Dirac equation perspective on higher-order topological insulators ⊘

Frank Schindler ✉ iD

Check for updates

View Online

Export Citation

CrossMark

Dirac equation perspective on higher-order topological insulators

View Online   Export Citation   CrossMark

Frank Schindler[a]

AFFILIATIONS

Princeton Center for Theoretical Science, Princeton University, Princeton 08544, New Jersey, USA

Note: This paper is part of the Special Topic on Topological Materials and Devices.
[a]**Author to whom correspondence should be addressed:** schindler@princeton.edu

ABSTRACT

In this Tutorial, we pedagogically review recent developments in the field of non-interacting fermionic phases of matter, focusing on the low-energy description of higher-order topological insulators in terms of the Dirac equation. Our aim is to give a mostly self-contained treatment. After introducing the Dirac approximation of topological crystalline band structures, we use it to derive the anomalous end and corner states of first- and higher-order topological insulators in one and two spatial dimensions. In particular, we recast the classical derivation of domain wall bound states of the Su–Schrieffer–Heeger (SSH) chain in terms of crystalline symmetry. The edge of a two-dimensional higher-order topological insulator can then be viewed as a single crystalline symmetry-protected SSH chain, whose domain wall bound states become the corner states. We never explicitly solve for the full symmetric boundary of the two-dimensional system but instead argue by adiabatic continuity. Our approach captures all salient features of higher-order topology while remaining analytically tractable.

11 August 2023 09:50:18

I. INTRODUCTION

We begin by introducing some basic notions of condensed matter physics, topological phases, and topological crystalline insulators. We illustrate symmetry-protected topological phases with a toy model in zero spatial dimensions. Various aspects of the material in this Tutorial are also covered in Refs. 1–6.

Units are chosen so that the speed of light and Planck's constant are dimensionless $c = \hbar = 1$.

A. Basic notions

Whereas particle physics aims to decompose the macroscopic material world into its elementary constituents, the goal of condensed matter physics is to understand how the properties of materials emerge from their microscopic degrees of freedom. In principle, this question can be posed not only for materials but also for any many-body system that is comprised of a macroscopic number of particles. However, condensed phases of matter are particularly interesting, because they are not *adiabatically* (that is, not without a phase transition) connected to the featureless phase at infinite temperature. This characterization implies that their behavior cannot be simply understood as the sum of their parts:[7]

for instance, the elementary particles that make out a condensed matter system often reorganize into collective excitations that behave like particles themselves, the so-called quasi-particles, and yet have markedly different properties, such as a different mass or even different exchange statistics.[8] One of the basic insights of condensed matter physics is that there is a multitude of condensed phases beyond the rather coarse classification of matter into solids, liquids, and gases, with particularly interesting cases at very low temperatures where quantum mechanical effects come into play.

A crystal is an example of a condensed phase that consists of atomic nuclei and electrons, which for our purposes are its elementary constituents. We assume that the atomic nuclei (often just called atoms) form a crystalline lattice in D spatial dimensions and move so slowly in comparison to the electrons—because of their much larger mass—that we can essentially treat them as a fixed background. Since the atomic lattice breaks spatial rotation and translation symmetry, the crystal realizes a condensed phase that is not adiabatically connected to the infinite-temperature phase (which is perfectly symmetric). From now on, we only focus on the electronic part, which at low temperatures is governed by quantum mechanics[9] and determines whether the crystal is an insulator or a conductor, among many other measurable properties.

Since electrons are fermionic particles obeying the Pauli exclusion principle, a conglomerate of electrons such as the one populating a crystal cannot be understood as the sum of its parts even when there are no further interactions: in contrast to bosons, electrons cannot all occupy the same single-particle eigenstate. This property already enables a great number of interesting electronic phases of matter, and so we will completely neglect further interactions between the electrons such as the Coulomb interaction. It turns out that this approximation is often surprisingly good and applies to many real materials.

Quantum properties dominate over thermal fluctuations at low temperatures, and therefore we will consider systems at zero temperature for simplicity, allowing us to focus on ground state properties only. N non-interacting electrons then fill up the N-lowest energy eigenstates of the single-particle Hamiltonian due to the Pauli exclusion principle. To be precise, let $c_\alpha^\dagger$ be the creation operator of an electron in the single-particle eigenstate labeled by α. Fermionic exchange statistics are implemented by the anticommutation relations

$$\{c_\alpha^\dagger, c_\beta\} = 2\delta_{\alpha\beta}, \quad \{c_\alpha^\dagger, c_\beta^\dagger\} = \{c_\alpha, c_\beta\} = 0. \quad (1)$$

We can write a generic non-interacting many-body Hamiltonian that preserves electronic charge as

$$H = \sum_\alpha (\epsilon_\alpha - \mu) c_\alpha^\dagger c_\alpha, \quad (2)$$

where ϵ_α are the eigenvalues of the single-particle Hamiltonian $\mathcal{H}$, and μ is the chemical potential. The N-particle ground state is then

given by

$$|\Omega\rangle = \frac{1}{\mathcal{N}} \prod_{\epsilon_\alpha < \mu} c_\alpha^\dagger |0\rangle, \quad (3)$$

where $\mathcal{N}$ is a normalization factor, $|0\rangle$ denotes the electronic vacuum, and each α appears exactly once as long as $\epsilon_\alpha < \mu$. $|\Omega\rangle$ is called a Slater-determinant state, because the fermionic anticommutation relations imply that its coefficients in a given basis are given by a determinant.[10]

The total number of single-particle states scales linearly with the system volume V. An insulator is defined by an energy gap between the ground state and the lowest excited state that survives in the thermodynamic limit, where we take $N \to \infty$, $V \to \infty$ while keeping the particle density N/V constant. The condition for a gap in the many-body spectrum of H translates to the absence of single-particle eigenstates with energies at the chemical potential μ. A gap implies that a small external electric field[11] cannot lift electrons to excited single-particle states with nonzero net momentum and, therefore, cannot create a current. In contrast, a conductor (a metal) is characterized by the absence of a gap. A (second-order) phase transition, which features a diverging correlation length and a discontinuous change in some ground state properties, implies a gap closing.

B. Topological phases

A condensed matter system realizes a gapped topological phase when its ground state cannot be continuously deformed to a trivial reference state adiabatically, that is, without closing the energy gap to the first excited state. In the context of a crystal, a trivial insulator can be defined as an *atomic limit* where all electrons are tightly bound to the atoms, that is, as a state that is adiabatically connected to a Slater product over single-electron wavefunctions whose weight is exponentially localized around the atomic sites in position space.[12] For this definition, it is important to assume periodic boundary conditions so that boundary effects do not play a role and only bulk crystal properties are probed. We refer to the gap of a crystal with periodic boundary conditions as the *bulk gap* and adiabatic manipulations as those preserving it. There may be multiple trivial phases: for instance, in a crystal containing multiple types of atoms (such as rock salt and NaCl), the states formed by electrons localized around each atom are equally valid atomic limits (in real materials, however, at most one atomic limit is energetically favored).

The concept of topological phases can be refined by demanding that symmetry preservation is necessary to prevent an adiabatic path to a trivial phase, one then speaks of a *symmetry-protected topological phase* (a pedagogical review is provided in Ref. 13) (see Fig. 1). For example, it may be possible to adiabatically deform the ground state of all electrons in a crystal to an atomic limit that breaks the crystalline inversion symmetry (which maps a coordinate x in the crystal lattice to $-x$) but to no other atomic limit. The state then realizes a *topological crystalline insulator*[14] protected by inversion symmetry. The adjective "crystalline" here just means that a crystalline symmetry, in addition to the bulk gap, is necessary to stabilize the topological phase. When the symmetry is not
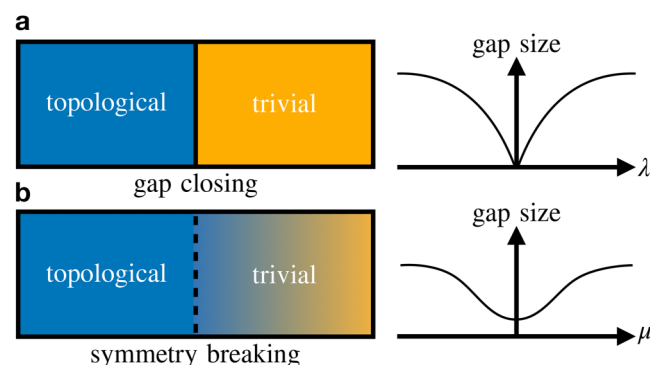

FIG. 1. Schematic phase diagrams of a symmetry-protected topological phase. (a) When all symmetries are maintained, interpolating from the topological to the trivial parameter regime necessitates a gap closing and thereby a phase transition. The right panel shows how the (absolute value of the) gap evolves with a tuning parameter λ, where the Hamiltonian $H(\lambda)$ is symmetric for all choices of λ. (b) Upon symmetry breaking, a smooth (adiabatic) transition without gap closing becomes possible. The right panel shows how the gap evolves with a tuning parameter μ, where $H(\mu)$ is symmetric for $\mu < 0$ but the symmetry is broken for $\mu > 0$.

crystalline, such as time reversal, the convention is to simply call the system a topological insulator. Historically, topological insulators protected by time-reversal symmetry came before topological crystalline insulators, here we nevertheless focus on the latter class of systems because they are conceptually much simpler. There is only one known example of a non-interacting topological phase that requires no symmetries at all for its protection: the Chern insulator or the integer quantum Hall effect.[15]

To illustrate these concepts, consider a zero-dimensional toy model of a spinless electron occupying one of the two orbitals in an atom, an s and a p orbital, with single-particle Hamiltonian

$$\mathcal{H} = \alpha\sigma_x + \beta\sigma_y + \gamma\sigma_z, \quad \epsilon_\alpha = \pm\sqrt{\alpha^2 + \beta^2 + \gamma^2}, \quad (4)$$

where σ_a, $a = x, y, z$, are the 2×2 Pauli matrices, and there are only two energy levels labeled by their sign $\alpha = \pm$. Let inversion symmetry be represented by $\mathcal{I} = \sigma_z$ so that the s orbital corresponds to the state $|s\rangle = (1, 0)^{\mathrm{T}}$, while the p orbital corresponds to $|p\rangle = (0, 1)^{\mathrm{T}}$ (T takes the transpose). For $|\gamma| \gg |\alpha|, |\beta|$, for instance, the spectrum of $\mathcal{H}$ is gapped, and therefore also that of the corresponding many-body Hamiltonian H in Eq. (2) at chemical potential $\mu = 0$. Let us now define $|s\rangle$ as the trivial reference state, it is the ground state of $\mathcal{H}$ for $\gamma = -\infty$.[16] Then, the system realizes an inversion symmetry-protected topological phase for large positive γ: as long as inversion symmetry is not enforced, it is possible to tune γ from positive to negative values without a gap closing, which happens only when $\alpha^2 + \beta^2 + \gamma^2 = 0$, as long as α and β are kept non-zero. On the other hand, enforcing inversion symmetry via the commutation relation $[\mathcal{H}, \mathcal{I}] = 0$ implies that α and β vanish, which in turn ensures a gap closing along the interpolation $\gamma(t) = 1 - 2t$, $t \in [0, 1]$ that exchanges $|p\rangle$ with $|s\rangle$ as the lowest energy eigenstate of $\mathcal{H}$. We therefore find that there are no topological insulators in zero spatial dimensions that are protected by the bulk gap alone, but that there do exist topological crystalline insulators protected by inversion symmetry.

Since the vacuum can be viewed as a trivial phase,[17] the boundary of a topological phase generically hosts a phase transition (a phase transition must occur to go between topological and trivial) and thereby a gap closing, as long as it respects the protecting symmetries.[18] The most experimentally relevant kinds of topological insulators are therefore characterized by an insulating (gapped) bulk and a conducting (gapless) boundary. If the protecting symmetry is not crystalline, all boundaries preserve it by default and are therefore gapless, implying that the D-dimensional insulating bulk is surrounded by a $(D-1)$ dimensional conducting surface—this is the celebrated *bulk-boundary correspondence*. Topological crystalline insulators, on the other hand, generically have surfaces that do not preserve the relevant crystalline symmetry on their own and can therefore be gapped. The crystalline symmetry may still enforce the presence of $(D-k)$-dimensional gapless states with $k > 1$, thereby giving rise to *higher-order topological insulators*, where k is referred to as the order.

In this Tutorial, our aim is to give a pedagogical introduction to first- and higher-order topological crystalline insulators from the point of view of their effective low-energy description. Our main tool will be the Dirac equation, which describes elementary electrons in particle physics. In a particularly tractable realization of the principle of emergence discussed at the beginning of this introduction, we will see how variants of the Dirac Hamiltonian arise in the low-energy description of topological (crystalline) insulators (see also Chap. V.5 of Ref. 19). We will then explain how they can be used to argue for the stability of (higher-order) topological phases and, in particular, for the presence of protected boundary states.

II. DIRAC APPROXIMATION OF TOPOLOGICAL CRYSTALLINE INSULATORS

In this section, we introduce the Dirac equation in the context of the tight-binding description of electron movement in crystals. We also fix some basic notation and conventions that will be extensively used later on. For simplicity, we only treat one-dimensional (1D) systems in this section, the generalization to higher dimensions being straightforward (it mostly involves making all position and momentum coordinates vector-valued and replacing products between them by dot products).

A. Tight-binding description

Consider a 1D crystal.[20] We may model it as an infinitely long chain of repeated atom arrangements, the unit cells. In the tight-binding approximation, each unit cell hosts a finite number of sites that can be occupied by electrons. Denote by $c_{xi}^\dagger$ the operator that creates an electron at site i, $i = 1, \ldots, M$, in the unit cell at position x, $x = 1, \ldots, L$ (we take x to be dimensionless so that L is just the total number of unit cells in the sample). Assuming only nearest-neighbor hoppings between the sites (in a microscopic treatment, these hoppings would be derived from transition matrix elements of the microscopic electronic Hamiltonian between states that are well-localized on the individual sites), the Hamiltonian takes the form

$$H = \sum_{xij} h_{ij}c_{xi}^\dagger c_{xj} + t_{ij}c_{x+1,i}^\dagger c_{xj} + t_{ij}^\star c_{xj}^\dagger c_{x+1,i}. \quad (5)$$

(Here and in the following, we will set the chemical potential μ to zero for simplicity.) By construction, H preserves a global $U(1)$ symmetry that is represented by

$$c_{xi}^\dagger \rightarrow e^{i\phi}c_{xi}^\dagger, \quad c_{xi} \rightarrow e^{-i\phi}c_{xi}, \quad (6)$$

with $\phi \in (0, 2\pi]$ an arbitrary phase. The infinitesimal generator of this symmetry is the total charge,

$$Q = \sum_{xi} c_{xi}^\dagger c_{xi}, \quad (7)$$

so that we have $e^{i\phi Q}c_{xi}^\dagger e^{-i\phi Q} = e^{i\phi}c_{xi}^\dagger$. Let $|\Omega\rangle$ be the many-body ground state of H and perform a $U(1)$ transformation by the angle 2π. Evidently, such a transformation must be equal to the identity transformation, and so the ground state, which transforms as $|\Omega\rangle \rightarrow e^{i2\pi Q}|\Omega\rangle$, must have integer charge $Q \in \mathbb{Z}$. We will see in Sec. III D how the boundary of a topological phase can break this

quantization constraint and host fractional charges at the expense of a topologically nontrivial bulk.

Let us for now return to exploring the general properties of the Hamiltonian H. Assuming periodic boundary conditions, we can exploit its discrete translational symmetry and partially diagonalize it by introducing the crystal momenta $k = (1, \ldots, L)\frac{2\pi}{L}$, and the correspondingly Fourier-transformed creation operators

$$c_{ki}^\dagger = \sum_x e^{-ikx} c_{xi}^\dagger, \qquad c_{xi}^\dagger = \frac{1}{L} \sum_k e^{ikx} c_{ki}^\dagger. \qquad (8)$$

Later on, we will often implicitly take $L \to \infty$, in which case all sums over k should be interpreted as integrals over the first Brillouin zone $k \in (0, 2\pi)$. Equation (5) becomes

$$H = \frac{1}{L} \sum_{kij} \mathcal{H}_{ij}(k) c_{ki}^\dagger c_{kj}, \qquad (9)$$

where we introduced the Bloch Hamiltonian

$$\mathcal{H}_{ij}(k) = h_{ij} + t_{ij} e^{ik} + t_{ji}^* e^{-ik}. \qquad (10)$$

The eigenvalues of $\mathcal{H}(k)$ give the single-particle energy spectrum as a function of crystal momentum k, the so-called electronic band structure. That is, in the case of a periodic one-dimensional crystal, the eigenstate label α that appears in Eqs. (2) and (3) is a composite label $\alpha = (k, n)$ that specifies the crystal momentum k together with the index of the nth eigenstate of $\mathcal{H}(k)$, where $n = 1, \ldots, M$ (there are M bands). [Equation (4) can be viewed as a Bloch Hamiltonian in zero dimensions, where there is no momentum k, and with $M = 2$ sites in the unit cell.]

B. Two-band models and topology in the absence of symmetry

In the remainder of this section, we consider the simplest non-trivial case, where the unit cell has two atomic sites, $M = 2$. We can then expand the Bloch Hamiltonian $\mathcal{H}(k)$ in terms of the identity matrix and the Pauli matrices (something that is possible for any Hermitian 2×2 matrix),

$$\mathcal{H}(k) = \epsilon_0(k) \dotdiv + \boldsymbol{d}(k) \cdot \boldsymbol{\sigma}, \qquad (11)$$

where we used scalar product notation and introduced the real-valued vector $\boldsymbol{d}(k)$ with components $d_a(k)$, $a = x, y, z$. The spectrum of $\mathcal{H}(k)$ can be easily inferred by calculating $[\mathcal{H}(k) - \epsilon_0(k)\dotdiv]^2$, which turns out to be proportional to the identity matrix $\dotdiv$. It is given by

$$\epsilon_\pm(k) = \epsilon_0(k) \pm \sqrt{\boldsymbol{d}(k) \cdot \boldsymbol{d}(k)}. \qquad (12)$$

We see that the condition for an energy gap between the two bands $\epsilon_+(k)$ and $\epsilon_-(k)$ translates to the inequality $d_a(k) \neq 0$, $a = x, y, z$. Moreover, we can always adiabatically tune $\epsilon_0(k) \to 0$ without closing the gap between the two energy levels, and so all two-band insulators that only differ by the choice of $\epsilon_0(k)$ are topologically equivalent. We set $\epsilon_0(k) = 0$ in the following.

We now come to topology. Due to the above criterion for an energy gap, the space of all one-dimensional two-band insulators is isomorphic to the space of all vector-valued functions $\boldsymbol{d}(k)$ that do not map any momentum $k \in (0, 2\pi)$ to the zero vector $\boldsymbol{0}$, that is, $\boldsymbol{d}(k) \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$. Since the Brillouin zone has the topology of a circle, or one-sphere S^1 (after all, the replacement $k \to k + 2\pi$ has no effect), a Bloch Hamiltonian is characterized by a collection of vectors $\boldsymbol{d}(k)$ that satisfies $\boldsymbol{d}(2\pi) = \boldsymbol{d}(0)$. It is easy to see that all such loops in $\mathbb{R}^3$ can be contracted to a point $\boldsymbol{d}(k) = \boldsymbol{d}^*$ without ever touching the origin $\boldsymbol{0}$. Comparing with Eq. (10), this implies that any constant $\boldsymbol{d}^*$ gives rise to a corresponding atomic limit Hamiltonian $\mathcal{H}(k) = \mathcal{H}^*$ that is independent of k and does not feature any hoppings between unit cells.[21] We deduce that all one-dimensional two-band insulators are topologically trivial absent symmetries.

C. Crystalline topology by inversion symmetry

Here and in the following, we will specify only a single kind of trivial reference phase: the atomic limits with a Bloch Hamiltonian independent of k. There may be other atomic limits (remember from Sec. I B that atomic limits are simply insulators whose ground state can be decomposed into exponentially localized electronic wavefunctions). For instance, the Hamiltonian belonging to an atomic limit where all electrons are exponentially localized around atomic sites at the boundary of the unit cell necessarily has to be k-dependent. From our point of view, we will treat such phases as topological as long as they cannot be adiabatically transformed into a trivial reference phase, similar to the more canonical examples of topological phases that do not have a representation in terms of exponentially localized electrons at all.[22]

In order to stabilize nontrivial 1D insulators, we need to incorporate symmetries. As mentioned in Sec. I B, the simplest example of a crystalline symmetry is given by spatial inversion that takes x to $-x$. When each unit cell contains two atoms that are mapped onto each other by inversion, it is represented by $\mathcal{I} = \sigma_x$. At the same time, inversion symmetry flips a crystal momentum k to $-k$. The condition for inversion symmetry in 1D is therefore that the Bloch Hamiltonian satisfies

$$\mathcal{I}\mathcal{H}(k)\mathcal{I}^\dagger = \mathcal{H}(-k). \qquad (13)$$

[This condition is equivalent to the requirement that the many-body inversion operator commutes with the many-body Hamiltonian in Eq. (2).] In our two-band example, this constraint enforces that $d_y(k) = -d_y(-k)$ and $d_z(k) = -d_z(-k)$ are odd functions of k, while $d_x(k) = d_x(-k)$ is even. Moreover, we find that the crystal momenta $\bar{k} = 0, \pi$, which satisfy $\bar{k} = -\bar{k} \bmod 2\pi$, are special in that, at these momenta, we have the commutation relation $[\mathcal{H}(\bar{k}), \mathcal{I}] = 0$.

These constraints are sufficient to protect topological crystalline insulators: Consider the many-body ground state at zero chemical potential. In our example [Eq. (11)], this situation corresponds to occupying a single site per unit cell with an electron (half-filling), and therefore, a single occupied energy band (each band has L available crystal momenta that may be occupied, and there are L unit cells with two sites each).[23] The occupied band is

11 August 2023 09:50:18

composed of Bloch states $|u(k)\rangle$ that satisfy

$$\mathcal{H}(k)|u(k)\rangle = \epsilon_-(k)|u(k)\rangle. \qquad (14)$$

At the high-symmetry momenta $\bar{k}$, these states can furthermore be chosen as eigenstates of inversion symmetry,

$$\mathcal{I}|u(\bar{k})\rangle = \lambda(\bar{k})|u(\bar{k})\rangle. \qquad (15)$$

Due to $\mathcal{I}^2 = 1$, we have $\lambda(\bar{k}) = \pm 1$—the inversion eigenvalues are quantized to only two allowed values (this quantization is a key ingredient for topology). Now, an insulating state where the two occupied inversion eigenvalues differ, $\lambda(0) = -\lambda(\pi)$, is clearly not adiabatically connected to any trivial atomic limit: all constant Bloch Hamiltonians $\mathcal{H}(k) = \mathcal{H}^*$ have equal occupied inversion eigenvalues, and these cannot be changed without either closing the gap to the unoccupied band with energy $\epsilon_+(k)$ or breaking inversion symmetry. As an example, consider the inversion-symmetric Bloch Hamiltonian for $[d_x(k), d_y(k), d_z(k)] = [\Delta + \cos k, \sin k, 0]$, which reads

$$\mathcal{H}_{\text{SSH}}(k) = (\Delta + \cos k)\sigma_x + \sin k\sigma_y. \qquad (16)$$

This Hamiltonian is in the topological phase for $|\Delta| < 1$, and in the trivial phase otherwise. [Comparing with Eq. (12), the energy dispersion reads

$$\epsilon_\pm(k) = \pm\sqrt{\Delta^2 + 2\Delta\cos k + 1}, \qquad (17)$$

and so there is a gap closing at $k = \pi$ for $\Delta = 1$ and one at $k = 0$ for $\Delta = -1$.] In fact, $\mathcal{H}_{\text{SSH}}(k)$ is closely related to the famous Su–Schrieffer–Heeger (SSH) model of polyacetylene[24] (for a derivation, see Sec. III). We will show in Sec. III B that the topological phase is characterized by the presence of end states in an open chain geometry. The model in Eq. (16) provides the archetypical model of topological insulators, from which many more phases can be constructed by dimensional and symmetry enhancement.

D. Derivation of the Dirac Hamiltonian

We have found that there are 1D topological crystalline insulators protected by inversion symmetry. To explore their physical properties with respect to boundaries, we would like to solve for the spectrum and eigenstates of Eq. (16) in geometries with open rather than periodic boundary conditions. This can be done numerically via exact diagonalization, but it is more illuminating to find an approximate analytical solution, in particular, in order to determine which properties are specific to the model at hand, and which are general features of the topological phase.[25]

In order to make the problem tractable, we first need to determine the salient feature distinguishing the topological phase from the trivial phase. Fixing $\Delta < 1$ for the trivial reference phase (that is, the atomic limit where all electrons are tightly bound in the center of each unit cell as a symmetric, even-parity superposition), this feature is clearly the inversion eigenvalue at $k = 0$: in the trivial phase, we have $\lambda(0) = \lambda(\pi) = 1$, while in the topological phase we have $\lambda(0) = -\lambda(\pi) = -1$ [remember that $\lambda(\bar{k})$ is the eigenvalue of inversion symmetry $\mathcal{I}$ in the occupied, lower-energy eigenstate of $\mathcal{H}_{\text{SSH}}(k)$]. To capture the topological phase, it is therefore sufficient to perform a first-order Taylor expansion of $\mathcal{H}_{\text{SSH}}(k)$ around $k = 0$, arriving at the Hamiltonian

$$\mathcal{H}_{\text{Dirac}}(k) = m\sigma_x + k\sigma_y, \qquad (18)$$

where we identified $m = \Delta + 1$. This is the Hamiltonian of a Dirac fermion with mass m that propagates in $1 + 1$ spacetime dimensions.[26] (See Fig. 2 for a comparison of the tight-binding and Dirac energy band spectrum.) It is somewhat curious that the relativistic Dirac equation appears in the context of an evidently non-relativistic (non-Lorentz-invariant) crystal. Nevertheless, the Dirac approximation (also called $\mathbf{k} \cdot \mathbf{p}$ approximation) can be studied in its own right for any tight-binding model and is particularly useful
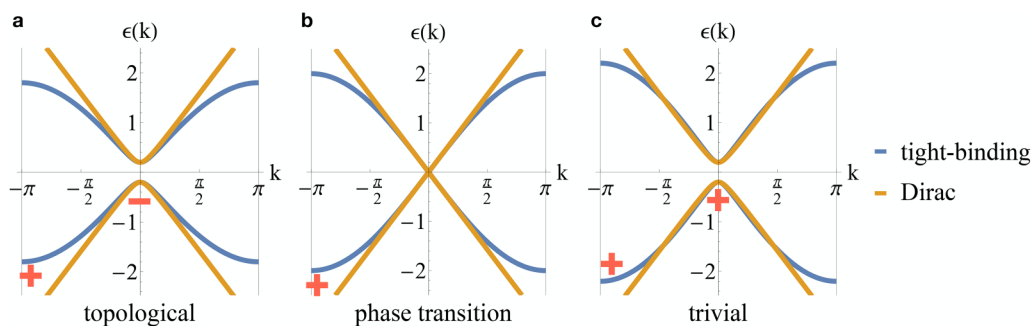

FIG. 2. Dirac approximation of a 1D inversion-symmetric tight-binding model at its topological phase transition. The original electronic structure [Eq. (16)] is shown in blue, while the Dirac spectrum [Eq. (18)] is shown in orange. The signs of the occupied inversion eigenvalues, when well-defined, are indicated in red (at the gap closing, we cannot unambiguously fix the sign). (a) The topological regime ($m < -1$) is characterized by inversion eigenvalues that have opposite sign at $k = 0$ and $k = \pi$. (b) At the phase transition ($m = 1$), the inversion eigenvalue at $k = 0$ is exchanged. (c) The trivial regime ($m > -1$) is characterized by equal occupied inversion eigenvalues. Importantly, the Dirac approximation reproduces the sign switch at $k = 0$, all other details of the band structure are unimportant from the point of view of crystalline topology.

in topological band theory: topological properties are insensitive to the non-universal details of the Bloch Hamiltonian that we discarded by approximating around $k = 0$.[27]

Note that by virtue of the Taylor expansion, the momentum variable k loses the Brillouin zone periodicity property it previously enjoyed. It should therefore be properly seen as taking values in $k \in (-\infty, \infty)$ (additionally, we may assume the thermodynamic limit $L \to \infty$ so that k becomes a continuous variable). However, large values of k are unphysical (the Taylor expansion breaks down), we will therefore only use the predictions of the Dirac approximation in a small window around $k = 0$ and deduce the behavior at the remaining momenta by arguments of adiabatic continuity. The advantage that the Dirac approximation has over the original tight-binding model is that it effectively "forgets" the microscopic structure of the crystalline lattice and lives in continuous space:[28] in order to go to real space (which we need to study finite sample geometries and boundary effects), we can just make the canonical replacement $k \to -i\partial/\partial x$.

In the following, we will use the Dirac equation approach to capture the boundary physics of the one-dimensional SSH model as well as that of a higher-order topological phase in two dimensions. The general strategy will be the same,

1. Write down the tight-binding model.
2. Identify the momentum at which the gap closes between trivial and topological parameter regimes.
3. Perform a Taylor expansion around this momentum to first order, thus arriving at a Dirac-type Hamiltonian.
4. Use the Dirac framework to derive the boundary physics of the respective topological phase.

III. CRYSTALLINE TOPOLOGY IN THE SU–SCHRIEFFER–HEEGER MODEL

In this section, we introduce the inversion-symmetric Su–Schrieffer–Heeger Hamiltonian as the elementary topological insulator in one spatial dimension and use its Dirac approximation (derived in Sec. II D) to study its anomalous boundary physics.

While previous tutorials have discussed the SSH model as a topological phase protected by chiral symmetry,[2,4,29] here we instead focus its role as the simplest topological crystalline insulator, protected by inversion symmetry alone.[5,30]

A. Motivation of the Hamiltonian

Consider polyacetylene, a polymer chain of C_2H_2 molecules. To minimize potential energy, the structure exhibits an alternation of weak and strong bonds between the two carbon atoms of the unit cell, A and B, which effectively each provide one atomic site that electrons can occupy [see Fig. 3(a)]. At half-filling, there is therefore one electron per unit cell. The Hamiltonian contains two terms: the weak, intra-unit cell hopping t that connects A with B, and the strong, inter-unit cell hopping t' that connects B with the A atom in the next unit cell. Inversion symmetry maps one carbon atom to another and should therefore again be represented by σ_x.
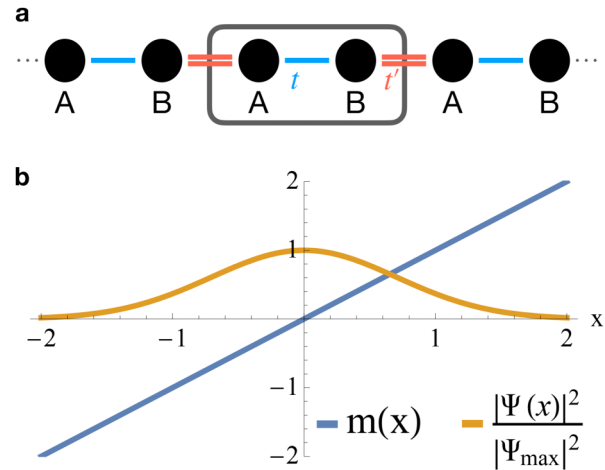

FIG. 3. Domain wall bound states in the Su–Schrieffer–Heeger model. (a) The unit cell contains two atomic sites, A and B, with alternating hopping amplitudes $t \neq t'$ for intra- and inter-unit cell hoppings. (b) A mass domain wall in the Dirac approximation of the SSH model [Eq. (18)], modeled by a linear mass profile $m(x) = x$ and shown in blue, gives rise to an exponentially localized domain wall bound state, shown in orange.

Equation (10) becomes

$$
\begin{aligned}
\mathcal{H}_{\text{SSH}}(k) &= \begin{pmatrix} 0 & t \\ t & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ t' & 0 \end{pmatrix} e^{ik} + \begin{pmatrix} 0 & t' \\ 0 & 0 \end{pmatrix} e^{-ik} \\
&= (t + t'\cos k)\sigma_x + t'\sin k\,\sigma_y \\
&= (\Delta + \cos k)\sigma_x + \sin k\,\sigma_y,
\end{aligned} \tag{19}
$$

where in the last line we identified $\Delta = t$ and set $t' = 1$ in order to recover Eq. (16). We have seen in Sec. II that the spectrum of $\mathcal{H}_{\text{SSH}}(k)$ is fully gapped when $|\Delta| \neq 1$ or equivalently when $|t| \neq |t'|$ (when there is a mismatch in bond strength). Furthermore, we identified the regime $|\Delta| < 1$ as topological and derived the effective Dirac Hamiltonian, Eq. (18), that captures the sign flip of inversion eigenvalues at $k = 0$ as Δ is tuned. For convenience, we reproduce the Dirac Hamiltonian here,

$$
\mathcal{H}_{\text{Dirac}}(k) = m\sigma_x + k\sigma_y, \tag{20}
$$

Keeping in mind $m = \Delta + 1$ and that we fixed $\Delta < -1$ as the trivial reference phase, we can identify the parameter range $m < 0$ as trivial and $m > 0$ as topological in the Dirac approximation.

B. Derivation of the domain wall bound state

The Dirac model $\mathcal{H}_{\text{Dirac}}(k)$, just like the original tight-binding Hamiltonian $\mathcal{H}_{\text{SSH}}(k)$, is gapped at all values of k, and therefore bulk-insulating. To study the effect of boundaries in the topological phase, we need to model an interface between the trivial and topological regimes. This is most easily achieved by assigning a position-dependence to the mass $m \to m(x) = mx$, with the

convention $m > 0$, so that $m(x)$ is negative (trivial) in the left half of the system and positive (topological) in the right half. The particular form of the real space dependence beyond such a sign flip is arbitrary—we might as well have chosen $m(x) = m\,\mathrm{sign}(x)$, for instance. The reason is that we are only interested in topological properties that survive continuous deformations on either side of the gap-closing point at $x = 0$. The only notable effect of our particular choice $m(x) = mx$ is that it is differentiable and so gives smoother results than, for instance, $m(x) = m\,\mathrm{sign}(x)$.

We conclude that the following real-space Hamiltonian effectively describes the interface between two SSH chains, one trivial and one topological,

$$\mathcal{H}_{\mathrm{bdry}} = \sigma_x m x - \mathrm{i}\sigma_y \frac{\partial}{\partial x}, \tag{21}$$

where we expressed $k = -\mathrm{i}\partial/\partial x$ as an operator acting on real space, because its eigenvalue k is not anymore a good quantum number [any non-constant $m(x)$ breaks translation symmetry].

Recall the heuristic argument from Sec. I B that posited that the boundary between a topological and a trivial phase hosts a phase transition, characterized by a gapless energy spectrum. [This argument can be made mathematically rigorous,[31] here we will find it more illuminating to derive it in the Dirac approximation context.] Guided by this intuition, we will attempt to solve for zero-energy states at the interface, which are the two-component wavefunctions $|\Psi(x)\rangle = [\psi(x), \phi(x)]^{\mathrm{T}}$ that satisfy the equation

$$\mathcal{H}_{\mathrm{bdry}}|\Psi(x)\rangle = 0 \Leftrightarrow \frac{\partial}{\partial x}|\Psi(x)\rangle = -\sigma_z m x |\Psi(x)\rangle. \tag{22}$$

There is only one normalizable solution to this equation

$$|\Psi(x)\rangle = \frac{1}{\mathcal{N}} e^{-\frac{1}{2}mx^2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{23}$$

where $\mathcal{N} = (\pi/m)^{\frac{1}{4}}$ is a normalization factor. [The solution that multiplies the vector $(0, 1)^{\mathrm{T}}$ comes with a plus sign in the exponential and is therefore not normalizable.] In conclusion, we find a single zero-energy state that is exponentially localized to the boundary of the topological phase. Its wavefunction decay is shown in Fig. 3(b).

C. Protection by sublattice symmetry: End zeromodes

What remains is to interpret the zero-energy end state from the point of view of topology. At this point, it is illuminating to first recount the historical point of view, according to which the SSH model realizes a topological phase protected by a local rather than a crystalline symmetry.[4] [These conceptually distinct interpretations of topology are both consistent with the Hamiltonian $\mathcal{H}_{\mathrm{SSH}}(k)$, that is, there is more than one symmetry that prevents adiabatic transformations of the SSH ground state to the trivial reference state.] This local symmetry is a sublattice symmetry $\mathcal{C} = \sigma_z$ (also known as chiral symmetry) of $\mathcal{H}_{\mathrm{SSH}}(k)$, and it is

realized via

$$\mathcal{C}\mathcal{H}_{\mathrm{SSH}}(k)\mathcal{C}^\dagger = -\mathcal{H}_{\mathrm{SSH}}(k). \tag{24}$$

This relation enforces that the Hamiltonian only couples the two inequivalent sites of the unit cell (the two sublattices) to each other but not to themselves (as would be the case for local potentials or longer-range hoppings). [Somewhat confusingly, on the level of the Bloch Hamiltonian, sublattice symmetry implies anticommutation rather than commutation with $\mathcal{C}$. It is nevertheless a valid symmetry of the physical system, because it commutes with the many-body Hamiltonian, Eq. (2)—the two minus signs, one deriving from the anti-commutation, and another from the fermionic exchange statistics, cancel.]

Sublattice symmetry is preserved by the Dirac approximation and the interface geometry in Eq. (21). Crucially, it implies that for every eigenstates of $\mathcal{H}_{\mathrm{bdry}}$ at energy E, with $|E| > 0$, there exists a partner eigenstate at energy $-E$. Only a state at $E = 0$ may be unpaired. But this constraint ensures that the single zero-energy mode that we derived in Eq. (23) cannot be moved away from $E = 0$ by any local perturbation: a single eigenstate cannot continuously split into two. The edge of a nontrivial SSH insulator therefore hosts a zero-energy bound state that is topologically protected by the bulk gap and sublattice symmetry.[32]

D. Protection by inversion symmetry: Filling anomaly

Let us now return to crystalline symmetries. Whereas it is difficult to guarantee sublattice symmetry in realistic systems, inversion symmetry is a property of many naturally arising crystal structures. We understood in Sec. II that the SSH model also realizes an inversion symmetry-protected topological insulator and derived the presence of an end state in Eq. (23). It is then natural to ask what properties of the end state are universal and guaranteed by inversion symmetry alone, that is, not specific to the particular choice of Hamiltonian $\mathcal{H}_{\mathrm{SSH}}(k)$. Clearly, absent sublattice symmetry there is no preferred notion of "zero energy,"[33] so that lying at zero-energy cannot be a universal property of the domain wall bound state.

At this point, it is important to note that in order to connect bulk crystalline topology to boundary states, we must study a boundary termination that preserves inversion symmetry. Our choice of the mass profile $m(x)$ from before explicitly broke inversion symmetry—we only modeled one (the left) end of the topological phase. In order to restore inversion symmetry, we must also take into account the second zero-energy state that is located at the opposite (the right) end of the sample. The inversion-symmetric and finite system will then have two zero-energy states, but as explained above, nothing prevents the addition of a local edge potential to $\mathcal{H}_{\mathrm{SSH}}(k)$ that shifts them to finite energies (at least as long as the local potential is the same on the left and right end of the sample, such a manipulation respects inversion symmetry). The property that remains invariant is therefore not the energy of the end states, but their filling: imagine that we gradually tune a finite system from trivial to topological in an $\mathcal{I}$-symmetric manner. This

process can, for instance, be modeled via the Hamiltonian

$$\mathcal{H}_{\text{bdry},\mathcal{I}}(T) = \sigma_x m(x, T) - \mathrm{i}\sigma_y \frac{\partial}{\partial x}, \qquad (25)$$

where $T \in [0, 1]$,

$$m(x, 0) = -m(m > 0), \quad m(x, 1) = -(|x| - L/2)m, \qquad (26)$$

and $m(x, T)$ interpolates continuously between $m(x, 0)$ and $m(x, 1)$ in an inversion-symmetric fashion: $m(x, T) = m(-x, T)$. $\mathcal{H}_{\text{bdry},\mathcal{I}}(0)$ then describes a fully trivial system without boundaries, while $\mathcal{H}_{\text{bdry},\mathcal{I}}(1)$ describes a topological system of length L that is embedded in a trivial environment. For large L, we may take over the result we obtained for a single edge, Eq. (23), and deduce that $\mathcal{H}_{\text{bdry},\mathcal{I}}(1)$ hosts two inversion-related end states as discussed above. The only important feature of the interpolating Hamiltonian $\mathcal{H}_{\text{bdry},\mathcal{I}}(T)$ is that it also respects sublattice symmetry at all values of T, ensuring that the two end states that are present at $T = 1$ but not at $T = 0$ must come in with opposite energies $E_1(T) = -E_2(T)$ as T is tuned from 0 to 1, with $E_1(1) = E_2(1) = 0$ [recall Eq. (24) and the surrounding discussion]. Correspondingly, only one state is filled (occupied by an electron in the system's ground state), namely, the one deriving from the negative-energy manifold of states.[34] We conclude that the system hosts two end states that together have only one electron's worth of electric charge [recall how the total charge Q is defined in Eq. (7) as the number of

a

$\mathcal{I}$

b

FIG. 4. Filling anomaly of the inversion-symmetric SSH model with open boundary conditions. (a) In the topological phase, each end of an inversion-symmetric sample hosts a single bound state worth of half an electronic charge. (b) Exact diagonalization spectrum of the SSH tight-binding model, Eq. (19), with open boundary conditions ($L = 100$). Only the states close to $E = 0$ are shown. The presence of two midgap states in the exact model validates the Dirac approximation [discussion around Eq. (25)]. Half-filling implies that there are L occupied states, highlighted in orange. Importantly, only one of the two midgap bound states is occupied, giving rise to a filling anomaly.

occupied electrons]. The only inversion-symmetric way to distribute this charge is to assign half a charge to either end of the system, essentially fractionalizing the electronic charge.

The trade-off between the global $U(1)$ symmetry, which naively implies charge quantization for a Hamiltonian of the form of Eq. (5), and the crystalline inversion symmetry, constitutes the so-called *filling anomaly*.[35] It is the only universal boundary characteristic of inversion-protected topological insulators in one dimension [see Fig. 4(a)]. The spectrum of Eq. (19) with open boundary conditions, shown in Fig. 4(b), confirms these results.

While we have used sublattice symmetry to show that the edge states of $\mathcal{H}_{\text{SSH}}(k)$ are (filling) anomalous, it should be emphasized that sublattice symmetry is *not* required to stabilize the anomaly: we used sublattice symmetry simply to shortcut an otherwise lengthy calculation of the domain wall charge. It can be broken without destroying either the crystalline bulk topology or the edge state filling anomaly, which are both protected by inversion symmetry and $U(1)$ charge conservation alone. Note also that the filling anomaly, even though it derives from the zero-dimensional edge states, crucially relies on the topologically nontrivial one-dimensional bulk: as we have shown in Sec. II A, there is no well-defined (regularizable) Hamiltonian that preserves $U(1)$ symmetry and has a ground state with fractional charge.

IV. SECOND-ORDER TOPOLOGICAL INSULATORS IN TWO DIMENSIONS

We next leverage what we have learned from the SSH model to understand the simplest possible example of a higher-order topological insulator: the inversion-symmetric case in two spatial dimensions. Again, the Dirac approximation provides the most direct route to a crystalline bulk-boundary correspondence and filling anomaly.

Historically, higher-order phases were discovered much later[37–41] than their first-order counterparts such as the Chern insulator[42] or the time-reversal symmetric topological insulator.[43] Ironically however, at least in 2D, the higher-order case is conceptually simpler. This allows us to bypass the discussion of first-order phases, which have been extensively reviewed elsewhere.[1–4,29]

A. Concept

In Sec. III, we have seen how zero-dimensional states can arise at the ends of a one-dimensional insulator that preserves inversion symmetry. This kind of topological bulk-boundary correspondence is called first order, because the gapless boundaries have a dimension that is one lower than that of the bulk. Similarly, higher-order topological insulators in D spatial dimensions are defined by the presence of gapless states along boundary segments with dimension $D - K$, where $K > 1$ is the order. The simplest example is corner states in two dimensions, which realize a corner (zero-dimensional) filling anomaly similar to what was discussed in Sec. III D, while the bulk (2D) and edges (1D) are gapped. The filling anomaly gives rise to fractional corner charge, similar to how the SSH filling anomaly gave rise to fractional end charge. There also exist 2D first-order topological insulators, most notably the Chern insulator (the integer quantum Hall effect) and the time-reversal symmetric
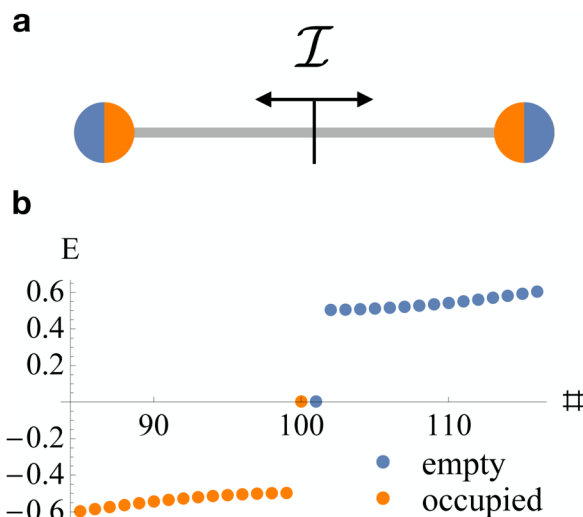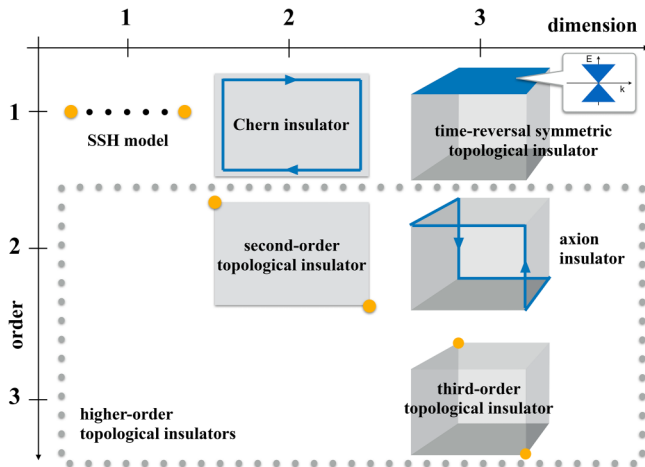
FIG. 5. Archetypes of topological insulators. The SSH model, as well as the second- and third-order topological insulators, host 0D states on their boundaries (here, the inversion-symmetric case is shown in all three dimensions—let us note that in 2D and 3D, there are further crystalline symmetries beyond inversion that stabilize higher-order topology). The boundaries of Chern and axion insulators[36] host 1D conducting states (1D chiral metals). The surface of the 3D time-reversal symmetric topological insulator features a gapless Dirac cone surface state. Importantly, although the gapless boundary states of nth order phases in n dimensions [and $(n − 1)$th order phases in n dimensions] are of the same type, these phases are genuinely distinct: for instance, a second-order topological insulator in 2D is not adiabatically related to a stack of 1D SSH models.

topological insulator (the quantum spin Hall effect). An overview of first- and higher-order topological phases is given in Fig. 5.

First-order phases realize insulators (gapped band structures) in the bulk and anomalous metals [gapless band structures that cannot be obtained from well-defined, regularizable $(D − 1)$-dimensional Hamiltonians] on the boundary. Second-order phases, on the other hand, realize insulators in the bulk and *anomalous insulators* [gapped band structures that cannot be obtained from well-defined, regularizable $(D − 1)$-dimensional Hamiltonians] on the boundary. These anomalous boundaries can then themselves be viewed as first-order topological phases that host their own gapless end states of dimension $D − 2$.

Crucially, we cannot simply view the $(D − 1)$-dimensional boundaries as first-order topological insulators that have been glued to the boundary of an otherwise trivial D-dimensional system:[44] even though they are gapped, they are anomalous and can only be realized in the presence of a topologically nontrivial bulk. These qualitative considerations will be made concrete in the following.

B. Bloch and Dirac Hamiltonian

Unlike in Sec. III A, where we made at least some attempt at motivating the SSH Hamiltonian from the microscopic chemistry of polyacetylene, we will from now on study the properties of toy model Hamiltonians without further physical legitimization. The

reason is that these capture the essential physics without distracting us too much with the non-universal properties of any given material. Moreover, higher-order topological phases were only discovered recently, and as of yet only very few "nice" material realizations are known.[45–47] For our purposes, it will therefore be sufficient to study these novel phases of matter on a conceptual level—it is one of the great luxuries of condensed matter theory that very often initially theoretical concepts eventually do turn out to be realized in nature.

Following up on our general characterization of two-dimensional second-order topological insulators in Sec. IV A, let us therefore ask the question: how do we construct a two-dimensional tight-binding that hosts corner states? Evidently, we will again need a crystalline symmetry to ensure topological protection: with only local symmetries in place, any corner states could be brought together along the sample boundary and annihilated with each other—this process would be adiabatic with respect to the bulk gap. We note that the requirement of crystalline symmetries (symmetries that do not leave all spatial positions x invariant) is particular to higher-order phases. First-order phases *can* be protected by crystalline symmetry operations, but do not *have* to—we have seen this already for the SSH insulator that can be stabilized by either sublattice or inversion symmetry. For simplicity, we will again consider inversion symmetry here, which in 2D maps both coordinates $\boldsymbol{x} = (x, y) \rightarrow (−x, −y)$ to their opposite values.

In particular, inversion symmetry maps the one-dimensional boundary of a two-dimensional, inversion-symmetric sample to itself. A natural starting point to construct a second-order topological insulator is then to engineer a phase that on its boundary hosts a phase akin to the one-dimensional, inversion-symmetric SSH model—domain walls in the SSH parameter m [Eq. (20)] then bind zero-dimensional "corner" states. We will show in the following that such a dimensional hierarchy is achieved by the four-band tight-binding Hamiltonian

$$\mathcal{H}_{\text{SOTI}}(\boldsymbol{k}) = (\Delta + \cos k_x + \cos k_y)\tau_x\sigma_0 + \sin k_x\tau_z\sigma_x + \sin k_y\tau_z\sigma_y + \delta O, \tag{27}$$

where τ_i and σ_i, $i = 0, x, y, z$ are two sets of Pauli matrices, and we abbreviated the Kronecker product by $\tau_i \otimes \sigma_j \equiv \tau_i\sigma_j$. Inversion symmetry is given by $\mathcal{I} = \tau_x\sigma_0$, and the topological parameter regime is delimited by $|\Delta| < 2$. The small constant $\delta \ll 1$ multiplies a collection of symmetry-allowed perturbation terms O that will be essential in gapping out the one-dimensional boundary.[48] We may again define inversion eigenvalues at the four high-symmetry momenta $\bar{\boldsymbol{k}} = (0, 0), (0, \pi), (\pi, 0), (\pi, \pi)$, just as was done in Sec. II C. Declaring $\Delta < −2$ as the trivial reference phase,[49] the topological phase is then characterized by a sign flip of two occupied inversion eigenvalues at $\bar{\boldsymbol{k}} = (0, 0)$ [see Fig. 6(a)], prompting us to expand around that momentum in a Dirac approximation:

$$\mathcal{H}_{\text{2D,Dirac}}(\boldsymbol{k}) = m\tau_x\sigma_0 + k_x\tau_z\sigma_x + k_y\tau_z\sigma_y + \delta O. \tag{28}$$

This Dirac Hamiltonian captures the trivial-to-topological gap-closing transition of $\mathcal{H}_{\text{SOTI}}(\boldsymbol{k})$ when $m = \Delta + 2$ is tuned from negative to positive values.
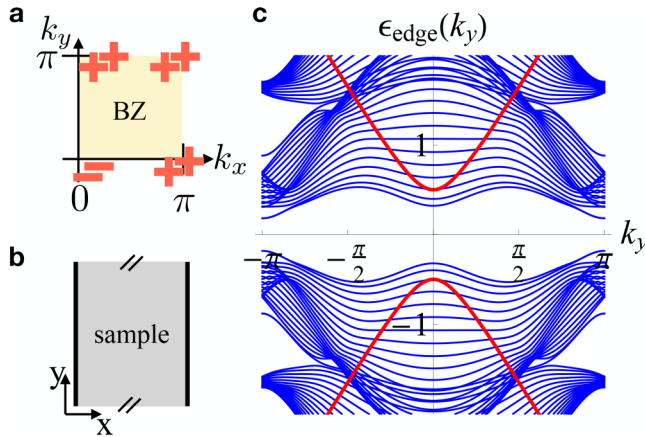
FIG. 6. Bulk and edge band structure of the 2D higher-order topological insulator modeled by Eq. (27). (a) Bulk inversion eigenvalues in the 2D Brillouin zone. There is a double band inversion at $\mathbf{k} = (0, 0)$. (b) Inversion-symmetric slab geometry with open boundary conditions in the x direction and periodic boundary conditions in the y direction. (c) Exact diagonalization spectrum of the tight-binding model, Eq. (27), in the slab geometry with $L_x = 20$ (shown in blue). The dispersion of the edge states derived from the Dirac approximation, Eq. (31), is shown in red.

C. Gapped edge states

As for the SSH model, we will first derive the band structure of $\mathcal{H}_{\mathrm{SOTI}}(\mathbf{k})$ in the presence of a single edge before arguing for the topological signature of a full inversion-symmetric boundary. Without loss of generality, we keep periodic boundary conditions in the y direction so that the edge under consideration has its normal along the (negative) x direction. A mass profile $m(\mathbf{x}) = mx$, $m > 0$ then locates the left edge of the topological phase at $x = 0$. Similar to Eq. (21), any boundary-localized zero-energy states are determined by the Hamiltonian

$$\mathcal{H}_x(k_y) = \tau_x \sigma_0 mx - i\tau_z \sigma_x \frac{\partial}{\partial x} + k_y \tau_z \sigma_y + \delta O. \tag{29}$$

There are two normalizable solutions to the zero-energy condition $\mathcal{H}_x(0)|\Psi(x)\rangle = 0$ at $k_y = \delta = 0$, given by

$$|\Psi_1(x)\rangle = \frac{1}{\mathcal{N}} e^{-\frac{1}{2}mx^2} \begin{pmatrix} i \\ 0 \\ 0 \\ 1 \end{pmatrix}, |\Psi_2(x)\rangle = \frac{1}{\mathcal{N}} e^{-\frac{1}{2}mx^2} \begin{pmatrix} 0 \\ i \\ 1 \\ 0 \end{pmatrix}, \tag{30}$$

where $\mathcal{N} = (4\pi/m)^{\frac{1}{4}}$ is a normalization factor.

We may now deduce the form of the edge Hamiltonian (the Hamiltonian governing the edge degrees of freedom) by degenerate first-order perturbation theory in k_y and δ. Let us for the moment fix $O = \tau_0 \sigma_x$—this choice preserves inversion symmetry. The edge Hamiltonian is given by the matrix elements of $\mathcal{H}_x(k_y)$ in the basis

$\{|\Psi_1(x)\rangle, |\Psi_2(x)\rangle\}$ spanned by the zero-energy edge states,

$$\begin{aligned}[\mathcal{H}_{\mathrm{edge}}(k_y)]_{mn} &= \langle \Psi_m(x)|\mathcal{H}_x(k_y)|\Psi_n(x)\rangle \\ &= [\delta\xi_x + k_y\xi_y]_{mn}, \end{aligned} \tag{31}$$

where we introduced another set of Pauli matrices ξ_i, $i = 0, x, y, z$. We thus find that the edge of $\mathcal{H}_{\mathrm{SOTI}}(\mathbf{k})$ is governed by the SSH Dirac Hamiltonian, Eq. (18), with the identification $m = \delta$. In particular, any finite $|\delta| > 0$ implies that the edge is gapped out, and so there is no protected filling anomaly stemming from the one-dimensional edge states. The edges of a second-order topological insulator are therefore gapped and seemingly trivial. While some choices of the perturbation O do not lead to a gap (for instance, $\tau_0 \sigma_z$), the existence of a single symmetry-allowed gapping term is already enough to trivialize the edge from a topological point of view (also, such a term will be generically present in real materials). The Dirac results are confirmed by a full diagonalization of $\mathcal{H}_{\mathrm{SOTI}}(\mathbf{k})$ in a slab geometry, see Figs. 6(b) and 6(c).

D. Gapless corner states

We have shown that the Hamiltonian $\mathcal{H}_{\mathrm{SOTI}}(\mathbf{k})$ is gapped in bulk and on edges with normal in the x direction. Similarly, one can show that the states along edges with normal in the y direction are gapped and described by a Hamiltonian just like Eq. (31). We might thus conclude that the entire system is gapped and featureless. There is, however, one catch: the edge termination we considered is not inversion symmetric. Recall how we first derived the SSH end state for a single edge in Sec. III B, before deducing that an inversion-symmetric pair of edges hosts a pair of edge states of which only one is filled. To determine the anomalous boundary physics of second-order topological insulators, we likewise have to consider an inversion-symmetric edge termination, for instance, that of a rectangular sample. A rectangle has four edges (call them left and right, top and bottom) that all come with their respective version of Eq. (29)—in Sec. IV C, we have explicitly treated the left edge only. It does, however, not suffice to repeat the calculation for the remaining three edges separately: due to the gauge freedom in defining the edge states (we may multiply either state with a phase factor, or take arbitrary superpositions, to arrive at an equally valid set of edge states), there would be no meaningful way of comparing the resultant edge Hamiltonians. Instead, we must find the bound states of an arbitrary edge, labeled by an angle ϕ—these then continuously evolve with ϕ as we interpolate from one edge to another, and thereby provide a consistent choice of gauge.

For an arbitrary edge that has its normal at an angle ϕ to the x axis, it makes sense to decompose the vector of crystal momenta as

$$\mathbf{k} = k_\perp \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} + k_\parallel \begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix}, \tag{32}$$

where $k_\perp$ is the momentum along the edge normal and $k_\parallel$ is the momentum along the edge itself (note that we are orienting $k_\parallel$ at a

positive right angle from $k_\perp$). The Dirac Hamiltonian then reads

$$\mathcal{H}_{\text{2D,Dirac}}^\phi(k_\perp, k_\parallel) = m\tau_x\sigma_0 + k_\perp\tau_z\sigma_1(\phi) + k_\parallel\tau_z\sigma_2(\phi) + \delta O, \quad (33)$$

where $\sigma_1(\phi) = \sigma_x\cos\phi + \sigma_y\sin\phi$ and $\sigma_2(\phi) = \sigma_y\cos\phi - \sigma_x\sin\phi$. After decomposing the position coordinate $\boldsymbol{x}$ into $x_\perp$ and $x_\parallel$ similar to Eq. (32), we again implement the edge via a domain wall mass profile $m(x_\perp) = mx_\perp$, $m > 0$ so that the topological phase lies at positive values of $x_\perp$ (the edge normal points into the topological region). This setup retains $k_\parallel$ as a good momentum quantum number, while we replace $k_\perp$ by $-i\partial/\partial x_\perp$,

$$\mathcal{H}_{x_\perp}^\phi(k_\parallel) = \tau_x\sigma_0 mx_\perp - i\tau_z\sigma_1(\phi)\frac{\partial}{\partial x_\perp} + k_\parallel\tau_z\sigma_2(\phi) + \delta O. \quad (34)$$

Edge states at $k_\parallel = \delta = 0$ then satisfy

$$[\tau_x\sigma_0 mx_\perp - i\tau_z\sigma_1(\phi)\frac{\partial}{\partial x_\perp}]|\Psi_n^\phi(x_\perp)\rangle = 0. \quad (35)$$

Noting that $[\sigma_1(\phi)]^2 = \div$, the eigenvalues of $\sigma_1(\phi)$ are constant and given by -1 and $+1$. The equation can then be solved to give

$$
\begin{aligned}
|\Psi_1^\phi(x_\perp)\rangle &= \frac{1}{\mathcal{N}} e^{-\frac{1}{2}mx_\perp^2}\begin{pmatrix} e^{-i\phi} \\ 0 \\ 0 \\ i \end{pmatrix}, \\
|\Psi_2^\phi(x_\perp)\rangle &= \frac{1}{\mathcal{N}} e^{-\frac{1}{2}mx_\perp^2}\begin{pmatrix} 0 \\ e^{+i\phi} \\ i \\ 0 \end{pmatrix},
\end{aligned}
\quad (36)
$$

where $\mathcal{N} = (4\pi/m)^{\frac{1}{4}}$ is a normalization factor. The dispersion of these states with edge momentum $k_\parallel$, and their gapping, can again be derived in degenerate first-order perturbation theory by evaluating the matrix elements

$$
\begin{aligned}
[\mathcal{H}_{\text{edge}}^\phi(k_\parallel)]_{mn} &= \langle\Psi_m^\phi(x_\perp)|\mathcal{H}_{x_\perp}^\phi(k_\parallel)|\Psi_n^\phi(x_\perp)\rangle \\
&= [k_\parallel(\xi_x\sin\phi + \xi_y\cos\phi) + \delta\zeta(\phi)]_{mn},
\end{aligned}
\quad (37)
$$

where

$$[\zeta(\phi)]_{mn} = \langle\Psi_m^\phi(x_\perp)|O|\Psi_n^\phi(x_\perp)\rangle. \quad (38)$$

We will in fact not use any explicit form of the perturbation terms O in the following—it will be sufficient to note that these terms will generically gap the two edge states as long as they project into a nonzero $\zeta(\phi)$ (we have already seen this for the particular choice of $O = \tau_0\sigma_x$ and $\phi = 0$). Importantly, however, we now show that inversion symmetry implies stable zeroes of $\zeta(\phi)$. First, note that it acts on the edge states as follows:

$$
\begin{aligned}
\iota_{mn}^\phi &= \langle|\Psi_m^{\phi+\pi}(x_\perp)\mathcal{I}|\Psi_n^\phi(x_\perp)\rangle \\
&= [-i(\xi_x\cos\phi - \xi_y\sin\phi)]_{mn}.
\end{aligned}
\quad (39)
$$

The constraint of inversion symmetry then implies

$$\iota^\phi\mathcal{H}_{\text{edge}}^\phi(k_\parallel)\iota^{\phi\dagger} = \mathcal{H}_{\text{edge}}^{\phi+\pi}(k_\parallel). \quad (40)$$

(Recall that the momentum $k_\parallel$ is defined at a right angle to the edge normal, it rotates with ϕ and so is not separately flipped by inversion.) Now, any perturbation in Eq. (37) that anticommutes with the kinetic term (other terms do not lead to a gap and can be absorbed into a redefinition of $k_\parallel$) can be decomposed as

$$\zeta(\phi) = a(\phi)(\xi_x\cos\phi - \xi_y\sin\phi) + b(\phi)\xi_z, \quad (41)$$

so that the eigenvalues of $\mathcal{H}_{\text{edge}}^\phi(k_\parallel)$ are given by $\epsilon_\pm^\phi(k_\parallel) = \pm\sqrt{k_\parallel^2 + a(\phi)^2 + b(\phi)^2}$. Equation (40) yields the constraints

$$a(\phi + \pi) = a(\phi), \quad b(\phi + \pi) = -b(\phi), \quad (42)$$

which together imply

$$\zeta(\phi + \pi) = -\zeta(\phi). \quad (43)$$

This relation necessitates the presence of domain walls in the Dirac mass, $\zeta(\phi) = 0$, at some angles $\bar\phi$ and $\bar\phi + \pi$ where $\zeta(\phi)$ switches sign [see Fig. 7(a)]. The domain walls are movable yet irremovable: perturbations may change the precise location $\bar\phi$, but domain walls can only be annihilated in pairs. Since the partner of the domain wall at $\bar\phi$ is fixed to the location $\bar\phi + \pi$ by inversion symmetry, the two can never come together and cancel.

We conclude that, in a rectangular geometry, the edge states on the left and right edges, as well as on the top and bottom edges, have to come with the opposite signs of their Dirac mass. Viewing the full rectangular boundary as a 1D system governed by a SSH type Hamiltonian [Eq. (37)], it then hosts Dirac mass domain walls at two oppositely related corners [see Figs. 7(b) and 7(c)].[50] The situation is therefore just that of Eq. (25) at $t = 1$, where we
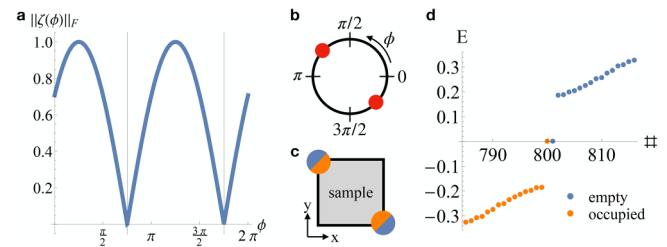

FIG. 7. Corner states in 2D. (a) The projected edge mass, Eq. (35), has (at least) two stable domain walls at inversion-related momenta $\bar\phi$ and $\bar\phi + \pi$. As an example, here the Frobenius norm of $\zeta(\phi)$ is shown for $O = \tau_0\sigma_x + \tau_0\sigma_y$. [(b) and (c)] The domain wall bound states at $\bar\phi$ and $\bar\phi + \pi$ become corner states in a rectangular geometry. (d) Exact diagonalization spectrum of Eq. (27), with open boundary conditions in two directions ($L_x = L_y = 20$). Only the states close to $E = 0$ are shown. The half-filled corner midgap states give rise to an inversion symmetry-protected filling anomaly [see discussion around Eq. (25)].

considered a SSH model with two inversion-related domain walls and showed that these host end states with a filling anomaly. Equivalently, the corners of $\mathcal{H}_{\mathrm{SOTI}}(\boldsymbol{k})$ bind gapless states that lead to a filling anomaly of second-order: while the infinite slab geometry of Fig. 6(b) is gapped, a filling anomaly arises when introducing open boundary conditions in two directions. Just like for the SSH model, the presence of the filling anomaly is guaranteed by the topologically nontrivial bulk gap and inversion symmetry alone, and neither relies on the particular structure of $\mathcal{H}_{\mathrm{SOTI}}(\boldsymbol{k})$, nor on the rectangular sample geometry we chose to derive it. The Dirac result is confirmed by a full diagonalization of the tight-binding Hamiltonian, shown in Fig. 7(d).

The corner states discussed here have been predicted to naturally occur in 2D monolayers of antimony and arsenic at finite buckling.[51]

ACKNOWLEDGMENTS

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] E. Fradkin, *Field Theories of Condensed Matter Physics*, 2nd ed. (Cambridge University Press, 2013).

[2] J. K. Asbóth, L. Oroszlány, and A. Pályi, *A Short Course on Topological Insulators* (Springer, 2016), Vol. 919.

[3] B. A. Bernevig and T. L. Hughes, *Topological Insulators and Topological Superconductors* (Princeton University Press, 2013).

[4] A. Bernevig and T. Neupert, "Topological superconductors and category theory: Lecture Notes of the Les Houches Summer School," in *Topological Aspects of Condensed Matter Physics* (Oxford University Press, 2017), pp. 63–121.

[5] T. Neupert and F. Schindler, "Topological crystalline insulators," in *Topological Matter: Lectures from the Topological Matter School 2017*, edited by D. Bercioux, J. Cayssol, M. G. Vergniory, and M. Reyes Calvo (Springer International Publishing, Cham, 2018), pp. 31–61.

[6] H. C. Po, "Symmetry indicators of band topology," J. Phys. Condens. Matter 32, 263001 (2020).

[7] P. W. Anderson, "More is different," Science 177, 393–396 (1972).

[8] F. Wilczek, "Quantum mechanics of fractional-spin particles," Phys. Rev. Lett. 49, 957 (1982).

[9] The formation of the crystalline lattice, on the other hand, can be understood in classical physics via potential energy minimization.

[10] In contrast, non-interacting bosons created by $b_\alpha^\dagger$ condense into a Bose-Einstein condensate, where they all populate the same single-particle wavefunction: $|\Omega\rangle \propto (b_{\tilde\alpha}^\dagger)^N |0\rangle$, where $\tilde\alpha$ labels the lowest energy state of the single-particle Hamiltonian.

[11] By a small electric field, we mean that its energy per particle is much smaller than the gap.

[12] B. Bradlyn, L. Elcoro, J. Cano, M. G. Vergniory, Z. Wang, C. Felser, M. I. Aroyo, and B. A. Bernevig, "Topological quantum chemistry," Nature 547, 298–305 (2017).

[13] T. Senthil, "Symmetry-protected topological phases of quantum matter," Annu. Rev. Condens. Matter Phys. 6, 299–324 (2015).

[14] L. Fu, "Topological crystalline insulators," Phys. Rev. Lett. 106, 106802 (2011).

[15] 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, 405–408 (1982).

[16] Note that in zero dimensions, there is no notion of wavefunction localization or atomic limit, we therefore have to fix the trivial phase by hand. This makes the notion of topological phase somewhat artificial but allows us to demonstrate it in its simplest possible variant.

[17] The vacuum corresponds to an atomic insulator where the atoms are infinitely far apart.

[18] One might ask why the usual non-topological phases, such as ferromagnets, do not host boundary phase transitions by the same argument. The reason is that these are characterized by local order parameters (such as magnetization) that can smoothly go to zero at the sample boundary, with an energy cost that scales sub-extensively with the system size. Topological insulators, on the other hand, are characterized by non-local and quantized topological invariants that cannot change continuously.

[19] A. Zee, *Quantum Field Theory in a Nutshell* (Princeton University Press, 2010).

[20] Naively, the Mermin–Wagner theorem prevents the formation of crystals, which break the continuous translational symmetry of space down to the discrete subgroup of lattice translations, in dimensions smaller than three. Nevertheless, one-dimensional crystals exist in nature, either because they are not perfectly one-dimensional (their atoms do not all lie along a single line) such as is the case for polymer chains, or because they are part of three-dimensional structures, such as is the case for adatom chains.

[21] Note that our notion of (trivial) atomic limit, which is characterized by vanishing hopping terms, depends on the choice of unit cell.

[22] In the literature, such additional inequivalent atomic limits are called obstructed atomic limits (OALs), here we gloss over the distinction between topological phases, OALs, and the recently introduced fragile phases[52]—these can all be viewed as topological by the same criterion, namely, that we cannot adiabatically deform them to a trivial atomic limit that has vanishing hopping terms.

[23] In the present example, half-filling implies that the bandgap of $\mathcal{H}(k)$ induces an energy gap above the many-body ground state of H. It is important to keep in mind that the filling (the number of electrons in the sample), or alternatively the chemical potential, is just as important as the presence of a bandgap in order for a system to be insulating.

[24] W. P. Su, J. R. Schrieffer, and A. J. Heeger, "Solitons in polyacetylene," Phys. Rev. Lett. 42, 1698–1701 (1979).

[25] In fact, an exact analytical solution is possible for the specific model considered here, but it is not particularly illuminating as it only works in this fine-tuned case. We are not interested in the minute details of solutions to one particular model, but rather want to determine the universal topological features of the corresponding phase of matter. For this purpose, the approximation we will make is sufficient and has the advantage that it can be generalized to arbitrary tight-binding models.

[26] The reader might be more familiar with the real-space Dirac equation in $3 + 1$ spacetime dimensions, which is constructed using $4 \times 4 \gamma$ matrices instead of the 2×2 Pauli matrices. The formal requirement for a Dirac Hamiltonian is that all matrices that appear anti-commute (form a Clifford algebra) and square to $+$. This requirement can be satisfied using only 2×2 matrices in $1 + 1$ dimensions but not in $3 + 1$ dimensions (where there are two more momenta that need to be accommodated).

[27] The expansion around $k = 0$ is often called long-wavelength or low-energy expansion. Here, it is however purely coincidental that the topological distinction between trivial and topological originates from $k = 0$. Had we chosen the atomic limit built from antisymmetric, odd-parity superpositions as our trivial reference phase, the relevant momentum would have been $k = \pi$. This underlines the fact that topological properties do not care about energetics or length scales, they are global properties of the full band structure, and not just of its long-wavelength part.

[28] In this Tutorial, we only treat so-called strong topological phases that do not require (discrete) translational symmetry for their protection. Correspondingly,

we do not care about the particular lattice structure of the crystal and may discard it with impunity when going over to the continuum Dirac description.

[29]M. Z. Hasan and C. L. Kane, "Colloquium: Topological insulators," Rev. Mod. Phys. **82**, 3045–3067 (2010).

[30]A. Alexandradinata, X. Dai, and B. A. Bernevig, "Wilson-loop characterization of inversion-symmetric topological insulators," Phys. Rev. B **89**, 155114 (2014).

[31]A. M. Essin and V. Gurarie, "Bulk-boundary correspondence of topological insulators from their respective Green's functions," Phys. Rev. B **84**, 125132 (2011).

[32]In fact, sublattice symmetry stabilizes an arbitrary number of end states, and $\mathcal{C}$-symmetric Hamiltonians that host more than one end-localized zeromode exist—for instance, we may just double all degrees of freedom and hopping elements per unit cell to arrive at a model with two protected end states. We do not consider such models in our Tutorial because we want to focus on the more physically relevant protection by inversion symmetry, with which only a single zeromode is stable.

[33]Recall that sublattice symmetry maps an energy E to $-E$ so that the point $E = 0$ is special.

[34]To fix the filling, note that the system at $T = 0$ is gapped and we are working at zero chemical potential, $\mu = 0$. This implies that the many-body ground state is a Slater determinant of the form of Eq. (3), containing only the negative-energy eigenstates. As T is tuned, it continuously evolves into the ground state at $T = 1$, which has two end states at zero energy (as long as the separation L between the two ends is sufficiently large). The filling (particle number in the ground state) is conserved by this interpolation, therefore the zeromode that started out at negative-energy must remain filled at $T = 1$, while the zeromode that started out at positive energy must remain empty.

[35]W. A. Benalcazar, T. Li, and T. L. Hughes, "Quantization of fractional corner charge in C_n-symmetric higher-order topological crystalline insulators," Phys. Rev. B **99**, 245151 (2019).

[36]C. Fang and L. Fu, "New classes of three-dimensional topological crystalline insulators: Nonsymmorphic and magnetic," Phys. Rev. B **91**, 161105 (R) (2015).

[37]W. A. Benalcazar, B. A. Bernevig, and T. L. Hughes, "Quantized electric multipole insulators," Science **357**, 61–66 (2017).

[38]Z. Song, Z. Fang, and C. Fang, "$(d-2)$-dimensional edge states of rotation symmetry protected topological states," Phys. Rev. Lett. **119**, 246402 (2017).

[39]J. Langbehn, Y. Peng, L. Trifunovic, F. von Oppen, and P. W. Brouwer, "Reflection-symmetric second-order topological insulators and superconductors," Phys. Rev. Lett. **119**, 246401 (2017).

[40]W. A. Benalcazar, B. A. Bernevig, and T. L. Hughes, "Electric multipole moments, topological multipole moment pumping, and chiral hinge states in crystalline insulators," Phys. Rev. B **96**, 245115 (2017).

[41]F. Schindler, A. M. Cook, M. G. Vergniory, Z. Wang, S. S. P. Parkin, B. A. Bernevig, and T. Neupert, "Higher-order topological insulators," Sci. Adv. **4**, eaat0346 (2018).

[42]F. D. M. Haldane, "Model for a quantum Hall effect without landau levels: Condensed-matter realization of the 'parity anomaly'," Phys. Rev. Lett. **61**, 2015–2018 (1988).

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

[44]L. Trifunovic and P. W. Brouwer, "Higher-order topological band structures," Phys. Status Solidi B **2020**, 2000090.

[45]F. Tang, H. C. Po, A. Vishwanath, and X. Wan, "Comprehensive search for topological materials using symmetry indicators," Nature **566**, 486–489 (2019).

[46]M. G. Vergniory, L. Elcoro, C. Felser, N. Regnault, B. A. Bernevig, and Z. Wang, "A complete catalogue of high-quality topological materials," Nature **566**, 480–485 (2019).

[47]T. Zhang, Y. Jiang, Z. Song, H. Huang, Y. He, Z. Fang, H. Weng, and C. Fang, "Catalogue of topological electronic materials," Nature **566**, 475–479 (2019).

[48]In principle, we can add any inversion-symmetric term, k-dependent or independent, to the Hamiltonian without changing any of the topological properties (assuming the term is small enough so as not to close the bulk gap). In doing so, we have to tread a fine line: adding no perturbations at all is dangerous because it may give a fine-tuned model that exhibits accidental properties not intrinsic to the topological phase itself, while adding too many perturbing terms makes an analytic treatment burdensome and thereby hinders an intuitive understanding of the physics. Hence, equipped with the benefit of hindsight, we only consider absolutely essential perturbations in this Tutorial. In principle, all symmetry-allowed perturbations will be present in realistic materials; however, they cannot affect topological properties as long as they are so small that they do not close the bulk gap.

[49]For $\Delta < -2$, all eight occupied inversion eigenvalues (there are two occupied bands and four inversion-symmetric momenta) are equal to $+1$, this configuration is compatible with an atomic limit that is constructed from two occupied s orbitals located at the center of the unit cell.

[50]The presence of two inversion-related zero-dimensional bound states is a general property of the nontrivial bulk and not specific to a rectangular geometry. For instance, we could also have considered a circular geometry with a sufficiently large radius of curvature: as long as a disk-shaped sample is large enough so that we can approximate the local electronic band structure at each point of its boundary by a Dirac equation of the type in Eq. (33), the two domain walls in $\zeta(\phi)$ give rise to two filling-anomalous bound states on the circle. The failure of our Dirac approximation at small radii has a very physical interpretation: for finite-sized circles, the angular momentum quantum number takes on quantized values so that in the exact solution of the Dirac equation on a circle there is a finite energy gap in the edge state dispersion, Eq. (37), that scales inversely proportional to the radius.

[51]F. Schindler, M. Brzezińska, W. A. Benalcazar, M. Iraola, A. Bouhon, S. S. Tsirkin, M. G. Vergniory, and T. Neupert, "Fractional corner charges in spin-orbit coupled crystals," Phys. Rev. Res. **1**, 033074 (2019).

[52]H. C. Po, H. Watanabe, and A. Vishwanath, "Fragile topology and Wannier obstructions," Phys. Rev. Lett. **121**, 126402 (2018).