}$). Therefore, assuming the global isolation condition for (the space spanned by) the $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ does not constrain the topology of $\mathcal{V} \xrightarrow{\pi_\mathcal{V}}$ BZ$_{3D}$. However, this does have implications for the topology of the vector bundle of all relevant Bloch states, which we denote $\mathcal{B} \xrightarrow{\pi_\mathcal{B}}$ BZ$_{3D}$. The total space $\mathcal{B}$ of that bundle is constructed in a manner similar to $\mathcal{V}$, but now with the fiber at each $\boldsymbol{k} \in$ BZ$_{3D}$ isomorphic to $\mathcal{B}_{\boldsymbol{k}} \equiv \text{span}_{\mathbb{C}}(\{|\tilde{\psi}_{(\alpha,\sigma),\boldsymbol{k}'}\rangle : \boldsymbol{k}' = \boldsymbol{k}$ and $\alpha \in \{0, 1\}$, $\sigma \in \{\uparrow, \downarrow\}\})$. (That $\mathcal{B} \xrightarrow{\pi_\mathcal{B}}$ BZ$_{3D}$ is a vector bundle follows from an analogous argument [45,73] to $\mathcal{V} \xrightarrow{\pi_\mathcal{V}}$ BZ$_{3D}$, which can indeed be applied to any set of isolated energy bands). In constructing the model, we have already explicitly assumed, as one always does in employing tight-binding models, that $\mathcal{B} \xrightarrow{\pi_\mathcal{B}}$ BZ$_{3D}$ has a smooth global frame and is therefore globally trivial.

---

[38]For example, even in the case of graphene where it is said that the model is written with respect to p_z orbitals, one centered at the position of each ion core, this is not sufficient because p_z orbitals centered at different lattice sites are generally nonorthogonal if they have common support. So these at least need to be orthogonalized. In this sense there is an important distinction between mathematically consistent tight-binding Hamiltonians constructed via WFs and the physically motivated ones constructed from a linear combination of atomic orbitals.

[39]See, e.g., Sec. II.C.1 of Xiao *et al.* [40].

[40]In a related work Ref. [63] in which we consider the TME in antiferromagnetic TIs, this approximation requires further consideration.

---

[41]If the states associated with all of the energy bands are employed, then in any material a $\boldsymbol{k} \cdot \boldsymbol{p}$ over the entire Brillouin zone exists (see, e.g., Chap. 2 of Winkler [60]) since the set $\{|u_{m,\boldsymbol{k}_0}\rangle\}_{n \in \mathbb{Z}}$ at any $\boldsymbol{k}_0$ span the space of Γ-periodic functions [76]. In the application of $\boldsymbol{k} \cdot \boldsymbol{p}$ theory it is typically assumed (often accurately; see, e.g., Cardona *et al.* [77]) that the truncation of that basis is possible.

The assumption that we can label the WFs (30) by $I = (\alpha, \sigma)$ implies the existence of a smooth global frame of the form $(|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle)_{\alpha\in\{0,1\},\sigma\in\{\uparrow,\downarrow\}}$. In particular, this means that $\mathcal{B} \xrightarrow{\pi_{\mathcal{B}}} \text{BZ}_{3D}$ can be decomposed as a product of vector bundles that correspond with spin-up ($\sigma = \uparrow$) and spin-down ($\sigma = \downarrow$) sectors,[42] and that each sector admits a smooth global frame $(|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle)_{\alpha\in\{0,1\}}$ for a given σ. Thus the triple of Chern numbers (since $d = 3$) that characterizes each sector vanishes and $\mathcal{B} \xrightarrow{\pi_{\mathcal{B}}} \text{BZ}_{3D}$ is characterized by a trivial $\mathbb{Z}_2$ invariant [78,79]. Therefore, the topology of $\mathcal{B} \xrightarrow{\pi_{\mathcal{B}}} \text{BZ}_{3D}$ does not forbid the existence of a smooth and symmetric global frame. In addition, if we assume the global isolation condition then for any $\boldsymbol{k}, \boldsymbol{k}' \in \text{BZ}_{3D}$ we have $\text{Ran}P_{\mathcal{B}}(\boldsymbol{k}) = \text{Ran}P_{\mathcal{B}}(\boldsymbol{k}')$. Then, while we do not make a general existence argument for a frame with components $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ that are constant over BZ_{3D}, assuming one to exist is (at least) not generically forbidden by the topology of $\mathcal{B} \xrightarrow{\pi_{\mathcal{B}}} \text{BZ}_{3D}$. This argument applies to any tight-binding model that is specified with respect to WFs whose type labels are assumed to involve a spin index and satisfies the global isolation condition.

B. Calculation of α_{CS}

As mentioned above, in order to compute α_{CS} as the BZ_{3D}-integral of the Chern-Simons 3-form, which is valid

strictly for band insulators with corresponding $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \text{BZ}_{3D}$ that is globally trivial, one is required to work in a smooth global gauge of $\mathcal{V}$. Since Bloch energy eigenvectors are not generically smooth over BZ_{3D}, particularly in $\mathbb{Z}_2$ TIs, a Hamiltonian gauge choice is not typically viable. In past works [11,15,30,38], a gauge of that type is induced by the choice of WFs, where WFs are there assumed to be constructed from the (un)occupied energy eigenvectors alone. However, the WFs employed to construct tight-binding models are not typically of this type. In particular, the gauge $(|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle)_{\alpha\in\{0,1\},\sigma\in\{\uparrow,\downarrow\}}$ is not valid for this calculation.

As described in Sec. IV A 2, the model we employ has the positive feature that we are guaranteed a smooth global Hamiltonian gauge due to the combination of TRS and each set of isolated bands being completely degenerate over the entire BZ_{3D}. Using the $|u_{n,\boldsymbol{k}}\rangle$ related to (37), we can define a suitable frame $\mathfrak{u}$ of $\mathcal{B} \xrightarrow{\pi_{\mathcal{B}}} \text{BZ}_{3D}$ pointwise by $\mathfrak{u}_{\boldsymbol{k}} \equiv (|u_{n,\boldsymbol{k}}\rangle)_{n\in\{1,2,3,4\}}$; although $\mathfrak{u}$ is technically a frame of $\mathcal{B}$, since it is Hamiltonian it projects to a frame of $\mathcal{V}$. This gauge choice corresponds to taking the bulk electronic polarization and orbital magnetization to be defined with respect to the set of WFs $|W_{n,\boldsymbol{R}}\rangle = [\Omega_{uc}/(2\pi)^3]^{1/2} \int_{\text{BZ}_{3D}} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}|\psi_{n,\boldsymbol{k}}\rangle d^3k$. It has been shown [10,11,15,30,38] that, with respect to any smooth global Hamiltonian gauge $\mathfrak{u}$ defined by $\mathfrak{u}_{\boldsymbol{k}} \equiv (|u_{n,\boldsymbol{k}}\rangle)_{n\in\mathbb{N}}$,

$$\alpha_{\text{CS}}^{\mathfrak{u}} = -\frac{e^2}{2\hbar c}\epsilon^{abd}\int_{\text{BZ}_{3D}}\frac{d^3k}{(2\pi)^3}\left(\sum_{vv'}\xi_{vv'}^a(\boldsymbol{k})\frac{\partial}{\partial k^b}\xi_{v'v}^d(\boldsymbol{k}) - \frac{2i}{3}\sum_{vv'v_1}\xi_{vv'}^a(\boldsymbol{k})\xi_{v'v_1}^b(\boldsymbol{k})\xi_{v_1v}^d(\boldsymbol{k})\right), \tag{40}$$

where the sums are over the initially occupied band indices (here $v, v', v_1 \in \{1, 2\}$) and ξ_{nm}^a are the components of the non-Abelian Berry connection induced by $\mathfrak{u}$,

$$\xi_{nm}^a(\boldsymbol{k}) \equiv i\langle u_{n,\boldsymbol{k}}|\frac{\partial}{\partial k^a}u_{m,\boldsymbol{k}}\rangle. \tag{41}$$

Although Eq. (40) is gauge dependent, transforming from one appropriate gauge of $\mathcal{V}$ to another, be it a Hamiltonian gauge or otherwise, can only change its value by an integer multiple of e^2/hc [10,11]; that is, there is a quantum of indeterminacy associated with α_{CS}.

We now calculate $\alpha_{\text{CS}}^{\mathfrak{u}}$ via Eq. (40). Assuming that $\frac{\partial}{\partial k^a}|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle \equiv 0$ (see Sec. IV A 3), Eq. (41) becomes

$$\xi_{nm}^a(\boldsymbol{k}) = -\mathcal{T}_{nm}^a(\boldsymbol{k}), \tag{42}$$

where we have used $|u_{n,\boldsymbol{k}}\rangle = \sum_{\alpha,\sigma} |\bar{u}_{(\alpha,\sigma),\boldsymbol{k}}\rangle T_{(\alpha,\sigma),n}^\dagger(\boldsymbol{k})$ (obtained from (38) for $\alpha \in \{1, 2\}$ and $\sigma \in \{\uparrow, \downarrow\}$) and defined the Hermitian matrix $\mathcal{T}^a$ populated by elements

$$\mathcal{T}_{nm}^a(\boldsymbol{k}) \equiv i\sum_{\alpha,\sigma}\left(\frac{\partial}{\partial k^a}T_{n,(\alpha,\sigma)}(\boldsymbol{k})\right)T_{(\alpha,\sigma),m}^\dagger(\boldsymbol{k}). \tag{43}$$

Using Eqs. (37) and (38) in Eq. (43), we obtain

$$\mathcal{T}_{11}^x(\boldsymbol{k}) = -\frac{A^2\text{c}k_x\text{s}k_y + 2B\Delta_{\text{D}}\text{s}k_x\text{s}ak_z}{2\varepsilon(\boldsymbol{k})^2}, \quad \mathcal{T}_{11}^y(\boldsymbol{k}) = \frac{A^2\text{s}k_x\text{c}k_y - 2B\Delta_{\text{D}}\text{s}k_y\text{s}ak_z}{2\varepsilon(\boldsymbol{k})^2}, \quad \mathcal{T}_{11}^z(\boldsymbol{k}) = -\frac{a\Delta_{\text{D}}[\Delta_{\text{D}} + \Delta_{\text{S}}(k_x, k_y)\text{c}ak_z]}{2\varepsilon(\boldsymbol{k})^2},$$

---

[42]To construct these bundles, the prescription of Prodan [71] could be followed. Here, however, we are interested in spin sectors of both the total Bloch bundle and its valence subbundle. In the language of Prodan, the former corresponds with the projector being the identity map and the latter with the projector mapping to the occupied states of the electronic ground state.

$$\mathcal{T}_{12}^x(\boldsymbol{k}) = \frac{A[2B + ck_x(e^{iak_z}\Delta_D + \Delta_S - 4B + 2Bck_y) - 2iBsk_xsk_y]}{2\varepsilon(\boldsymbol{k})^2},$$

$$\mathcal{T}_{12}^y(\boldsymbol{k}) = -\frac{iA[2B + ck_y(e^{iak_z}\Delta_D + \Delta_S - 4B + 2Bck_x) + 2iBsk_xsk_y]}{2\varepsilon(\boldsymbol{k})^2}, \quad \mathcal{T}_{12}^z(\boldsymbol{k}) = -\frac{aAe^{iak_z}\Delta_D(isk_x + sk_y)}{2\varepsilon(\boldsymbol{k})^2}. \tag{44}$$

All other components are related to those given above. Note in particular that $\mathcal{T}^a$ is Hermitian, $\mathcal{T}_{nm}^a = (\mathcal{T}_{mn}^a)^*$, and that by explicit calculation it can be shown that

$$\mathcal{T}_{11}^a = \mathcal{T}_{33}^a, \quad \mathcal{T}_{22}^a = \mathcal{T}_{44}^a, \quad \mathcal{T}_{11}^a = -\mathcal{T}_{22}^a,$$

$$\mathcal{T}_{13}^a = -\mathcal{T}_{11}^a, \quad \mathcal{T}_{24}^a = -\mathcal{T}_{22}^a, \quad \mathcal{T}_{34}^a = \mathcal{T}_{12}^a, \quad \mathcal{T}_{14}^a = \mathcal{T}_{32}^a = -\mathcal{T}_{12}^a. \tag{45}$$

The relations (45) need not be satisfied for a different choice of energy eigenvectors. Using Eqs. (42), (44), and (45), the integrand of the first term of Eq. (40) is here found to be

$$\epsilon^{abc} \sum_{v,v'=1}^2 \xi_{vv'}^a \frac{\partial}{\partial k^b} \xi_{v'v}^c = -\frac{3aA^2\Delta_D\{\Delta_Dck_xck_y + [2B(ck_x + ck_y) + (\Delta_S - 4B)ck_xck_y]cak_z\}}{\varepsilon(\boldsymbol{k})^4}. \tag{46}$$

The integrand of the second term of Eq. (40) is (in this case) $-1/3$ of the first. Indeed, applying the general relation $\epsilon^{abc}\partial_b\xi_{nm}^c = i\epsilon^{abc}\sum_r \xi_{nr}^b\xi_{rm}^c$ (we now adopt the shorthand $\partial_a \equiv \partial/\partial k^a$) to this case (such that $r \in \{1, 2, 3, 4\}$) in combination with the relations (45), one can explicitly show that $\epsilon^{abc}\sum_{v,v'=1}^2 \xi_{vv'}^a \partial_b\xi_{v'v}^c = 12i\epsilon^{abc}\xi_{11}^a\xi_{12}^b\xi_{21}^c = 2i(\epsilon^{abc}\sum_{v,v',v_1=1}^2 \xi_{vv'}^a\xi_{v'v_1}^b\xi_{v_1v}^c)$. With this, we obtain the final expression

$$\alpha_{CS}^u = \frac{e^2}{\hbar c} \int_{-\pi/a}^{\pi/a} \frac{d(ak_z)}{2\pi} \int_{-\pi}^{\pi} \frac{dk_xdk_y}{(2\pi)^2} \frac{\Delta_D'\{\Delta_D'ck_xck_y + [2B'(ck_x + ck_y) + (\Delta_S' - 4B')ck_xck_y]cak_z\}}{\varepsilon'(\boldsymbol{k})^4}. \tag{47}$$

We have noted that α_{CS}^u is invariant under a change in energy scale and used this to remove the explicit dependence on A in Eq. (47) through the introduction of scaled parameters $B' \equiv B/A$, $\Delta_D' \equiv \Delta_D/A$ and $\Delta_S' \equiv \Delta_S/A$, and $\varepsilon'(\boldsymbol{k}) \equiv \varepsilon(\boldsymbol{k})/A$. In Table I we list values of α_{CS}^u obtained by extrapolating numerical estimates of the integrals over BZ_{3D} in Eq. (47) to convergence for various combinations of model parameters; we were not able to perform an analytical integration. We are careful to avoid sets of parameters for which the bandgap (at half filling) vanishes since Eq. (40) applies only to band insulators.

To obtain the topological phase diagram (Fig. 5) from the data in Table I, we use that the numerical value of the α_{CS}^u integral is piecewise constant in regions of the model's parameter space over which the bandgap is nonzero. That value can change when the parameters are varied in such a way that the bandgap vanishes at some point in BZ_{3D}. From Eq. (36) we see that this can happen only when $\varepsilon(\boldsymbol{k}) = 0$, that $\varepsilon(\boldsymbol{k})$ can vanish only at those $\boldsymbol{k} \in BZ_{3D}$ for which $k_x, k_y \in \{0, \pm\pi\}$ and $k_z \in \{0, \pm\pi/a\}$, and that the condition for bandgap clos-

TABLE I. Representative set of values of α_{CS}^u obtained from numerical integration of Eq. (47) for (arbitrarily chosen) $\Delta_S = 190$ meV. From these data the topological phase diagram (Fig. 5) can be deduced. When the values of the crystal parameters are varied such that the bandgap does not close, the value to which α_{CS}^u evaluates is unchanged.

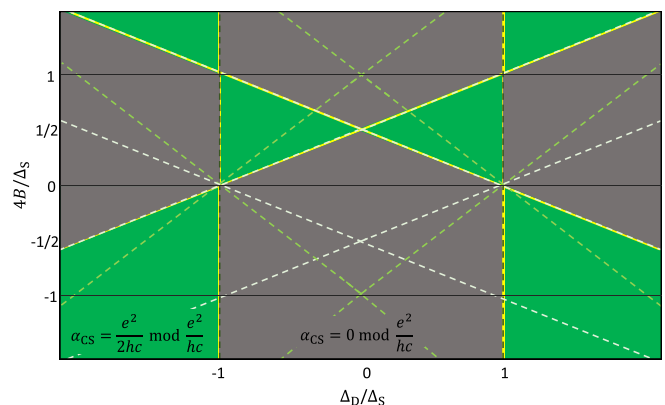| Δ_D/Δ_S	$4B/\Delta_S$	α_{CS}^u ($\frac{e^2}{hc}$)	Δ_D/Δ_S	$4B/\Delta_S$	α_{CS}^u ($\frac{e^2}{hc}$)
All	0	0	0	All	0
±0.1	0.5	1/2	±0.5	0.3	1/2
	-0.5	0		-0.3	0
±0.1	0.7	0	±0.8	0.3	$-1/2$
	-0.7	0		-0.3	0
±0.1	1.0	-1	±1.1	0.3	0
	-1.0	0		-0.3	1/2
±0.1	1.2	0	±1.4	0.3	0
	-1.2	0		-0.3	$-1/2$
±1.1	1.2	$-1/2$	±1.7	0.3	0
	-1.2	1/2		-0.3	0

FIG. 5. Phase diagram plotting the $\mathbb{Z}_2$ topological classification of the band insulating electronic ground state of $\hat{H}_{reg}^{(3D)}$ [specified in Eq. (35) and Sec. IV A 3] at half filling. The darkest (vertical) dashed lines identify the points in parameter space at which a bandgap closing occurs due to one of Eqs. (48a) and (49a) evaluating to zero, the medium dashed lines identify points at which one of Eqs. (48b) and (49b) evaluates to zero, and the lightest dashed lines identify points at which one of Eqs. (48c) and (49c) evaluates to zero.

ings depends only on the ratio of the scaled parameters. The end result is that bandgap closings occur along lines in (Δ'_D/Δ'_S)–(B'/Δ'_S) space [equivalently, (Δ_D/Δ_S)–(B/Δ_S) space]. Specifically, bandgaps close only when at least one of the following energies vanishes:

$$\varepsilon(0,0,0) = |\Delta_S + \Delta_D|, \tag{48a}$$

$$\varepsilon(\pi,0,0) = \varepsilon(0,\pi,0) = |\Delta_S + \Delta_D - 4B|, \tag{48b}$$

$$\varepsilon(\pi,\pi,0) = |\Delta_S + \Delta_D - 8B|, \tag{48c}$$

$$\varepsilon(0,0,\pi/a) = |\Delta_S - \Delta_D|, \tag{49a}$$

$$\varepsilon(\pi,0,\pi/a) = \varepsilon(0,\pi,\pi/a) = |\Delta_S - \Delta_D - 4B|, \tag{49b}$$

$$\varepsilon(\pi,\pi,\pi/a) = |\Delta_S - \Delta_D - 8B|. \tag{49c}$$

The topological phase diagram in (Δ_D/Δ_S)–(B/Δ_S) space is plotted in Fig. 5. If $\Delta_D/\Delta_S = \mp 1$ then from (48a) and (49a) $\varepsilon(\mathbf{k}) = 0$ at either $(k_x, k_y, k_z) = (0,0,0)$ or $(0,0,\pi/a)$. If $\Delta_D/\Delta_S = \mp(1 - 4B/\Delta_S)$ then from (48b) and (49b) $\varepsilon(\mathbf{k}) = 0$ at either $(\pi,0,0)$ and $(0,\pi,0)$ or $(\pi,0,\pi/a)$ and $(0,\pi,\pi/a)$. Finally, if $\Delta_D/\Delta_S = \mp(1 - 8B/\Delta_S)$ then from (48c) and (49c) $\varepsilon(\mathbf{k}) = 0$ at either $(\pi,\pi,0)$ or $(\pi,\pi,\pi/a)$. In Fig. 5 we use dashed lines to identify the sets of points in parameter space where at least one of the energies in (48) or (49) vanishes. In the region of parameter space that we consider, we find that a topological phase transition can occur only if a bandgap at $(0,0,0)$ or $(0,0,\pi/a)$ closes and re-opens, or one at $(\pi,\pi,0)$ or $(\pi,\pi,\pi/a)$ closes and re-opens [i.e., only along the lines defined by the vanishing of Eqs. (48a), (49a)) or (48c), (49c)]. When bandgaps at $(\pi,0,0)$ and $(0,\pi,0)$ or $(\pi,0,\pi/a)$ and $(0,\pi,\pi/a)$ close and re-open, the $\mathbb{Z}_2$ topological classification of $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \mathrm{BZ}_{3D}$ remains unchanged, but the value of $\alpha^{\mathrm{u}}_{\mathrm{CS}}$ can shift by an integer multiple of e^2/hc.

In Sec. III, the parameter B did not appear. This can be understood by recalling that the tight-binding model we employ was developed as a lattice regularization of an effective model that assumes that the low-energy states of the materials of interest are only near $(k_x, k_y) = (0,0)$; this is a property of our lattice model only when $B/\Delta_S < 0$ and $|\Delta_S| \approx |\Delta_D|$. In that case our expectation is that the largest contributions to the integrand of (47) come from near the line $(k_x, k_y) = (0,0)$, since $\varepsilon(\mathbf{k})$ is smallest there. Expanding our expression (47) for $\alpha^{\mathrm{u}}_{\mathrm{CS}}$ to first order in k_x and k_y about $(0,0)$, we find that the terms involving B cancel out and that

$$\alpha^{\mathrm{u}}_{\mathrm{CS}} \approx \frac{e^2}{\hbar c} \int_{-\pi/a}^{\pi/a} \frac{d(ak_z)}{2\pi} \int_{\mathbb{R}^2} \frac{dk_x dk_y}{(2\pi)^2} \frac{\Delta'_D(\Delta'_D + \Delta'_S cak_z)}{(k_x^2 + k_y^2 + |\Delta'_{k_z}|^2)^2}$$

$$= \begin{cases} 0 & \text{if } |\Delta_S| > |\Delta_D| \\ \frac{e^2}{2hc} & \text{if } |\Delta_S| < |\Delta_D|, \end{cases} \tag{50}$$

where we have artificially extended the domain of integration for k_x, k_y from the subset of BZ_{2D} near the expansion line to $\mathbb{R}^2$. For $B/\Delta_S < 0$ and $|\Delta_D| \approx |\Delta_S|$, Eq. (50) is consistent with Fig. 5. A similar approximate expression for α_{CS} was previously derived by Rosenberg and Franz [80] for models hosting Dirac-like low-energy states. We caution, however, that topological index calculations like this one, which focus only on contributions from regions near certain lines or points in $\mathbf{k}$ space, can fail. An explicit example is provided in Appendix A.

V. ATOMIC-LIKE AND ITINERANT CONTRIBUTIONS TO α_{CS}

In Sec. III we employed a coupled-Dirac cone model of a multilayer thin film—which had a vanishing magnetic exchange mass (i.e., no magnetic dopants) in the interior layers, but allowed a finite exchange mass in outermost layers—and found (see Fig. 3) that a perpendicular magnetic field can induce a charge density only near the surfaces, that this occurs only in the presence of local magnetic dopants, and that the corresponding magnetoelectric coefficient is quantized in sufficiently thick films only when the dopant configuration in the top and bottom layers is opposite. In Sec. II we argued from physical grounds that a meaningful bulk $\mathbf{P}^{(B)}$, particularly in nonmagnetic $\mathbb{Z}_2$ TIs, must (at least) involve an implicit consideration of surfaces to account for these magnetic-dopant requirements. We now ask whether there is a *purely* bulk manifestation of those requirements.

To begin, we recall the equivalence of the susceptibilities in Eq. (6), which is at least valid in bulk insulators for which $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \mathrm{BZ}_{3D}$ is globally trivial. We focus on the macroscopic electronic orbital magnetization $\mathbf{M}$ and recall from the modern theory [24,25] (or from other formulations of the microscopic response theory [27]) that

$$\mathbf{M} = \bar{\mathbf{M}} + \tilde{\mathbf{M}}, \tag{51}$$

where $\bar{\mathbf{M}}$ is called the atomic-like contribution and $\tilde{\mathbf{M}}$ the itinerant contribution. (In the modern theory these are typically denoted $\mathbf{M}_{\mathrm{LC}}$ and $\mathbf{M}_{\mathrm{IC}}$, respectively.) In an unperturbed bulk insulator, or when the orbital magnetization is induced by a uniform electric and/or magnetic Maxwell field, $\mathbf{M}$ is uniform and static, as are $\bar{\mathbf{M}}$ and $\tilde{\mathbf{M}}$ [31]. In the modern theory it has been argued [24] that in ferromagnetic ($\mathbf{M}^{(0)} \neq \mathbf{0}$) bulk insulators with isolated energy bands, interior and surface currents in finite samples thereof can be used to uniquely determine $\bar{\mathbf{M}}^{(0)}$ and $\tilde{\mathbf{M}}^{(0)}$, respectively;[43] we use the superscript (0) to label spontaneous orbital magnetizations and later use (E) to label magnetization that arises in linear response to the electric field. However, when there are degeneracies within the band structure, the decomposition of $\mathbf{M}^{(0)}$ as a sum of $\bar{\mathbf{M}}^{(0)}$ and $\tilde{\mathbf{M}}^{(0)}$ is gauge dependent [25]. In this case, interior and surface currents do not uniquely determine $\bar{\mathbf{M}}^{(0)}$ and $\tilde{\mathbf{M}}^{(0)}$; additional data related to the complete set of WFs being employed is required. Still, the partition into contributions associated with interior and surface currents may retain meaning. To investigate this possibility, we explicitly study the expressions for $\bar{\mathbf{M}}^{(E)}$ and $\tilde{\mathbf{M}}^{(E)}$, which, like their unperturbed counterparts, are individually gauge dependent, but their sum $\mathbf{M}^{(E)}$ is unique only modulo $e^2\mathbf{E}/hc$ [10,11].

In the approach of the modern theory, the expressions for the atomic-like $\bar{\mathbf{M}}^{(E)}$ and itinerant $\tilde{\mathbf{M}}^{(E)}$ contributions to $\mathbf{M}^{(E)}$ [Eq. (6) of Malashevich *et al.* [38]] are given in terms

---

[43]In the case of a topologically trivial bulk insulator, the expressions for the electronic polarization and orbital magnetization that result from the different approaches [24,25] and [27] agree. However, the proposed relation to interior and surface currents only arises in the former.

of states that implicitly involve the electric field $\boldsymbol{E}$. This is inconvenient for our purposes. As mentioned in Sec. II, there exists a microscopic approach [27], motivated by that of PZW, that identifies atomic-like and itinerant contributions to $\boldsymbol{M}$ by different means, but yields the same expressions in an unperturbed bulk insulator [30]. Indeed the microscopic approach reproduces the magnetoelectric susceptibility (9) derived via the modern theory. At linear response to a uniform dc electric field the expressions resulting from those distinct approaches are similar in that $\bar{\boldsymbol{M}}^{(E)}$ originates from the effect of $\boldsymbol{E}$ on the electronic dynamics directly [called dynamical modifications; see Sec. 3. Chap. 1 of Mahon and Sipe (MS) [30]], while $\tilde{\boldsymbol{M}}^{(E)}$ originates from a combination of dynamical modifications and changes in the form of the relevant matrix elements due to $\boldsymbol{E}$ (called compositional modifications). A familiar example of this is paramagnetic and diamagnetic response of an atom to an electromagnetic field,[44] which can be understood as arising from dynamical and compositional modifications, respectively. Below we employ the expressions derived by MS [30].

Working in the smooth global Hamiltonian gauge $\mathfrak{u}$ defined by the Bloch energy eigenvectors (37) and using the degeneracy of the energy bands [see Eq. (36)], Eq. (70) of MS for the atomic-like contribution evaluates to[45]

$$
\begin{aligned}
\bar{M}_{\mathfrak{u}}^{i(E)} = \frac{e^2}{2\hbar c}\epsilon^{iab} E^l \sum_{cv} \int_{\mathrm{BZ_{3D}}} \frac{d^3 k}{(2\pi)^3} \\
\times \left(-\sum_{v'} \mathrm{Re}\big[i\xi_{vv'}^a \xi_{v'c}^b \xi_{cv}^l \big] + \sum_{c'} \mathrm{Re}\big[i\xi_{cc'}^a \xi_{c'v}^b \xi_{vc}^l \big] \right) \\
= 0,
\end{aligned}
\tag{52}
$$

where $v, v' \in \{1, 2\}$ and $c, c' \in \{3, 4\}$. Here and elsewhere we often keep the $\boldsymbol{k}$ dependence of quantities implicit. In this gauge the atomic-like magnetization vanishes as an immediate result of the relations (45) between the band components of ξ^a. The itinerant contribution $\tilde{M}^{i(E)}$ has both a dynamical modification $\tilde{M}^{i(E;\mathrm{I})}$ and a compositional modification $\tilde{M}^{i(E;\mathrm{II})}$ [given by Eqs. (72) and (71) of MS, respectively]. In any smooth global Hamiltonian gauge $\mathfrak{u}'$ of our model, $\tilde{M}_{\mathfrak{u}'}^{i(E;\mathrm{I})} = 0$ since $E_{c,\boldsymbol{k}} + E_{v,\boldsymbol{k}}$ is constant over $\mathrm{BZ_{3D}}$. In the particular Hamiltonian gauge $\mathfrak{u}$ that we employ, Eq. (71) of

MS evaluates to

$$
\begin{aligned}
\tilde{M}_{\mathfrak{u}}^{i(E;\mathrm{II})} &= \frac{e^2}{2\hbar c}\epsilon^{iab} E^l \sum_{vm} \int_{\mathrm{BZ_{3D}}} \frac{d^3 k}{(2\pi)^3} \mathrm{Re}\big[\xi_{vm}^l \partial_b \xi_{mv}^a \big] \\
&= \alpha_{\mathrm{CS}}^{\mathfrak{u}} \delta^{il} E^l.
\end{aligned}
\tag{53}
$$

To reach the final equality in Eq. (53) (in our case, $m \in \{1, 2, 3, 4\}$) we have used (44) and (45) for ξ^a and recognized the result as the analytic expression (47) for $\alpha_{\mathrm{CS}}^{\mathfrak{u}}$. We find, therefore, that in the gauge $\mathfrak{u}$ the topological magnetoelectric response is entirely itinerant, i.e., that

$$
M_{\mathfrak{u}}^{(E)} = \tilde{M}_{\mathfrak{u}}^{i(E;\mathrm{II})}.
\tag{54}
$$

Now consider working in some other smooth global Hamiltonian gauge $\mathfrak{u}'$. (Recall the discussion at the end of Sec. IV A 2.) Then at each $\boldsymbol{k} \in \mathrm{BZ_{3D}}$ there exists a unitary matrix $U(\boldsymbol{k})$ relating the components $|u'_{n,\boldsymbol{k}}\rangle$ and $|u_{n,\boldsymbol{k}}\rangle$ of the gauges $\mathfrak{u}'$ and $\mathfrak{u}$,[46]

$$
|u'_{n,\boldsymbol{k}}\rangle = \sum_{m=1}^{4} |u_{m,\boldsymbol{k}}\rangle U_{m,n}(\boldsymbol{k}),
\tag{55}
$$

where $U_{c,v}(\boldsymbol{k}) \equiv 0$. The components of the non-Abelian Berry connection induced by $\mathfrak{u}'$ are

$$
(\xi_{\mathfrak{u}'})_{nm}^a(\boldsymbol{k}) = i\langle u'_{n,\boldsymbol{k}} | \partial_a u'_{m,\boldsymbol{k}} \rangle
$$

and are related to those induced by $\mathfrak{u}$ [denoted in Eq. (42) as $\xi_{nm}^a \equiv (\xi_{\mathfrak{u}})_{nm}^a$] by

$$
(\xi_{\mathfrak{u}'})_{nm}^a(\boldsymbol{k}) = \sum_{l,s=1}^{4} U_{n,l}^\dagger(\boldsymbol{k})\big[\xi_{ls}^a(\boldsymbol{k}) + \mathcal{W}_{ls}^a(\boldsymbol{k}) \big] U_{s,m}(\boldsymbol{k}),
\tag{56}
$$

where

$$
\mathcal{W}_{nm}^a(\boldsymbol{k}) \equiv i\sum_{r=1}^{4} [\partial_a U_{n,r}(\boldsymbol{k})] U_{r,m}^\dagger(\boldsymbol{k}).
\tag{57}
$$

Notably, $U_{c,v}(\boldsymbol{k}) \equiv 0$ implies $\mathcal{W}_{cv}^a(\boldsymbol{k}) \equiv 0$. Employing Eq. (56), we are able to relate the atomic-like and itinerant contributions to the orbital magnetization that are identified with respect to the distinct gauges $\mathfrak{u}$ and $\mathfrak{u}'$. In particular, using Eq. (52) we find

$$
\begin{aligned}
\bar{M}_{\mathfrak{u}'}^{i(E)} = \bar{M}_{\mathfrak{u}}^{i(E)} + \frac{e^2}{2\hbar c}\epsilon^{iab} E^l \int_{\mathrm{BZ_{3D}}} \frac{d^3 k}{(2\pi)^3} \\
\times \mathrm{Re}\left[i\sum_{cvv'} \mathcal{W}_{vv'}^b \xi_{v'c}^a \xi_{cv}^l - i\sum_{vcc'} \mathcal{W}_{cc'}^b \xi_{c'v}^a \xi_{vc}^l \right] \\
\equiv \bar{\alpha}_{\mathfrak{u}'}^{li} E^l.
\end{aligned}
\tag{58}
$$

In general, the only restriction on $U(\boldsymbol{k})$ is that it be smooth and Γ_H^* periodic, thus there is no reason to expect $\bar{\alpha}_{\mathfrak{u}'}^{li}$ to vanish.

As an example, consider a gauge transformation $U(\boldsymbol{k})$ such that $U_{n,m}(\boldsymbol{k}) \neq 0$ only if $n = m = 1$. Since $U(\boldsymbol{k})$ is unitary, there exists a smooth function $\lambda(\boldsymbol{k})$ such that $\forall~\boldsymbol{k} \in$

---

[44]See, e.g., Sec. 2.2 of Swiecicki [29].

[45]Working in a smooth and periodic Hamiltonian gauge, which, we reiterate, need not always exist, amounts to taking $U(\boldsymbol{k}) = \mathrm{id}_{4\times 4}$ (where $v \in \{1, 2\}$, $c \in \{3, 4\}$) and therefore $\mathcal{W}_{nm}^a(\boldsymbol{k}) \equiv 0$ in the expressions derived in Ref. [30,31]. In those works it was assumed the electronic ground state is such that there exists a Γ^*-periodic unitary matrix $U(\boldsymbol{k})$ that maps the (un)occupied Bloch energy eigenvectors $|\psi_{n,\boldsymbol{k}}\rangle$ to a set of smooth and orthogonal Bloch-type functions $|\tilde{\psi}_{\alpha,\boldsymbol{k}}\rangle$ that therefore live in the (un)occupied electronic Hilbert space. In the present work we find a set of smooth energy eigenvectors, so we can take that $U(\boldsymbol{k})$ as identity. That is, in this work the Bloch energy eigenvectors each map to a WF, with respect to which the bulk polarization and orbital magnetization are defined.

[46]Since the components of these gauges are mutually orthogonal pointwise over BZ, $\langle u'_{n,\boldsymbol{k}} | u'_{m,\boldsymbol{k}} \rangle = \delta_{nm}$ and $\langle u_{n,\boldsymbol{k}} | u_{m,\boldsymbol{k}} \rangle = \delta_{nm}$, $T(\boldsymbol{k})$ is unitary.

$\mathrm{BZ_{3D}}: \forall \ \boldsymbol{G} \in \Gamma_H^* : \lambda(\boldsymbol{k}+\boldsymbol{G}) = \lambda(\boldsymbol{k}) + 2\pi j$ for $j \in \mathbb{Z}$, and $U_{1,1}(\boldsymbol{k}) = e^{-i\lambda(\boldsymbol{k})}$. Taking for example $\lambda(\boldsymbol{k}) \equiv \sin(k_x)\cos(ak_z)$ results in $\bar{\alpha}_{\mathfrak{u}'}^{zx} \neq 0$ and taking $\lambda(\boldsymbol{k}) \equiv \cos(k_x)\sin(ak_z)$ results in $\bar{\alpha}_{\mathfrak{u}'}^{zz} \neq 0$. Then, since TRS implies that in any smooth global Hamiltonian gauge $\mathfrak{u}'$ we have $\boldsymbol{M}_{\mathfrak{u}'}^{(E)} = \bar{\boldsymbol{M}}_{\mathfrak{u}'}^{(E)} + \tilde{\boldsymbol{M}}_{\mathfrak{u}'}^{(E)} = \alpha_{\mathrm{CS}}^{\mathfrak{u}'}\boldsymbol{E}$, it is generally *not* the case that $\tilde{M}_{\mathfrak{u}'}^{i(E;\mathrm{II})}$ equals $\alpha_{\mathrm{CS}}^{\mathfrak{u}'}\delta^{il}E^l$.

The only feature that we *might* think is special about the nonmagnetic case considered here is that there exists a smooth global Hamiltonian gauge $\mathfrak{u}$ in which $\bar{\boldsymbol{M}}_{\mathfrak{u}}^{i(E)} = \boldsymbol{0}$ and $\tilde{\boldsymbol{M}}_{\mathfrak{u}}^{(E)} = \alpha_{\mathrm{CS}}^{\mathfrak{u}}\boldsymbol{E}$. In a related work [63] we investigate this further by considering the lattice regularization of the coupled-Dirac cone model for the $\mathrm{MnV_2VI_4}$ family of antiferromagnetic $\mathbb{Z}_2$ TIs. In that work we show that there *does not* generically exist a smooth global Hamiltonian gauge in which the topological magnetoelectric response is entirely itinerant. Indeed whether such a gauge exists is related to the geometry of $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \mathrm{BZ_{3D}}$ for the insulator under consideration—a smooth global gauge of a vector bundle is equivalent to a global trivialization.

Given that the topological magnetoelectric response in nonmagnetic $\mathbb{Z}_2$ TIs is (with important qualifications) itinerant, from these purely bulk considerations alone we might expect it to originate entirely from surface currents. In cases for which this is true, there is good reason to regard the identification of the physically meaningful bulk magnetoelectric response as α_{CS} with suspicion, especially when we know that any finite-size material sample will host gapless surface states whose role must also be considered. Surface magnetic dopants can generate gaps in the energy dispersion of the surface states and puts this identification on safer footing. To generate a gap everywhere on the surface, it is also required that the outward normal of the magnetization characterizing the dopant configuration never change sign. For a thin film this leads to the requirement that this magnetization retains a perpendicular component on the side walls and that it have opposite orientations on the top and bottom surfaces. Indeed these are the same magnetic-dopant requirements that were mentioned at the beginning of this section and it is only when they are satisfied that microscopic finite-size calculations agree with the simple interpretation of the bulk magnetoelectric linear-response calculation. When these requirement are not satisfied, there is no physically meaningful magnetoelectric response in nonmagnetic $\mathbb{Z}_2$ TIs.

VI. SUMMARY AND DISCUSSION

In this paper we study the topological magnetoelectric effect in 3D nonmagnetic $\mathbb{Z}_2$ topological insulators. We begin Sec. II with a summary of the considerations that underlie notions of polarization and (orbital) magnetization in material media, and of magnetoelectric response. In doing so, we illustrate the precarious relationship between that response and time-reversal symmetry, in nonmagnetic $\mathbb{Z}_2$ TIs in particular. Focused on materials of that type, we present a slightly modified interpretation of the argument [15] used to derive an expression for the bulk topological magnetoelectric coefficient, one that is manifestly consistent with time-reversal symmetry. We conclude that the polarization adjacent to a particular surface position is activated locally (see below) by time-reversal symmetry breaking at that position. This

interpretation is supported by explicit calculations in finite sized (Sec. III) and bulk (Sec. V) materials. In the process of formulating this interpretation, we clarify a partially inconsistent conclusion related to the adiabatic current expression from which the magnetoelectric response was previously [15] deduced. The central difference between our interpretation and the present one is that ours accounts for the fact that a quantized magnetoelectric effect occurs in nonmagnetic $\mathbb{Z}_2$ TIs only if static magnetic surface dopants are present and are configured such that their out-of-plane magnetization orientations at the top and bottom surfaces are nonzero and opposite. This resolves the tension between the seemingly contradictory roles played by TRS in finite sized and bulk insulators.

The model that we employ in Sec. III is for an infinite quasi-2D thin film with two perfectly flat surfaces on top and bottom. The explicit calculations on which our interpretation is based show that a magnetic-field dependent charge density can arise only at those surfaces and depends on the surface normal of the magnetization on that surface. This conclusion can be generalized to arbitrary surfaces by noting that only the sign of the surface normal magnetization component is important, and that the nonlocality length along lines where the surface normal changes sign [81], $\lambda = \hbar v_D/m$, is finite. As long as the typical length scales of surface regions with a fixed sign of the surface normal of the magnetization is large compared with λ, the properties we have calculated for uniform surfaces apply to arbitrary surface magnetization profiles, and the term local here refers to averages over finite regions of area λ^2. One copy of the infinite flat thin-film model in Sec. III can be associated with each locally flat spatial region of the sample. Within each such region, our interpretation of the adiabatic current calculation is that $\partial P^i/\partial B^i$ is constrained by bulk topology to have a quantized value. The value that is realized at a given position on a sample boundary is locally determined by the surface magnetization at the same position on the boundary [10]. The magnitude of the polarization response to magnetic field tends to be the smallest allowed quantized value, so changes from position to position tend to be changes in sign only. Sign changes in the surface magnetization profile will produce locally ungapped regions on the surface that will support [81] chiral edge states that separate polarization domains.

As we have explained, the topological magnetoelectric effect (TME) in our picture is a local surface charge-density property. If this is the meaning ascribed to the TME, then the quantum Hall effect that appears in the film as a whole when the magnetizations on top and bottom surfaces have the same orientation can be viewed to be a natural partner of the TME. This magnetic configuration implies opposite dependence of polarization on magnetic field on opposite surfaces of the sample, and via Eq. (4) to a total charge density in the insulator that varies linearly with magnetic field. A dependence of charge density on magnetic field in a quasi-2D insulator is equivalent to the quantum Hall effect via the Streda formula. (The relationship between the quantum Hall effect in TI thin films observed optically and the TME has been controversial [82–86].) When the magnetizations are antiparallel on the two surfaces of a thin film, the same argument implies that in our interpretation the field-dependent polarization inside the material is spatially uniform, so that the total charge density

is independent of magnetic field. Thermodynamic identities nevertheless imply that the orbital magnetization must vary with the electric field applied across the system, implying that the orbital magnetization is sensitive to an energy shift between top and bottom surfaces in the local density of states even though no chiral edge states are present at the Fermi energy, as argued in Ref. [51].

In Sec. V we consider whether these surface magnetic-dopant qualifications on the magnetoelectric response of nonmagnetic $\mathbb{Z}_2$ TIs manifest in the bulk expressions. If they did, one might guess to use the equivalence of the susceptibility tensors (6) and investigate the properties of the atomic-like $\bar{M}$ and itinerant $\tilde{M}$ contributions to the electric-field-induced magnetic dipole moment. Indeed $\tilde{M}$ was originally associated with surface currents [24]. The identification of $\bar{M}$ and $\tilde{M}$ is generally gauge dependent, even in unperturbed insulators [25], thus associating any physical meaning with that partitioning is suspect. Nevertheless, with respect to a particular smooth global Hamiltonian gauge we analytically demonstrate that the topological response $\delta^{il}\alpha_{\mathrm{CS}}$ arises entirely through that of $\tilde{M}$ (independent of a particular choice of parameter values, in all regions of parameter space for which there is always a bandgap). However, this is not true in a generic Hamiltonian gauge. In a related work [63] we consider antiferromagnetic (AFM) $\mathbb{Z}_2$ TIs that exhibit a generalized TRS. In that case we show that no gauge exists for which the response is entirely itinerant (in the above sense). In AFM materials, surface magnetic dopants are unnecessary for a magnetoelectric effect to manifest since the termination of the material already breaks the generalized TRS. Thus, it may be the case that the surface magnetic-dopant qualification in thin films is equivalent to the existence of a smooth global Hamiltonian gauge with respect to which the bulk response is entirely itinerant. Indeed, this may be understood as a type of microscopic bulk-boundary correspondence relating the bulk geometry of $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \mathrm{BZ}_{3\mathrm{D}}$ to the surface magnetic-dopant qualifications for the TME to manifest in a particular material. Although the nonexistence of certain smooth global gauges has been related [16] to the topology of $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \mathrm{BZ}_{3\mathrm{D}}$, we do not suspect this to be the case here; we anticipate that in an AFM TI there could exist some smooth global gauges $\mathfrak{u}$ for which there exists a finely tuned point in the parameter space where $\tilde{M}_{\mathfrak{u}}^{(E)}$ equals $\alpha_{\mathrm{CS}}^{\mathfrak{u}}E$. Thus, whether such a gauge generically exists is taken to be a geometric property of $\mathcal{V} \xrightarrow{\pi_{\mathcal{V}}} \mathrm{BZ}_{3\mathrm{D}}$.

That magnetic surface dopants activate the topological magnetoelectric effect in thin films of nonmagnetic $\mathbb{Z}_2$ TIs can be understood on physical grounds, as we now describe. Assume that the magnetoelectric coefficient in a bulk nonmagnetic $\mathbb{Z}_2$ TI is given by $\alpha_{\mathrm{CS}} = (n + 1/2)e^2/hc$ for some $n \in \mathbb{Z}$. We seek the corresponding susceptibility tensor $\alpha_{\mathrm{finite}}(\boldsymbol{r})$ in a finite sample thereof, which we assume can be partitioned into contributions from the sample's interior and surface regions, such that $\alpha_{\mathrm{finite}}(\boldsymbol{r}) = \alpha_{\mathrm{CS}}f_{\mathrm{int}}(\boldsymbol{r}) + \alpha_{\mathrm{surf}}(\boldsymbol{r})f_{\mathrm{surf}}(\boldsymbol{r})$; $\{f_{\mathrm{int}}(\boldsymbol{r}), f_{\mathrm{surf}}(\boldsymbol{r}), f_{\mathrm{ext}}(\boldsymbol{r})\}$ is a partition of unity over $\mathbb{R}^3$, where we take $f_{\mathrm{int}}(\boldsymbol{r})$ to have support over the entire interior region of the sample and decay from 1 to 0 on some length scale near its boundary, $f_{\mathrm{surf}}(\boldsymbol{r})$ to be nonvanishing only near the sample boundary, and $f_{\mathrm{ext}}(\boldsymbol{r})$ to identify the region outside of the material.

If we assume that bulk physics always determines the properties of finite-sized samples such that $\alpha_{\mathrm{surf}}(\boldsymbol{r})$ can be neglected, then $\alpha_{\mathrm{finite}}(\boldsymbol{r}) = \alpha_{\mathrm{CS}}f_{\mathrm{int}}(\boldsymbol{r})$. We now show that this assumption leads to a contradiction. Implementing this assumption in Eq. (4), we find[47]

$$J_{\mathrm{finite}}^{i(1)}(\boldsymbol{r}) = c\alpha_{\mathrm{CS}}\epsilon^{iab}\frac{\partial f_{\mathrm{int}}(\boldsymbol{r})}{\partial r^a}E^b(\boldsymbol{r}) + \cdots,$$

$$\varrho_{\mathrm{finite}}^{(1)}(\boldsymbol{r}) = -\alpha_{\mathrm{CS}}\frac{\partial f_{\mathrm{int}}(\boldsymbol{r})}{\partial r^a}B^a(\boldsymbol{r}) + \cdots.$$

That is, although in bulk insulators that exhibit TRS the linearly induced charge and current densities are insensitive to α_{CS}, physical consequence of α_{CS} can manifest at the surface of a finite sample thereof. To reach this conclusion we assumed that $\boldsymbol{J}_{\mathrm{finite}}^{(1)}(\boldsymbol{r})$ and $\varrho_{\mathrm{finite}}^{(1)}(\boldsymbol{r})$ are found by restricting the bulk $\boldsymbol{P}^{(1)}(\boldsymbol{r})$ and $\boldsymbol{M}^{(1)}(\boldsymbol{r})$ to a finite region of space, and not the bulk $\boldsymbol{J}^{(1)}(\boldsymbol{r})$ and $\varrho^{(1)}(\boldsymbol{r})$ themselves. For if it were the latter, then α_{CS} would again not contribute. This is strange—who decides which quantity should be restricted? It has also been pointed out[48] that the above expression for $\boldsymbol{J}^{(1)}(\boldsymbol{r})$ corresponds to a half-quantized quantum anomalous Hall current at each surface. But how can this be if the sample exhibits TRS? Something has gone wrong. Indeed, the result of Sec. III was that a magnetic-field-induced polarization can only occur in nonmagnetic TI thin films when TRS is broken by magnetic dopants at the surface, which is moreover consistent with the usual symmetry analysis. In particular, if $m = 0$ in Eq. (22) prior to taking the bulk limit, then $\alpha_{\mathrm{me}} = 0$.

Thus, the assumption that $\alpha_{\mathrm{surf}}(\boldsymbol{r})$ can be neglected leads to contradictions. This illustrates that, while it may be true that the bulk α_{CS} can manifest at the surface of finite-size samples, one must also consider the response of the surface states that exist as a consequence of the bulk topology. This conclusion is consistent with topological field theoretic considerations, where it has been demonstrated that the surface states implied by a nontrivial 3D bulk $\mathbb{Z}_2$ invariant (via the topological magnetoelectric coefficient α_{CS}) are related to the restoration of the parity anomaly and their response identically cancels that of the bulk [12–14]. This contrasts the situation for the integer quantum Hall effect related to a nontrivial 2D bulk Chern invariant, in which case the bulk and surface responses add [87–89].

Ultimately, a ubiquitous aspect of theories considering polarization and orbital magnetization in crystalline solids has been the importance of interpretation [39]. This work illustrates that in order to have a physically meaningful notion of bulk magnetoelectric response, in particular in insulators that exhibit TRS, the implicit involvement of a surface is unavoidable. Otherwise, as has been the case until now, the topological magnetoelectric response generally falls outside of the Peierls paradigm. Our hope is that this work will help to clarify misconceptions related to the topological magnetoelectric effect and thereby assist with its experimental

---

[47]Expressions that are similar to the following have been presented in previous works, including in Sec. IV of Qi *et al.* [10] and in Chap. 6.4 of Vanderbilt [39].

[48]See, e.g., Sec. IV of Qi *et al.* [10] or Chaps. 6.2.2 and 6.3.2 of Vanderbilt [39].

demonstration. We expect that a quantized TME will be observable in nonmagnetic TI thin films only when the surface normal of the magnetization does not change direction on either top or bottom surfaces and the magnetic gap induced at the surface is larger than the disorder potential on that surface. Changes in sign of the surface normal and reductions in local gap size on either surface will generically alter the global charge at which the density of states reaches a minimum and make it magnetic field dependent, swamping the topological magnetoelectric effect. It follows that the robustness against disorder that is critical to accurate observation of the integer quantum Hall effect, does not apply to the topological magnetoelectric effect.

APPENDIX A: FLAWED ATTEMPT TO CALCULATE α_{CS} VIA A LOW-ENERGY EFFECTIVE HAMILTONIAN

For illustrative purposes, in this section we perform a too-simplistic calculation of a low-energy approximation of α_{CS} using the bulk (i.e., periodic in z) version of the coupled-Dirac cone model introduced in Sec. III. In contrast with the quasi-2D models used in that section, this model is genuinely 3D. Notably, in the smooth local gauge that motivates the eigenvectors (37) of the tight-binding model obtained from a lattice regularization of this model, we find the same result (50). However, it is generally the case that one *does not* produce the correct result for α_{CS} using a low-energy effective model alone. Indeed if one somehow knows that the local gauge choice on the $\mathrm{BZ_{3D}}$ subset on which the effective model is valid can be extended to a smooth and periodic gauge over the entire $\mathrm{BZ_{3D}}$ of some lattice regularization thereof, then the correct quantization can result; this was used implicitly in past work [80]. Otherwise, incorrect half-quantization of the

susceptibility tensor can arise, similar to the situation when too-simplistic attempts are made to calculate the 2D Hall conductivity in a Chern insulator.

The effective low-energy Hamiltonian of 3D bulk nonmagnetic and magnetic multilayer TIs that was previously studied by Lei *et al.* [57] is essentially a $\boldsymbol{k} \cdot \boldsymbol{p}$ model about Γ in the $\boldsymbol{b}_1$-$\boldsymbol{b}_2$ ($\boldsymbol{x}$-$\boldsymbol{y}$) plane and a lattice model in $\boldsymbol{b}_3$ (z); by construction, the model is valid only on a subset of the 3D Brillouin zone. The model is written with respect to operators that generate basis vectors $|v_{(\alpha,\sigma),\boldsymbol{k}}\rangle$ that are products of 2D $\boldsymbol{k} \cdot \boldsymbol{p}$ basis vectors $|\bar{u}_{(\alpha,\sigma),\boldsymbol{k}_0}\rangle$, with the $\boldsymbol{k} \cdot \boldsymbol{p}$ expansion point $\boldsymbol{k}_0$ taken to be $(0,0)$, and 1D Bloch-type functions $\bar{\psi}_{(\alpha,\sigma),k_z}(z)$ that are associated with WFs $W_{(\alpha,\sigma),na_3}(z)$ that are localized in $\boldsymbol{a}_3 \parallel \boldsymbol{z}$ and taken to coincide with a bottom (top) layer "surface state" for $\alpha = $ even (odd). Then, in this model the operators generate vectors $|v_{(\alpha,\sigma),\boldsymbol{k}}\rangle = |v_{(\alpha,\sigma),(0,0,k_z)}\rangle \equiv |v_{(\alpha,\sigma),k_z}\rangle$ and, in fact, the possibility for variation in k_z of the associated 1D cell-periodic functions $\bar{u}_{(\alpha,\sigma),k_z}(z) \propto e^{-ik_z z}\bar{\psi}_{(\alpha,\sigma),k_z}(z)$ is neglected. In the nonmagnetic case, $\alpha \in \{0, 1\}$ and the effective Hamiltonian is defined by its matrix representation

$$\mathcal{H}_{\mathrm{eff}}^{(3D)}(\boldsymbol{k}) = \begin{pmatrix} 0 & \hbar v_{\mathrm{D}}ik_- & \Delta_{k_z}^* & 0 \\ -\hbar v_{\mathrm{D}}ik_+ & 0 & 0 & \Delta_{k_z}^* \\ \Delta_{k_z} & 0 & 0 & -\hbar v_{\mathrm{D}}ik_- \\ 0 & \Delta_{k_z} & \hbar v_{\mathrm{D}}ik_+ & 0 \end{pmatrix}$$

(A1)

in the basis $(|v_{(0,\uparrow),k_z}\rangle, |v_{(0,\downarrow),k_z}\rangle, |v_{(1,\uparrow),k_z}\rangle, |v_{(1,\downarrow),k_z}\rangle)$, where $k_\pm \equiv k_x \pm ik_y \equiv k_\perp e^{\pm i\theta_{k_\perp}}$, $\Delta_{k_z} \equiv \Delta_{\mathrm{S}} + e^{iak_z}\Delta_{\mathrm{D}}$; here k_x and k_y have units of 1/length. The relevant velocity scale v_{D} of the individual 2D layers is determined by the material parameters of one such layer. The eigenvalues of Eq. (A1) are $E_{1,2}(\boldsymbol{k}) = -\epsilon(\boldsymbol{k})$ and $E_{3,4}(\boldsymbol{k}) = \epsilon(\boldsymbol{k})$, where

$$\epsilon(\boldsymbol{k}) \equiv \sqrt{\hbar^2 v_{\mathrm{D}}^2(k_x^2 + k_y^2) + |\Delta_{k_z}|^2}. \quad \text{(A2)}$$

In this model, the condition of a band insulator at half filling is $\forall\, k_z \in [-\pi/a, \pi/a] : \Delta_{k_z} \neq 0$. Thus, as the ratio $|\Delta_{\mathrm{D}}|/|\Delta_{\mathrm{S}}|$ is varied, the bandgap can vanish; in particular, if $\Delta_{\mathrm{D}} = -\Delta_{\mathrm{S}}$ then $\epsilon(0,0,0) = 0$ and if $\Delta_{\mathrm{D}} = \Delta_{\mathrm{S}}$ then $\epsilon(0,0,\pi/a) = 0$. These are the points in parameter space at which a topological phase transition may occur.

Due to the double degeneracy (A2), at every $\boldsymbol{k}$ for which the model applies, the set of eigenvectors is highly nonunique. However, (A1) has natural solutions in the $ak_\perp \to 0, \infty$ limits $[k_\perp \equiv (k_x^2 + k_y^2)^{1/2}]$, which motivates us to make a particular choice,

$$|\phi_{1,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(-\frac{\Delta_{k_z}^*}{\epsilon(\boldsymbol{k})}|v_{(0,\uparrow),k_z}\rangle + |v_{(1,\uparrow),k_z}\rangle - \frac{i\hbar v_{\mathrm{D}}k_+}{\epsilon(\boldsymbol{k})}|v_{(1,\downarrow),k_z}\rangle\right),$$

$$|\phi_{2,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(-\frac{i\hbar v_{\mathrm{D}}k_-}{\epsilon(\boldsymbol{k})}|v_{(0,\uparrow),k_z}\rangle + |v_{(0,\downarrow),k_z}\rangle - \frac{\Delta_{k_z}}{\epsilon(\boldsymbol{k})}|v_{(1,\downarrow),k_z}\rangle\right),$$

$$|\phi_{3,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(\frac{\Delta_{k_z}^*}{\epsilon(\boldsymbol{k})}|v_{(0,\uparrow),k_z}\rangle + |v_{(1,\uparrow),k_z}\rangle + \frac{i\hbar v_{\mathrm{D}}k_+}{\epsilon(\boldsymbol{k})}|v_{(1,\downarrow),k_z}\rangle\right),$$

$$|\phi_{4,\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2}}\left(\frac{i\hbar v_{\mathrm{D}}k_-}{\epsilon(\boldsymbol{k})}|v_{(0,\uparrow),k_z}\rangle + |v_{(0,\downarrow),k_z}\rangle + \frac{\Delta_{k_z}}{\epsilon(\boldsymbol{k})}|v_{(1,\downarrow),k_z}\rangle\right). \quad \text{(A3)}$$

In the $ak_\perp \to 0$ limit, the Hamiltonian (A1) involves only interlayer transitions and the spin index becomes preserved, while in the $ak_\perp \to \infty$ limit the spin-orbit interaction dominates and the layer index becomes preserved; these limits are evident in the chosen form (A3). Notably, the set of vectors (A3) forms an orthonormal basis of the relevant electronic Hilbert space, which are smooth in $\boldsymbol{k}$ for an insulator [$\epsilon(\boldsymbol{k}) > 0$].

From (A3) we can identify a unitary matrix $\check{T}(\boldsymbol{k})$ analogous to that appearing in Eq. (38), here with elements $\check{T}_{n,(\alpha,\sigma)}(\boldsymbol{k}) = \langle \phi_{n,\boldsymbol{k}} | v_{(\alpha,\sigma),k_z} \rangle = \langle v_{(\alpha,\sigma),k_z} | \phi_{n,\boldsymbol{k}} \rangle^*$. With this we explicitly find the matrices $\check{\mathcal{T}}^a(\boldsymbol{k}) \equiv i[\partial_a \check{T}(\boldsymbol{k})]\check{T}^\dagger(\boldsymbol{k})$ (for $a = x, y,$ or z) analogous to those appearing in Eq. (43). Unsurprisingly, we find that these matrices satisfy relations analogous to (45), but now we find

$$\check{\mathcal{T}}_{11}^x(\boldsymbol{k}) = -\frac{\hbar^2 v_{\mathrm{D}}^2 k_y}{2\epsilon(\boldsymbol{k})^2}, \quad \check{\mathcal{T}}_{11}^y(\boldsymbol{k}) = \frac{\hbar^2 v_{\mathrm{D}}^2 k_x}{2\epsilon(\boldsymbol{k})^2}, \quad \check{\mathcal{T}}_{11}^z(\boldsymbol{k}) = -\frac{a\Delta_{\mathrm{D}}(\Delta_{\mathrm{D}} + \Delta_{\mathrm{S}} c a k_z)}{2\epsilon(\boldsymbol{k})^2},$$

$$\check{\mathcal{T}}_{12}^x(\boldsymbol{k}) = \frac{\hbar v_{\mathrm{D}} \Delta_{k_z}}{2\epsilon(\boldsymbol{k})^2}, \quad \check{\mathcal{T}}_{12}^y(\boldsymbol{k}) = -\frac{i\hbar v_{\mathrm{D}} \Delta_{k_z}}{2\epsilon(\boldsymbol{k})^2}, \quad \check{\mathcal{T}}_{12}^z(\boldsymbol{k}) = -\frac{\hbar v_{\mathrm{D}} a \Delta_{\mathrm{D}} e^{iak_z}(ik_x + k_y)}{2\epsilon(\boldsymbol{k})^2}. \tag{A4}$$

Indeed, the matrix elements (A4) are consistent with those (44) that are obtained from the lattice regularized model; taking $B = 0$ and k_x, k_y to be small, Eq. (44) reduces to Eq. (A4).

In principle, this model, like any $\boldsymbol{k} \cdot \boldsymbol{p}$ description, is insufficient to compute α_{CS}. That is, while low-energy models may accurately describe the physics of the states associated with those crystal momentum $\boldsymbol{k}$ near the $\boldsymbol{k} \cdot \boldsymbol{p}$ expansion point, the topology of $\mathcal{V}$ imposes constraints on global properties of Bloch energy eigenvectors and of smooth gauge choices over all of BZ_d; those constraints are generally not relevant on subsets of BZ_d. Nevertheless, it is insightful to calculate a low-energy approximation of α_{CS} to see where issues arise. We take the initial electronic state of the insulator to be the zero-temperature ground state and there to be two electrons per unit cell; that is, we take $|\phi_{1,\boldsymbol{k}}\rangle$ and $|\phi_{2,\boldsymbol{k}}\rangle$ to be initially occupied and $|\phi_{3,\boldsymbol{k}}\rangle$ and $|\phi_{4,\boldsymbol{k}}\rangle$ to be initially unoccupied. The components of the Berry connection induced by (A3) can be found via $\check{\xi}_{nm}^a(\boldsymbol{k}) = -\check{\mathcal{T}}_{nm}^a(\boldsymbol{k})$ [recall the discussion above Eq. (A1)]. After some algebra we find

$$\epsilon^{abc} \sum_{vv'} \check{\xi}_{vv'}^a \partial_b \check{\xi}_{v'v}^c = -3\frac{a\hbar^2 v_{\mathrm{D}}^2 \Delta_{\mathrm{D}}(\Delta_{\mathrm{D}} + \Delta_{\mathrm{S}} c a k_z)}{\epsilon(\boldsymbol{k})^4},$$

$$-\frac{2i}{3}\epsilon^{abc} \sum_{vv'v_1} \check{\xi}_{vv'}^a \check{\xi}_{v'v_1}^b \check{\xi}_{v_1 v}^c = \frac{a\hbar^2 v_{\mathrm{D}}^2 \Delta_{\mathrm{D}}(\Delta_{\mathrm{D}} + \Delta_{\mathrm{S}} c a k_z)}{\epsilon(\boldsymbol{k})^4}. \tag{A5}$$

To calculate quantities that are generally written as Brillouin-zone integrals [such as α_{CS} (40)] via a $\boldsymbol{k} \cdot \boldsymbol{p}$ description, one usually restricts the domain of integration to the subset of $\mathrm{BZ}_{3\mathrm{D}}$ on which the model is valid [here to a limited range of k_x, k_y about (0,0)] based on the assumption that outside of this region the integrand is negligible. In-line with this assumption, when k_x and/or k_y are large, the relevant integrand (A5) tends to zero; consequently, the domain of integration in the "$\boldsymbol{k} \cdot \boldsymbol{p}$ dimensions" is usually artificially extended to the entire plane and we do so here. With this, and moving to the cylindrical coordinates $(k_\perp, \theta_{k_\perp}, k_z)$, we have

$$\alpha_{\mathrm{CS}} \doteq \frac{e^2}{hc} \int_{-\pi/a}^{\pi/a} \frac{dk_z}{2\pi} \int_0^\infty dk_\perp k_\perp \frac{a\Delta_{\mathrm{D}}'(\Delta_{\mathrm{D}}' + \Delta_{\mathrm{S}}' c a k_z)}{\epsilon'(\boldsymbol{k})^4}, \tag{A6}$$

where $\Delta_{\mathrm{S}}' \equiv \Delta_{\mathrm{S}}/\hbar v_{\mathrm{D}}$, $\Delta_{\mathrm{D}}' \equiv \Delta_{\mathrm{D}}/\hbar v_{\mathrm{D}}$, etc., and $\doteq$ denotes an equality under the above described assumptions of the $\boldsymbol{k} \cdot \boldsymbol{p}$ approach. Evaluating the integral,

$$\int_{-\pi/a}^{\pi/a} \frac{dk_z}{2\pi} \int_0^\infty dk_\perp k_\perp \frac{a\Delta_{\mathrm{D}}'(\Delta_{\mathrm{D}}' + \Delta_{\mathrm{S}}' c a k_z)}{\left(k_\perp^2 + |\Delta_{k_z}'|^2\right)^2}$$

$$= \begin{cases} 0 & \text{if } |\Delta_{\mathrm{S}}| > |\Delta_{\mathrm{D}}| \\ \frac{1}{2} & \text{if } |\Delta_{\mathrm{S}}| < |\Delta_{\mathrm{D}}|, \end{cases} \tag{A7}$$

yields the expected quantization of α_{CS}. This is consistent with the result of the tight-binding calculation [see Eq. (50) and the surrounding text].

But what happens under a change of gauge? The result is that, in general, a change of gauge shifts the value of α_{CS} by an amount different than $n\frac{e^2}{hc}$ for $n \in \mathbb{Z}$, in contradiction to the known behavior for a model defined over the entire Brillouin zone [10,11]. This is an artifact of the use of a $\boldsymbol{k} \cdot \boldsymbol{p}$ description, which will not generally yield quantized topological invariants.[49] In the case of α_{CS} in particular, the issue is that one does not know whether the eigenvectors employed in the gauge choice on the BZ_d subset on which the $\boldsymbol{k} \cdot \boldsymbol{p}$ model is valid extend smoothly and periodically over the BZ_d of a lattice regularization thereof. For example, in the model studied here, consider a gauge constructed from $|v_{n,\boldsymbol{k}}'\rangle = \sum_{m=1}^4 |v_{m,k_z}\rangle U_{m,n}(\boldsymbol{k})$, where

$$U(\boldsymbol{k}) = U(k_z) = \begin{pmatrix} 0 & \Delta_{k_z}/|\Delta_{k_z}| & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta_{k_z}/|\Delta_{k_z}| \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{A8}$$

This appears harmless within the context of the model (A1) for an insulator ($\Delta_{k_z} \neq 0$). But, in the lattice regularization presented in the main text, $\Delta_{k_z} \to \Delta_{\boldsymbol{k}}$ and there may exist $\boldsymbol{k} \in \mathrm{BZ}_{3\mathrm{D}} : |\Delta_{\boldsymbol{k}}| = 0$ in a band insulator. Then attempting to define Bloch energy eigenvectors over the full Brillouin zone from the $|v_{n,\boldsymbol{k}}'\rangle$ [as was done in the main text (37) for the $|v_{n,\boldsymbol{k}}\rangle$] is pathological at any $\boldsymbol{k} \in \mathrm{BZ}_{3\mathrm{D}} : |\Delta_{\boldsymbol{k}}| = 0$; those vectors cannot be used to construct a smooth gauge over $\mathrm{BZ}_{3\mathrm{D}}$. Indeed a calculation of α_{CS} in that gauge would be invalid, since

---

[49]See, e.g., Jotzu [90].

a smooth global gauge was assumed in deriving the BZ_{3D} integral form of α_{CS}.

APPENDIX B: ANTIFERROMAGNETIC 3D MULTILAYER TIGHT-BINDING HAMILTONIAN

We now modify the considerations of Sec. IV A 2 to the antiferromagnetic (AFM) case ($M_{l_z} = -M_{l_z \pm 1}$). In this case the magnetic unit cell is twice the volume of the chemical unit cell. That is, the single-particle Bloch Hamiltonian $H(\mathbf{r}, \mathfrak{p}(\mathbf{r}))$ underlying a tight-binding description would not be invariant under translations by az, but would be invariant under translation by $2az$; it is useful to define $\Gamma_{3D} \equiv \Gamma_{2D} \times a\mathbb{Z}$ and $\Gamma_H \equiv \Gamma_{2D} \times 2a\mathbb{Z}$. Then $\mathbf{R} + az \notin \Gamma_H$ and strictly speaking Eq. (33) is not valid. Nevertheless we can modify that expression by relabelling the operators in such a way that the physical content is unchanged and the operators are identified in a technically correct manner. This achieved through the following: for $l_z \in \mathbb{Z}$ and $\mathbf{R} \in \Gamma_{2D} \times \{0\}$, take

$$\left| W_{(\alpha,\sigma),\mathbf{R}+l_z az} \right\rangle \to \begin{cases} \left| W_{(\alpha,\sigma),\mathbf{R}+l_z az} \right\rangle & \text{for } l_z \text{ even} \\ \left| W_{(\alpha+2,\sigma),\mathbf{R}+(l_z-1)az} \right\rangle & \text{for } l_z \text{ odd.} \end{cases} \quad \text{(B1)}$$

Under this transformation Eq. (33) is consistent; operators that are associated with WFs are labeled by elements of Γ_H rather than by elements of Γ_{3D}. This relabeling results in eight types (α, σ) of WFs $W_{(\alpha,\sigma),\mathbf{R}}(\mathbf{r})$ associated with each Bravais lattice vector $\mathbf{R} \in \Gamma_H$. Separating out the $l_z = $ odd contributions and implementing the necessary relabeling (B1), the tight-binding Hamiltonian (33) is rewritten as

$$\hat{H}^{(3D)} = \sum_{l_z \in 2\mathbb{Z}} \hat{H}'^{(2D)}(l_z)$$

$$+ \Delta_D \sum_{\substack{\mathbf{R} \in \Gamma_{2D} \times \{0\} \\ l_z \in 2\mathbb{Z} \\ \sigma \in \{\uparrow,\downarrow\}}} \left(\hat{c}^\dagger_{(2,\sigma),\mathbf{R}+l_z az} \hat{c}_{(1,\sigma),\mathbf{R}+l_z az} \right.$$

$$\left. + \hat{c}^\dagger_{(0,\sigma),\mathbf{R}+(l_z+2)az} \hat{c}_{(3,\sigma),\mathbf{R}+l_z az} + \text{H.c.} \right), \quad \text{(B2)}$$

where $\hat{H}'^{(2D)}(l_z)$ is obtained by implementing the relabeling (B1) in $\hat{H}^{(2D)}(l_z) + \hat{H}^{(2D)}(l_z + 1)$.

There is another, more physical, complication in the case when $M_{l_z} \neq 0$ in 3D, which is the possibility of exchange coupling between the dynamical degrees of freedom of a given layer with the static magnetization of all other layers. If we assume that the most relevant exchange interactions for a given WF $W_{(\alpha,\sigma),\mathbf{R}}(\mathbf{r})$ is that with the static magnetization of the layer for which it is associated (via J_S) and that with the static magnetization of the layer with which it is nearest (via J_D)—for $\alpha = $ odd (even) this is the layer below (above) that for which $W_{(\alpha,\sigma),\mathbf{R}}(\mathbf{r})$ is associated—and we focus on the AFM case such that both orbitals within a given layer feel the same $M_{l_z \pm 1} = -M_{l_z}$, we can account for this by taking $J_S M_{l_z} \to (J_S - J_D) M_{l_z} \equiv m_{l_z}$ in Eq. (31). Taking $m_{l_z} = m$ for l_z even and $= -m$ for l_z odd, writing (B2) as a BZ_{3D} integral via Eq. (30), and using Eq. (34), we have

$$\hat{H}^{(3D)}_{\text{reg}} = \int_{\text{BZ}_{3D}} \frac{d^3 k}{(2\pi)^3} \hat{c}^\dagger(\mathbf{k}) \mathcal{H}^{(3D)}_{\text{reg}}(\mathbf{k}) \hat{c}(\mathbf{k}),$$

which is of the same form as (29), but in this case $\hat{c}(\mathbf{k}) \equiv (\hat{c}_{(0,\uparrow),\mathbf{k}}, \hat{c}_{(0,\downarrow),\mathbf{k}}, \hat{c}_{(1,\uparrow),\mathbf{k}}, \hat{c}_{(1,\downarrow),\mathbf{k}}, \hat{c}_{(2,\uparrow),\mathbf{k}}, \hat{c}_{(2,\downarrow),\mathbf{k}}, \hat{c}_{(3,\uparrow),\mathbf{k}}, \hat{c}_{(3,\downarrow),\mathbf{k}})^T$ and $\hat{c}^\dagger(\mathbf{k}) \equiv (\hat{c}^\dagger_{(0,\uparrow),\mathbf{k}}, \hat{c}^\dagger_{(0,\downarrow),\mathbf{k}}, \hat{c}^\dagger_{(1,\uparrow),\mathbf{k}}, \hat{c}^\dagger_{(1,\downarrow),\mathbf{k}}, \hat{c}^\dagger_{(2,\uparrow),\mathbf{k}}, \hat{c}^\dagger_{(2,\downarrow),\mathbf{k}}, \hat{c}^\dagger_{(3,\uparrow),\mathbf{k}}, \hat{c}^\dagger_{(3,\downarrow),\mathbf{k}})$, and the general $\mathcal{H}^{(d)}(\mathbf{k})$ is taken to be

$$\mathcal{H}^{(3D)}_{\text{reg}}(\mathbf{k})$$
$$= \begin{pmatrix} m & iA(sk_x - isk_y) & \Delta(k_x,k_y) & 0 & 0 & 0 & e^{-2iak_z}\Delta_D & 0 \\ -iA(sk_x + isk_y) & -m & 0 & \Delta(k_x,k_y) & 0 & 0 & 0 & e^{-2iak_z}\Delta_D \\ \Delta(k_x,k_y) & 0 & m & -iA(sk_x - isk_y) & \Delta_D & 0 & 0 & 0 \\ 0 & \Delta(k_x,k_y) & iA(sk_x + isk_y) & -m & 0 & \Delta_D & 0 & 0 \\ 0 & 0 & \Delta_D & 0 & -m & iA(sk_x - isk_y) & \Delta(k_x,k_y) & 0 \\ 0 & 0 & 0 & \Delta_D & -iA(sk_x + isk_y) & m & 0 & \Delta(k_x,k_y) \\ e^{2iak_z}\Delta_D & 0 & 0 & 0 & \Delta(k_x,k_y) & 0 & -m & -iA(sk_x - isk_y) \\ 0 & e^{2iak_z}\Delta_D & 0 & 0 & 0 & \Delta(k_x,k_y) & iA(sk_x + isk_y) & m \end{pmatrix}. \quad \text{(B3)}$$

The eigenvalues of (B3) are

$$E_{1,2}(\mathbf{k}) = -E_{7,8}(\mathbf{k}) = -\sqrt{\frac{1}{2}A^2(2 - c2k_x - c2k_y) + \Delta_S(k_x,k_y)^2 + \Delta_D^2 + m^2 + 2\sqrt{\Delta_S(k_x,k_y)^2(\Delta_D^2 \cos^2(ak_z) + m^2)}},$$

$$E_{3,4}(\mathbf{k}) = -E_{5,6}(\mathbf{k}) = -\sqrt{\frac{1}{2}A^2(2 - c2k_x - c2k_y) + \Delta_S(k_x,k_y)^2 + \Delta_D^2 + m^2 - 2\sqrt{\Delta_S(k_x,k_y)^2(\Delta_D^2 \cos^2(ak_z) + m^2)}}. \quad \text{(B4)}$$

The double degeneracy of the energy bands at each $\mathbf{k} \in \text{BZ}$ follows from the combination of an inversion and a (fermionic) time-reversal symmetry (see Appendix C). Notice that taking $m = 0$ in (B4) does *not* reproduce (A2). This is not surprising because taking this limit implements an incorrect identification of the Bravais lattice for the nonmagnetic Hamiltonian; in deriving (B4) we explicitly use $\Gamma_H = \Gamma_{2D} \times 2a\mathbb{Z}$ rather than the correct $\Gamma_H = \Gamma_{3D} = \Gamma_{2D} \times a\mathbb{Z}$ implemented in deriving (35) in the nonmagnetic case.

APPENDIX C: SYMMETRY ANALYSIS OF THE NONMAGNETIC AND ANTIFERROMAGNETIC TIGHT-BINDING MODELS

In this Appendix we demonstrate that both the nonmagnetic and antiferromagnetic tight-binding Hamiltonian operators, specified by the matrix kernels (35) and (B3), respectively, have a center-of-inversion symmetry and a time-reversal symmetry. We use the language of Refs. [64,91] and we begin by reiterating (in a loose manner) some contents of those works to set notation. Starting with a single-particle Hamiltonian $H : \mathcal{H} \rightarrow \mathcal{H}$, which is a linear and self-adjoint operator in the electronic Hilbert space $\mathcal{H}$, and an orthonormal set of basis vectors $|\phi_A\rangle$ of $\mathcal{H}$, for A a general multi-index, we have $H = \sum_{A,B} H_{A,B} |\phi_A\rangle \otimes \langle\phi_B|$, where $H_{A,B} \equiv \langle\phi_A | H \phi_B\rangle \in \mathbb{C}$ for a given pair (A, B). H can be generalized to a single-body operator $\hat{\mathscr{H}} = \sum_{A,B} H_{A,B} \hat{c}_A^\dagger \hat{c}_B$ in the electronic Fock space $\mathcal{F}$, where $\hat{c}_A^\dagger$, $\hat{c}_A$ are fermionic operators that satisfy $\{\hat{c}_A^\dagger, \hat{c}_B\} = \delta_{A,B}$ and $\hat{c}_A^\dagger |\text{vac}\rangle \equiv |\phi_A\rangle$, and $\mathcal{F}$ is constructed by antisymmetrizing sums of products of $\mathcal{H}$. $\hat{\mathscr{H}}$ agrees with H when acting on one-particle states; this will be assumed of any generalization of an operator on $\mathcal{H}$ to one on $\mathcal{F}$. Products of operators on $\mathcal{H}$ or $\mathcal{F}$ is by composition, which we usually keep implicit.

Wigner's theorem dictates that any symmetry transformation $S : \mathcal{H} \rightarrow \mathcal{H}$ is either linear and unitary or antilinear and anti-unitary; we here neglect symmetry transformations that do not preserve particle number, for example, particle-hole symmetry. By definition, $\hat{\mathscr{S}}$ is a symmetry of $\hat{\mathscr{H}}$ if and only if $\hat{\mathscr{S}} \hat{\mathscr{H}} \hat{\mathscr{S}}^{-1} = \hat{\mathscr{H}}$ is satisfied, where $\hat{\mathscr{S}}$ is the generalization of S to an operator on $\mathcal{F}$. An insight is gained by noting that both $\hat{\mathscr{S}} \hat{c}_A \hat{\mathscr{S}}^{-1}$ and $\hat{c}_A$ map $\mathcal{H}^{(N)} \rightarrow \mathcal{H}^{(N-1)}$, and that both $\hat{\mathscr{S}} \hat{c}_A^\dagger \hat{\mathscr{S}}^{-1}$ and $\hat{c}_A^\dagger$ map $\mathcal{H}^{(N)} \rightarrow \mathcal{H}^{(N+1)}$ within $\mathcal{F}$. Thus, for every symmetry transformation S there exists collections of numbers $(U_S)_{A,B}$, $(V_S)_{A,B} \in \mathbb{C}$ that satisfy

$$\hat{\mathscr{S}} \hat{c}_A \hat{\mathscr{S}}^{-1} = \sum_B (U_S)_{A,B} \hat{c}_B, \quad \hat{\mathscr{S}} \hat{c}_A^\dagger \hat{\mathscr{S}}^{-1} = \sum_B \hat{c}_B^\dagger (V_S)_{B,A}. \tag{C1}$$

It follows from $\hat{\mathscr{S}} \{\hat{c}_A, \hat{c}_B^\dagger\} \hat{\mathscr{S}}^{-1} = \hat{\mathscr{S}} \{\hat{c}_A^\dagger, \hat{c}_B\} \hat{\mathscr{S}}^{-1} = \delta_{A,B}$ that the matrices V_S and U_S are each others inverse; that is, the matrices U_S and V_S constructed from the coefficients in Eq. (C1) satisfy $U_S V_S = V_S U_S = \mathbf{1}$, where matrix products are taken as the usual matrix multiplication. If S is (anti-)unitary, then

$$\delta_{AB} = \langle\phi_A | \phi_B\rangle = \langle S\phi_A | S\phi_B\rangle^{(*)} \equiv \langle\text{vac}| (\hat{\mathscr{S}} \hat{c}_A^\dagger \hat{\mathscr{S}}^{-1})^\dagger \hat{\mathscr{S}} \hat{c}_B^\dagger \hat{\mathscr{S}}^{-1} |\text{vac}\rangle^{(*)}$$

$$= \left(\sum_{C,D} (V_S)_{C,A}^* (V_S)_{D,B} \langle\text{vac}| \hat{c}_C \hat{c}_D^\dagger |\text{vac}\rangle \right)^{(*)} = \left(\sum_C (V_S^\dagger)_{A,C} (V_S)_{C,B} \right)^{(*)}. \tag{C2}$$

Thus $V_S^\dagger = U_S = V_S^{-1}$ is unitary and $(V_S)_{B,A} = (U_S)_{A,B}^*$ in Eq. (C1). With this, if $\hat{\mathscr{S}}$ is a symmetry of $\hat{\mathscr{H}}$, then

$$\sum_{A,B} (\hat{\mathscr{S}} \hat{c}_A^\dagger \hat{\mathscr{S}}^{-1})(\hat{\mathscr{S}} H_{A,B} \hat{\mathscr{S}}^{-1})(\hat{\mathscr{S}} \hat{c}_B \hat{\mathscr{S}}^{-1}) = \sum_{I,J} \hat{c}_I^\dagger \left(\sum_{A,B} (U_S)_{A,I}^* H_{A,B}^{(*)} (U_S)_{B,J} \right) \hat{c}_J = \sum_{I,J} \hat{c}_I^\dagger H_{I,J} \hat{c}_J^\dagger$$

$$\Longleftrightarrow \sum_{A,B} (U_S^\dagger)_{I,A} H_{A,B}^{(*)} (U_S)_{B,J} = H_{I,J}, \tag{C3}$$

where $H_{A,B}^{(*)} = H_{A,B} (H_{A,B}^*)$ for S being (anti-)linear.

Any antilinear and anti-unitary symmetry $\hat{\mathscr{S}}$ of $\hat{\mathscr{H}}$ might be called a time-reversal symmetry because of its affect on the time-evolution operator $\hat{\mathscr{U}}(t) = e^{-i\hat{\mathscr{H}}t/\hbar}$; if $\hat{\mathscr{S}}$ is an anti-unitary symmetry, then $\hat{\mathscr{S}} \hat{\mathscr{U}}(t) \hat{\mathscr{S}}^{-1} = \hat{\mathscr{U}}(-t)$. See Footnote 13 of Ryu *et al.* [64] for a discussion on uniqueness. Nevertheless, there is typically one physically motivated operator named the time-reversal transformation. When S is anti-unitary we have

$$\hat{\mathscr{S}}^2 \hat{c}_A (\hat{\mathscr{S}}^{-1})^2 \equiv \hat{\mathscr{S}} (\hat{\mathscr{S}} \hat{c}_A \hat{\mathscr{S}}^{-1}) \hat{\mathscr{S}}^{-1} = \sum_{B,D} (U_S)_{A,B}^* (U_S)_{B,D} \hat{c}_D = \sum_D (U_S^* U_S)_{A,D} \hat{c}_D,$$

$$\hat{\mathscr{S}}^2 \hat{c}_A^\dagger (\hat{\mathscr{S}}^{-1})^2 \equiv \hat{\mathscr{S}} (\hat{\mathscr{S}} \hat{c}_A^\dagger \hat{\mathscr{S}}^{-1}) \hat{\mathscr{S}}^{-1} = \sum_{B,D} (U_S)_{A,B} (U_S)_{B,D}^* \hat{c}_D^\dagger = \sum_D (U_S^* U_S)_{A,D}^* \hat{c}_D^\dagger. \tag{C4}$$

It has been shown [64] that in general $U_S^* U_S \in \{-\mathbf{1}, \mathbf{1}\}$, thus either $\hat{\mathscr{S}}^2 \hat{c}_A (\hat{\mathscr{S}}^{-1})^2 = \hat{c}_A$ and $\hat{\mathscr{S}}^2 \hat{c}_A^\dagger (\hat{\mathscr{S}}^{-1})^2 = \hat{c}_A^\dagger$ or $\hat{\mathscr{S}}^2 \hat{c}_A (\hat{\mathscr{S}}^{-1})^2 = -\hat{c}_A$ and $\hat{\mathscr{S}}^2 \hat{c}_A^\dagger (\hat{\mathscr{S}}^{-1})^2 = -\hat{c}_A^\dagger$ are satisfied. This is familiar from the usual analysis on $\mathcal{H}$; i.e., in the case $U_S^* U_S = -\mathbf{1}$, Kramer's theorem results.

We now identify some symmetry operators $\hat{\mathscr{S}}$ (and the corresponding U_S) for which Eq. (C3) is satisfied in the case of the nonmagnetic (35) or the antiferromagnetic (B3) tight-binding Hamiltonian.

1. Inversion symmetry

An inversion operator $\mathcal{I}$ on the relevant electronic Hilbert space $\mathcal{H}$ is taken to be linear and unitary and satisfy the physically motivated relation $\mathcal{I}^2 = \text{id}_{\mathcal{H}}$. Any (anti-)linear operator can be specified by its action on a basis of the space in which it acts. For

the nonmagnetic Hamiltonian we define

$$\mathcal{I}|W_{(0,\sigma),\boldsymbol{R}}\rangle = |W_{(1,\sigma),-\boldsymbol{R}}\rangle \iff \mathcal{I}|\bar{\psi}_{(0,\sigma),\boldsymbol{k}}\rangle = |\bar{\psi}_{(1,\sigma),-\boldsymbol{k}}\rangle \tag{C5}$$

for $\sigma \in \{\uparrow, \downarrow\}$ and $I = ((\alpha, \sigma), \boldsymbol{k})$ is the relevant multi-index on the right-hand side. This transformation corresponds to inversion about the (single) layer center within the unit cell. Equation (C5) can be generalized to a relation involving Fock-space operators $\hat{c}^{\dagger}_{(\alpha,\sigma),\boldsymbol{k}}$,

$$\hat{\mathscr{I}}\hat{c}^{\dagger}_{(0,\sigma),\boldsymbol{k}}\hat{\mathscr{I}}^{-1} = \hat{c}^{\dagger}_{(1,\sigma),-\boldsymbol{k}} \Rightarrow (U_{\mathcal{I}})_{((\alpha,\sigma),\boldsymbol{k}),((\beta,\gamma),\boldsymbol{k}')} = (\tau_x)_{\alpha,\beta}\delta_{\sigma,\gamma}\delta(\boldsymbol{k} + \boldsymbol{k}'), \tag{C6}$$

where the identification of $U_{\mathcal{I}}$ follows from comparison with (C1) and using $(V_S)_{B,A} = (U_S)^*_{A,B}$; equivalently one could consider relations for the $\hat{c}_{(\alpha,\sigma),\boldsymbol{k}}$. To ease notation, we do not include factors to make this unitless, and take the unitarity to be understood as

$$\sum_{\substack{\beta \in \{1,\dots,N_{\text{orb}}\} \\ \gamma \in \{\uparrow,\downarrow\}}} \int_{\text{BZ}_d} d^d k' (U_S)_{((\alpha,\sigma),\boldsymbol{k}),((\beta,\gamma),\boldsymbol{k}')}(U_S^{\dagger})_{((\beta,\gamma),\boldsymbol{k}'),((\mu,\lambda),\boldsymbol{k}'')} = \delta_{\alpha,\mu}\delta_{\sigma,\gamma}\delta(\boldsymbol{k} - \boldsymbol{k}''), \tag{C7}$$

where $N_{\text{orb}} = 2(4)$ in the nonmagnetic (antiferromagnetic) case. For the AFM Hamiltonian we define

$$\mathcal{I}|W_{(0,\sigma),\boldsymbol{R}}\rangle = |W_{(1,\sigma),-\boldsymbol{R}}\rangle \impliedby \hat{\mathscr{I}}\hat{c}^{\dagger}_{(0,\sigma),\boldsymbol{k}}\hat{\mathscr{I}}^{-1} = \hat{c}^{\dagger}_{(1,\sigma),-\boldsymbol{k}},$$

$$\mathcal{I}|W_{(2,\sigma),\boldsymbol{R}}\rangle = |W_{(3,\sigma),-\boldsymbol{R}-\boldsymbol{a}_3}\rangle \impliedby \hat{\mathscr{I}}\hat{c}^{\dagger}_{(2,\sigma),\boldsymbol{k}}\hat{\mathscr{I}}^{-1} = e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_3}\hat{c}^{\dagger}_{(3,\sigma),-\boldsymbol{k}}, \tag{C8}$$

which leads to

$$(U_{\mathcal{I}})_{((\alpha,\sigma),\boldsymbol{k}),((\beta,\gamma),\boldsymbol{k}')} = \left(\begin{pmatrix} 1 & 0 \\ 0 & e^{i2ak_z} \end{pmatrix} \otimes \tau_x\right)_{\alpha,\beta} \delta_{\sigma,\gamma}\delta(\boldsymbol{k} + \boldsymbol{k}'), \tag{C9}$$

where $\boldsymbol{a}_3 = 2a\boldsymbol{z}$. This transformation corresponds to inversion about the center of the bottom layer in the unit cell; one could equally well define an inversion operation with respect to the center of the top layer. Physically, this corresponds to inversion about the position of a magnetic ion, which is indeed a symmetry of the crystal lattice [66].

Then one can explicitly show that

$$\hat{\mathscr{I}}\hat{H}^{(3D)}_{\text{reg}}\hat{\mathscr{I}}^{-1} = \hat{H}^{(3D)}_{\text{reg}} \tag{C10}$$

is satisfied for the $\hat{H}^{(3D)}_{\text{reg}}$ the appropriate nonmagnetic (35) or antiferromagnetic (B3) tight-binding Hamiltonian and $\hat{\mathscr{I}}$ the corresponding inversion operator (C6) or (C8) and (C9). In particular, in the nonmagnetic case the inversion operator (C6) is a symmetry of the Hamiltonian if and only if

$$(\tau_x \otimes \text{id}_{2\times2})\mathcal{H}^{(3D)}_{\text{reg}}(-\boldsymbol{k})(\tau_x \otimes \text{id}_{2\times2}) = \mathcal{H}^{(3D)}_{\text{reg}}(\boldsymbol{k}) \tag{C11}$$

[which plays the role of (C3) in this case], where we have equated Brillouin-zone integrands and taken $\boldsymbol{k} \to -\boldsymbol{k}$ in the BZ_{3D}-integral related to $\hat{\mathscr{I}}\hat{H}^{(3D)}_{\text{reg}}\hat{\mathscr{I}}^{-1}$ on the left-hand side (LHS) of Eq. (C10). Given the matrix kernel (35), one can explicitly show that Eq. (C11) is satisfied. Similarly, given the matrix kernel of the AFM Hamiltonian (B3), the inversion operator (C8) is a symmetry since one can explicitly show that

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & e^{-i2ak_z} \end{pmatrix} \otimes \tau_x \otimes \text{id}_{2\times2}\right)\mathcal{H}^{(3D)}_{\text{reg}}(\boldsymbol{k})\left(\begin{pmatrix} 1 & 0 \\ 0 & e^{i2ak_z} \end{pmatrix} \otimes \tau_x \otimes \text{id}_{2\times2}\right) = \mathcal{H}^{(3D)}_{\text{reg}}(-\boldsymbol{k}) \tag{C12}$$

is satisfied. To arrive at Eq. (C12) we have again equated BZ_{3D} integrands but have here taken $\boldsymbol{k} \to -\boldsymbol{k}$ in the BZ_{3D} integral related to $\hat{H}^{(3D)}_{\text{reg}}$ on the right-hand side (RHS) of Eq. (C10). Note that the AFM Hamiltonian (B3) is not symmetric with respect to inversion about the center point between the two layers within a unit cell.

2. Time-reversal symmetry

a. One option

A time-reversal operator $\mathcal{T}$ is always taken to be an antilinear and anti-unitary operator on the electronic Hilbert space $\mathcal{H}$. We begin by seeking a traditional time-reversal transformation $\mathcal{T}$, which takes a state with spin components (s_x, s_y, s_z) to one with $(-s_x, -s_y, -s_z)$. One option is[50]

$$\mathcal{T}|W_{(\alpha,\uparrow),\boldsymbol{R}}\rangle = -|W_{(\alpha,\downarrow),\boldsymbol{R}}\rangle, \quad \mathcal{T}|W_{(\alpha,\downarrow),\boldsymbol{R}}\rangle = |W_{(\alpha,\uparrow),\boldsymbol{R}}\rangle, \tag{C13}$$

---

[50]See, e.g., pp. 277–280 of Sakurai [61].

for $\alpha \in \{0, 1\}$ ($\alpha \in \{0, 1, 2, 3\}$) in the nonmagnetic (AFM) case. A generalization to a Fock-space relation is

$$\hat{\mathscr{T}} \hat{c}^{\dagger}_{(\alpha,\uparrow),\boldsymbol{k}} \hat{\mathscr{T}}^{-1} = -\hat{c}^{\dagger}_{(\alpha,\downarrow),-\boldsymbol{k}}, \qquad \hat{\mathscr{T}} \hat{c}^{\dagger}_{(\alpha,\downarrow),\boldsymbol{k}} \hat{\mathscr{T}}^{-1} = \hat{c}^{\dagger}_{(\alpha,\uparrow),-\boldsymbol{k}}$$

$$\Rightarrow \hat{\mathscr{T}}^2 \hat{c}^{\dagger}_{(\alpha,\sigma),\boldsymbol{k}} (\hat{\mathscr{T}}^{-1})^2 = -\hat{c}^{\dagger}_{(\alpha,\sigma),\boldsymbol{k}}, \tag{C14}$$

which is consistent with (C4) for $U_{\mathcal{T}}^* U_{\mathcal{T}} = -\mathbf{1}$. In particular, comparing the second line of (C4) and the second line of (C14) gives $(U_{\mathcal{T}}^* U_{\mathcal{T}})_{((\alpha,\sigma),\boldsymbol{k}),((\beta,\gamma),\boldsymbol{k}')} = -\delta_{\alpha,\beta}\delta_{\sigma,\gamma}\delta(\boldsymbol{k} - \boldsymbol{k}')$. Then from (C1) and the first line of (C14) we identify

$$(U_{\mathcal{T}})_{((\alpha,\sigma),\boldsymbol{k}),((\beta,\gamma),\boldsymbol{k}')} = \delta_{\alpha,\beta}(-i\sigma_y)_{\sigma,\gamma}\delta(\boldsymbol{k} + \boldsymbol{k}'). \tag{C15}$$

With this one can show

$$\hat{\mathscr{T}} \hat{H}^{(3D)}_{\text{reg}} \hat{\mathscr{T}}^{-1} = \hat{H}^{(3D)}_{\text{reg}} \tag{C16}$$

is satisfied for the nonmagnetic Hamiltonian (35), but not for the antiferromagnetic (AFM) Hamiltonian (B3).

In particular, given the matrix kernel of the nonmagnetic Hamiltonian (35), one can explicitly show that

$$(\text{id}_{2\times 2} \otimes (i\sigma_y))\mathcal{H}^{(3D)}_{\text{reg}}(-\boldsymbol{k})^*(\text{id}_{2\times 2} \otimes (-i\sigma_y)) = \mathcal{H}^{(3D)}_{\text{reg}}(\boldsymbol{k}), \tag{C17}$$

which is (C3) in this case, and again we have equated Brillouin-zone integrands in Eq. (C16).

In contrast, given the matrix kernel of the AFM Hamiltonian (B3), the time-reversal transformation (C13) is not a symmetry due to the terms related to the exchange interaction. Indeed, upon explicit evaluation one finds

$$[(\text{id}_{4\times 4} \otimes (i\sigma_y))\mathcal{H}^{(3D)}_{\text{reg}}(-\boldsymbol{k})^*(\text{id}_{4\times 4} \otimes (-i\sigma_y))] - \mathcal{H}^{(3D)}_{\text{reg}}(\boldsymbol{k}) = -2m(\text{id}_{4\times 4} \otimes \sigma_z). \tag{C18}$$

Nevertheless, there is a different time-reversal-like transformation that is a symmetry of the AFM Hamiltonian (B3).

b. A second option

In the AFM case we are physically motivated to consider an operation that involves a spin flip, as in the preceding section, followed by a translation by $a z$. This can be obtained by a product of operators, T_{az} and $\mathcal{T}$, where $\mathcal{T}$ is as before [Eq. (C15)] and we take $T_{az} : \mathcal{H} \to \mathcal{H}$ to be linear and unitary, and such that

$$T_{az}|W_{(\alpha,\sigma),\boldsymbol{R}}\rangle = \begin{cases} |W_{(\alpha+2,\sigma),\boldsymbol{R}}\rangle & \text{if } \alpha \in \{0, 1\} \\ |W_{(\alpha-2,\sigma),\boldsymbol{R}+\boldsymbol{a}_3}\rangle & \text{if } \alpha \in \{2, 3\} \end{cases} \impliedby \hat{T}_{az}\hat{c}^{\dagger}_{(\alpha,\sigma),\boldsymbol{k}}\hat{T}_{az}^{-1} = \begin{cases} \hat{c}^{\dagger}_{(\alpha+2,\sigma),\boldsymbol{k}} & \text{if } \alpha \in \{0, 1\} \\ e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_3}\hat{c}^{\dagger}_{(\alpha-2,\sigma),\boldsymbol{k}} & \text{if } \alpha \in \{2, 3\}. \end{cases} \tag{C19}$$

Then,

$$(\hat{T}_{az}\hat{\mathscr{T}})\hat{c}^{\dagger}_{(\alpha,\uparrow),\boldsymbol{k}}(\hat{T}_{az}\hat{\mathscr{T}})^{-1} = \begin{cases} -\hat{c}^{\dagger}_{(\alpha+2,\downarrow),-\boldsymbol{k}} & \text{if } \alpha \in \{0, 1\} \\ -e^{i\boldsymbol{k}\cdot\boldsymbol{a}_3}\hat{c}^{\dagger}_{(\alpha-2,\downarrow),-\boldsymbol{k}} & \text{if } \alpha \in \{2, 3\}, \end{cases}$$

$$(\hat{T}_{az}\hat{\mathscr{T}})\hat{c}^{\dagger}_{(\alpha,\downarrow),\boldsymbol{k}}(\hat{T}_{az}\hat{\mathscr{T}})^{-1} = \begin{cases} \hat{c}^{\dagger}_{(\alpha+2,\uparrow),-\boldsymbol{k}} & \text{if } \alpha \in \{0, 1\} \\ e^{i\boldsymbol{k}\cdot\boldsymbol{a}_3}\hat{c}^{\dagger}_{(\alpha-2,\uparrow),-\boldsymbol{k}} & \text{if } \alpha \in \{2, 3\}, \end{cases} \tag{C20}$$

and thus

$$(U_{T_{az}\mathcal{T}})_{((\alpha,\sigma),\boldsymbol{k}),((\beta,\gamma),\boldsymbol{k}')} = \left(\begin{pmatrix} 0 & 1 \\ e^{-i2ak_z} & 0 \end{pmatrix} \otimes \text{id}_{2\times 2}\right)_{\alpha,\beta} (-i\sigma_y)_{\sigma,\gamma}\delta(\boldsymbol{k} + \boldsymbol{k}'). \tag{C21}$$

Then, given the matrix kernel of the AFM Hamiltonian (B3), this time-reversal-like transformation is a symmetry,

$$(\hat{T}_{az}\hat{\mathscr{T}})\hat{H}^{(3D)}_{\text{reg}}(\hat{T}_{az}\hat{\mathscr{T}})^{-1} = \hat{H}^{(3D)}_{\text{reg}}, \tag{C22}$$

since one can explicitly show

$$\left(\begin{pmatrix} 0 & e^{i2ak_z} \\ 1 & 0 \end{pmatrix} \otimes \text{id}_{2\times 2} \otimes (i\sigma_y)\right)\mathcal{H}^{(3D)}_{\text{reg}}(\boldsymbol{k})^*\left(\begin{pmatrix} 0 & 1 \\ e^{-i2ak_z} & 0 \end{pmatrix} \otimes \text{id}_{2\times 2} \otimes (-i\sigma_y)\right) = \mathcal{H}^{(3D)}_{\text{reg}}(-\boldsymbol{k}). \tag{C23}$$

We have again equated Brillouin-zone integrands and here taken $\boldsymbol{k} \to -\boldsymbol{k}$ in the BZ$_{3D}$-integral related to $\hat{H}^{(3D)}_{\text{reg}}$ on the RHS of Eq. (C22). Indeed, in the AFM case it is this symmetry that leads to a $\mathbb{Z}_2$ topological classification of bulk insulating electronic ground states [52,66] similar to the nonmagnetic case [92,93].

In summary, both the nonmagnetic and antiferromagnetic Hamiltonian operators, specified by their matrix kernels (35) and (B3), respectively, have a center-of-inversion symmetry and a time-reversal symmetry.

[1] R. Ernst Peierls, *Surprises in Theoretical Physics* (Princeton University Press, Princeton, New Jersey, 1979), Vol. 107.

[2] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Quantized Hall conductance in a two-dimensional periodic potential, Phys. Rev. Lett. **49**, 405 (1982).

[3] B. Simon, Holonomy, the quantum adiabatic theorem, and Berry's phase, Phys. Rev. Lett. **51**, 2167 (1983).

[4] M. Kohmoto, Topological invariant and the quantization of the Hall conductance, Ann. Phys. (NY) **160**, 343 (1985).

[5] B. I. Halperin, Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential, Phys. Rev. B **25**, 2185 (1982).

[6] A. H. MacDonald and P. Středa, Quantized Hall effect and edge currents, Phys. Rev. B **29**, 1616 (1984).

[7] M. Büttiker, Absence of backscattering in the quantum Hall effect in multiprobe conductors, Phys. Rev. B **38**, 9375 (1988).

[8] D. J. Thouless, Edge voltages and distributed currents in the quantum Hall effect, Phys. Rev. Lett. **71**, 1879 (1993).

[9] D. Tong, Lectures on the quantum Hall effect, arXiv:1606.06687.

[10] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Topological field theory of time-reversal invariant insulators, Phys. Rev. B **78**, 195424 (2008).

[11] A. M. Essin, J. E. Moore, and D. Vanderbilt, Magnetoelectric polarizability and axion electrodynamics in crystalline insulators, Phys. Rev. Lett. **102**, 146805 (2009).

[12] H.-G. Zirnstein and B. Rosenow, Cancellation of quantum anomalies and bosonization of three-dimensional time-reversal symmetric topological insulators, Phys. Rev. B **88**, 085105 (2013).

[13] M. Mulligan and F. J. Burnell, Topological insulators avoid the parity anomaly, Phys. Rev. B **88**, 085104 (2013).

[14] E. Witten, Fermion path integrals and topological phases, Rev. Mod. Phys. **88**, 035001 (2016).

[15] A. M. Essin, A. M. Turner, J. E. Moore, and D. Vanderbilt, Orbital magnetoelectric coupling in band insulators, Phys. Rev. B **81**, 205104 (2010).

[16] D. Monaco, Chern and Fu–Kane–Mele invariants as topological obstructions, in *Advances in Quantum Mechanics: Contemporary Trends and Open Problems*, edited by A. Michelangeli and G. Dell'Antonio (Springer International Publishing, Cham, 2017), pp. 201–222.

[17] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Maximally localized Wannier functions: Theory and applications, Rev. Mod. Phys. **84**, 1419 (2012).

[18] S. D. Swiecicki, Ph.D. thesis, University of Toronto, 2017, available at https://hdl.handle.net/1807/80790.

[19] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[20] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms: Introduction to Quantum Electrodynamics* (John Wiley and Sons, Inc., 1989).

[21] W. P. Healy, *Non-Relativistic Quantum Electrodynamics* (Academic Press, 1982).

[22] R. Resta, Quantum-mechanical position operator in extended systems, Phys. Rev. Lett. **80**, 1800 (1998).

[23] R. D. King-Smith and D. Vanderbilt, Theory of polarization of crystalline solids, Phys. Rev. B **47**, 1651 (1993).

[24] T. Thonhauser, D. Ceresoli, D. Vanderbilt, and R. Resta, Orbital magnetization in periodic insulators, Phys. Rev. Lett. **95**, 137205 (2005).

[25] D. Ceresoli, T. Thonhauser, D. Vanderbilt, and R. Resta, Orbital magnetization in crystalline solids: Multi-band insulators, Chern insulators, and metals, Phys. Rev. B **74**, 024408 (2006).

[26] W. A. Wheeler, L. K. Wagner, and T. L. Hughes, Many-body electric multipole operators in extended systems, Phys. Rev. B **100**, 245135 (2019).

[27] P. T. Mahon, R. A. Muniz, and J. E. Sipe, Microscopic polarization and magnetization fields in extended systems, Phys. Rev. B **99**, 235140 (2019).

[28] S. D. Swiecicki and J. E. Sipe, Gauge-invariant Green function dynamics: A unified approach, Ann. Phys. (NY) **338**, 260 (2013).

[29] S. D. Swiecicki and J. E. Sipe, Linear response of crystals to electromagnetic fields: Microscopic charge-current density, polarization, and magnetization, Phys. Rev. B **90**, 125115 (2014).

[30] P. T. Mahon and J. E. Sipe, Magnetoelectric polarizability: A microscopic perspective, Phys. Rev. Res. **2**, 033126 (2020).

[31] P. T. Mahon and J. E. Sipe, From magnetoelectric response to optical activity, Phys. Rev. Res. **2**, 043110 (2020).

[32] P. T. Mahon, J. G. Kattan, and J. E. Sipe, Polarization and orbital magnetization in Chern insulators: A microscopic perspective, Phys. Rev. B **107**, 115110 (2023).

[33] P. T. Mahon and J. E. Sipe, Electric polarization and magnetization in metals, SciPost Phys. **14**, 058 (2023).

[34] A. H. Duff and J. E. Sipe, Magnetoelectric polarizability and optical activity: Spin and frequency dependence, Phys. Rev. B **106**, 085413 (2022).

[35] A. H. Duff, A. Lau, and J. E. Sipe, Microscopic theory of the magnetic susceptibility of insulators, arXiv:2306.03820.

[36] J. G. Kattan and J. E. Sipe, Multipolar quantum electrodynamics of localized charge-current distributions: Spectral theory and renormalization, Phys. Rev. A **107**, 032820 (2023).

[37] R. Resta, Macroscopic polarization in crystalline dielectrics: The geometric phase approach, Rev. Mod. Phys. **66**, 899 (1994).

[38] A. Malashevich, I. Souza, S. Coh, and D. Vanderbilt, Theory of orbital magnetoelectric response, New J. Phys. **12**, 053032 (2010).

[39] D. Vanderbilt, *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge University Press, 2018).

[40] D. Xiao, J. Shi, D. P. Clougherty, and Q. Niu, Polarization and adiabatic pumping in inhomogeneous crystals, Phys. Rev. Lett. **102**, 087602 (2009).

[41] M. Nakahara, *Geometry, Topology and Physics*, 2nd ed. (CRC Press, 2003).

[42] D. S. Freed and G. W. Moore, Twisted equivariant matter, Ann. Henri Poincaré **14**, 1927 (2013).

[43] R. M. Kaufmann, D. Li, and B. Wehefritz-Kaufmann, Notes on topological insulators, Rev. Math. Phys. **28**, 1630003 (2016).

[44] J. Lee, *Introduction to Smooth Manifolds* (Springer-Verlag, New York, 2012).

[45] G. Panati, Triviality of Bloch and Bloch-Dirac bundles, Ann. Henri Poincaré **8**, 995 (2007).

[46] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Exponential localization of Wannier functions in insulators, Phys. Rev. Lett. **98**, 046402 (2007).

[47] T. Olsen, M. Taherinejad, D. Vanderbilt, and I. Souza, Surface theorem for the Chern-Simons axion coupling, Phys. Rev. B **95**, 075137 (2017).

[48] D. Vanderbilt and R. D. King-Smith, Electric polarization as a bulk quantity and its relation to surface charge, Phys. Rev. B **48**, 4442 (1993).

[49] S. Coh, D. Vanderbilt, A. Malashevich, and I. Souza, Chern-Simons orbital magnetoelectric coupling in generic insulators, Phys. Rev. B **83**, 085108 (2011).

[50] J. Wang, B. Lian, X.-L. Qi, and S.-C. Zhang, Quantized topological magnetoelectric effect of the zero-plateau quantum anomalous Hall state, Phys. Rev. B **92**, 081107(R) (2015).

[51] N. Pournaghavi, A. Pertsova, A. H. MacDonald, and C. M. Canali, Nonlocal sidewall response and deviation from exact quantization of the topological magnetoelectric effect in axion-insulator thin films, Phys. Rev. B **104**, L201102 (2021).

[52] R. S. K. Mong, A. M. Essin, and J. E. Moore, Antiferromagnetic topological insulators, Phys. Rev. B **81**, 245209 (2010).

[53] L. Fu, C. L. Kane, and E. J. Mele, Topological insulators in three dimensions, Phys. Rev. Lett. **98**, 106803 (2007).

[54] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Topological insulators in Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface, Nat. Phys. **5**, 438 (2009).

[55] M. Z. Hasan and C. L. Kane, Colloquium: Topological insulators, Rev. Mod. Phys. **82**, 3045 (2010).

[56] X.-L. Qi and S.-C. Zhang, Topological insulators and superconductors, Rev. Mod. Phys. **83**, 1057 (2011).

[57] C. Lei, S. Chen, and A. H. MacDonald, Magnetized topological insulator multilayers, Proc. Natl. Acad. Sci. USA **117**, 27224 (2020).

[58] T. Morimoto, A. Furusaki, and N. Nagaosa, Topological magnetoelectric effects in thin films of topological insulators, Phys. Rev. B **92**, 085113 (2015).

[59] C. Lei and A. H. MacDonald, Gate-tunable quantum anomalous Hall effects in $MnBi_2Te_4$ thin films, Phys. Rev. Mater. **5**, L051201 (2021).

[60] R. Winkler, *Spin-orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*, 1st ed. (Springer, Berlin, Heidelberg, 2003).

[61] J. J. Sakurai, *Modern Quantum Mechanics*, revised ed. (Addison-Wesley, Reading, 1994).

[62] A. A. Burkov and L. Balents, Weyl semimetal in a topological insulator multilayer, Phys. Rev. Lett. **107**, 127205 (2011).

[63] C. Lei, P. T. Mahon, and A. H. MacDonald, Non-universal surface magnetoelectric response in antiferromagnetic topological insulators, arXiv:2311.17859.

[64] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig, Topological insulators and superconductors: Tenfold way and dimensional hierarchy, New J. Phys. **12**, 065010 (2010).

[65] R. Resta and D. Vanderbilt, Theory of polarization: A modern approach, in *Physics of Ferroelectrics: A Modern Perspective* (Springer, Berlin, Heidelberg, 2007), pp. 31–68.

[66] D. Zhang, M. Shi, T. Zhu, D. Xing, H. Zhang, and J. Wang, Topological axion states in the magnetic insulator $MnBi_2Te_4$ with the quantized magnetoelectric effect, Phys. Rev. Lett. **122**, 206401 (2019).

[67] C.-X. Liu, H. J. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S. C. Zhang, Oscillatory crossover from two-

[68] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Quantum spin Hall effect and topological phase transition in HgTe quantum wells, Science **314**, 1757 (2006).

[69] S. Urazhdin, D. Bilc, S. D. Mahanti, S. H. Tessmer, T. Kyratsi, and M. G. Kanatzidis, Surface effects in layered semiconductors Bi_2Se_3 and Bi_2Te_3, Phys. Rev. B **69**, 085313 (2004).

[70] A. A. Soluyanov and D. Vanderbilt, Wannier representation of $\mathbb{Z}_2$ topological insulators, Phys. Rev. B **83**, 035108 (2011).

[71] E. Prodan, Robustness of the spin-Chern number, Phys. Rev. B **80**, 125327 (2009).

[72] M. Hamilton, *Mathematical Gauge Theory* (Springer International Publishing, 2017).

[73] D. Monaco and G. Panati, Symmetry and localization in periodic crystals: Triviality of Bloch bundles with a fermionic time-reversal symmetry, Acta Appl. Math. **137**, 185 (2015).

[74] T. Ando, Theory of electronic states and transport in carbon nanotubes, J. Phys. Soc. Jpn. **74**, 777 (2005).

[75] P. Marconcini and M. Macucci, The $\boldsymbol{k} \cdot \boldsymbol{p}$ method and its application to graphene, carbon nanotubes and graphene nanoribbons: The Dirac equation, Riv. Nuovo Cimento **34**, 489 (2011).

[76] A. Elçi and E. D. Jones, Some consequences of the closure of the momentum Bloch functions, Phys. Rev. B **34**, 8611 (1986).

[77] M. Cardona and F. H. Pollak, Energy-band structure of germanium and silicon: The $\boldsymbol{k} \cdot \boldsymbol{p}$ method, Phys. Rev. **142**, 530 (1966).

[78] R. Roy, $\mathbb{Z}_2$ classification of quantum spin Hall systems: An approach using time-reversal invariance, Phys. Rev. B **79**, 195321 (2009).

[79] R. Roy, Characterization of three-dimensional topological insulators by two-dimensional invariants, New J. Phys. **12**, 065009 (2010).

[80] G. Rosenberg and M. Franz, Witten effect in a crystalline topological insulator, Phys. Rev. B **82**, 035105 (2010).

[81] R. Jackiw and C. Rebbi, Solitons with fermion number 1/2, Phys. Rev. D: Part. Fields **13**, 3398 (1976).

[82] L. Wu, M. Salehi, N. Koirala, J. Moon, S. Oh, and N. P. Armitage, Quantized Faraday and Kerr rotation and axion electrodynamics of a 3d topological insulator, Science **354**, 1124 (2016).

[83] K. N. Okada, Y. Takahashi, M. Mogi, R. Yoshimi, A. Tsukazaki, K. S. Takahashi, N. Ogawa, M. Kawasaki, and Y. Tokura, Terahertz spectroscopy on Faraday and Kerr rotations in a quantum anomalous Hall state, Nat. Commun. **7**, 12245 (2016).

[84] V. Dziom, A. Shuvaev, A. Pimenov, G. V. Astakhov, C. Ames, K. Bendias, J. Böttcher, G. Tkachov, E. M. Hankiewicz, C. Brüne *et al.*, Observation of the universal magnetoelectric effect in a 3d topological insulator, Nat. Commun. **8**, 15197 (2017).

[85] C. Beenakker, Topological magnetoelectric effect versus quantum Faraday effect, J. Club Condens. Matter Phys. (2016).

[86] N. P. Armitage and L. Wu, On the matter of topological insulators as magnetoelectrics, SciPost Phys. **6**, 046 (2019).

[87] Y. Hatsugai, Chern number and edge states in the integer quantum Hall effect, Phys. Rev. Lett. **71**, 3697 (1993).

[88] J. Böttcher, C. Tutschku, L. W. Molenkamp, and E. M. Hankiewicz, Survival of the quantum anomalous Hall effect in

orbital magnetic fields as a consequence of the parity anomaly, Phys. Rev. Lett. **123**, 226602 (2019).

[89] C. Tutschku, F. S. Nogueira, C. Northe, J. van den Brink, and E. M. Hankiewicz, Temperature and chemical potential dependence of the parity anomaly in quantum anomalous Hall insulators, Phys. Rev. B **102**, 205407 (2020).

[90] G. Jotzu, *Chern Insulators* (2014), Chap. 4.

[91] A. W. W. Ludwig, Topological phases: Classification of topological insulators and superconductors of non-interacting fermions, and beyond, Phys. Scr. **2016**, 014001 (2016).

[92] C. L. Kane and E. J. Mele, $\mathbb{Z}_2$ topological order and the quantum spin Hall effect, Phys. Rev. Lett. **95**, 146802 (2005).

[93] L. Fu and C. L. Kane, Time reversal polarization and a Z_2 adiabatic spin pump, Phys. Rev. B **74**, 195312 (2006).