

Single-point spin Chern number in a supercell framework

Roberta Favata  and Antimo Marrazzo* 

Dipartimento di Fisica, Università di Trieste, Strada Costiera 11, I-34151 Trieste, Italy

* Author to whom any correspondence should be addressed.

E-mail: antimo.marrazzo@units.it

Keywords: Chern, supercell, single-point, quantum spin Hall, topological invariant

Abstract

We present an approach for the calculation of the \mathbb{Z}_2 topological invariant in non-crystalline two-dimensional quantum spin Hall insulators. While topological invariants were originally mathematically introduced for crystalline periodic systems, and crucially hinge on tracking the evolution of occupied states through the Brillouin zone, the introduction of disorder or dynamical effects can break the translational symmetry and imply the use of larger simulation cells, where the \mathbf{k} -point sampling is typically reduced to the single Γ -point. Here, we introduce a single-point formula for the spin Chern number that enables to adopt the supercell framework, where a single Hamiltonian diagonalisation is performed. Inspired by the work of Prodan (2009 *Phys. Rev. B* **80** 125327), our single-point approach allows to calculate the spin Chern number even when the spin operator \hat{s}_z does not commute with the Hamiltonian, as in the presence of Rashba spin-orbit coupling. We validate our method on the Kane–Mele model, both pristine and in the presence of Anderson disorder. Finally, we investigate the disorder-driven transition from the trivial phase to the topological state known as topological Anderson insulator. Beyond disordered systems, our approach is particularly useful to investigate the role of defects, to study topological alloys and in the context of *ab-initio* molecular dynamics simulations at finite temperature.

1. Introduction

Two-dimensional (2D) topological insulators (TI) are materials with an insulating bulk and robust edge states protected by the non-trivial topology of the bulk electronic structure [1, 2]. These systems are discussed through topological invariants, integer quantities which characterise the ground-state electronic wavefunction in the bulk. As long as the topological invariant is non-trivial and, possibly, the symmetries needed to define that topology are preserved, the material is said to be in a topological phase. These invariants are geometrical properties of the electronic structure, as they are defined in terms of quantities such as the Berry phase or the Berry curvature, which involve derivatives of the occupied states in reciprocal space with respect to the quasi-momentum \mathbf{k} [2]. Standard geometrical formulas are usually discretised on a regular mesh of \mathbf{k} -points for numerical implementation. However, most electronic structure calculations for non-crystalline systems are normally performed by diagonalising the Hamiltonian at a single \mathbf{k} -point in a large supercell. Usually the Γ point at the centre of the Brillouin zone (BZ) is considered, although potentially more efficient choices based on the Baldereschi point [3] can be employed. The derivation of single-point formulas for geometrical and topological properties is not at all a trivial task, although successful single-point formalisms have been developed for the Berry phase [4], the orbital magnetisation and the Chern number [5].

In this work, we target the calculation of the topological invariant for non-crystalline 2D insulators with time-reversal (TR) symmetry. As we deal with electronic systems, the TR operator enjoys that $\mathcal{T}^2 = -1$ and in two dimensions the invariant ν is a \mathbb{Z}_2 number: if $\nu = 0$ the topology is trivial, otherwise if $\nu = 1$ we have a quantum spin Hall insulator, where topologically-protected gapless helical edge states cross the bulk gap [2]. Notably, the systems classified by the \mathbb{Z}_2 invariant encompass the large set of non-magnetic 2D materials [6].

Over the years, several methods have been developed to calculate the \mathbb{Z}_2 invariant in crystalline systems with periodic boundary conditions. In the following, we briefly outline some of the most popular and practical methods in the context of electronic structure simulations. If inversion symmetry is present, there is a particularly simple method introduced by Fu and Kane [7], which requires the knowledge of the parity of the occupied states at the four TR-invariant points in the BZ. In the more general case, the \mathbb{Z}_2 invariant can be obtained by tracking the evolution of hermaphrodite [8] (a.k.a. hybrid) Wannier charge centres [9–11], or equivalently the eigenvalues of the Wilson loop [12–14], over half BZ. More recently, generalisations of the Fu–Kane approach based on elementary band representation [15, 16] have been proposed to infer topological phases from the symmetry labels of the occupied bands at high-symmetry points in the BZ [17, 18]. This approach has been systematically formalised through the theories of symmetry-based indicators [19–22] and topological quantum chemistry [23–25], and these have been applied to non-magnetic materials databases to identify non-trivial materials of various classes [26–28]. The \mathbb{Z}_2 invariant can be also computed as an individual Chern number [2] on half of the Hilbert space [10, 11], where the split is performed by two projectors which are smooth and related by TR symmetry. Although several formulas to compute the \mathbb{Z}_2 invariant have been introduced, all the ones we mentioned, and most other existing approaches, require the knowledge of the occupied states at multiple \mathbf{k} -points and become ill-defined for non-crystalline systems; hence in the supercell framework they are of no avail.

Nonetheless, a number of methods have been proposed to deal with non-periodic systems. Some of these [29–31] calculate the \mathbb{Z}_2 invariant by means of a Pfaffian with twisted boundary conditions, as firstly advocated by Kane and Mele in their original discussion of the \mathbb{Z}_2 invariant in presence of disorder and electron–electron interactions [32]. A different method is based on constructing the \mathbb{Z}_2 invariant from the scattering matrix of the system at the Fermi level [33, 34]. Further, there exists a formulation relying on the non-commutative index theorem [35, 36], where the \mathbb{Z}_2 index for disordered topological insulators is computed from the discrete spectrum of a certain compact operator, which is defined as the difference of a proper pair of projection operators [37–39]. An alternative non-commutative approach was proposed by Loring and Hastings [40, 41] and relates the \mathbb{Z}_2 index to the topological obstruction to approximating almost commuting matrices by exactly commuting matrices; its robustness with respect to the introduction of disorder has been investigated in [42]. Very recently, a novel method based on the concept of spillage [43] has been proposed [44], where the identification of topological phases in a non-crystalline system is obtained by calculating the spillage with respect to a crystalline reference structure, whose topological characterization can be performed with standard methods. The most practical approach from the point of electronic structure simulations has been arguably put forward by Huang and Liu [45, 46], who addressed the problem of calculating the \mathbb{Z}_2 invariant for non-periodic system in the context of quantum spin Hall quasicrystals, and introduced the spin Bott index, which measures the commutativity of the projected position operators. The connection between the Bott indices and Chern or \mathbb{Z}_2 invariants has been investigated theoretically [40–42, 47], while numerical simulations [46, 48, 49] provided evidence that Bott indices can be used to study non-periodic topological systems. Still, it is conceptually rather unsatisfactory that the calculation of topological invariants in a supercell framework requires introducing radically different formalisms, which call for rather non-trivial equivalence proofs and extensive testing. As a matter of fact, the use of the primitive cell and \mathbf{k} -points is an arbitrary—although indeed very convenient—choice; there is no conceptual reason preventing bona fide \mathbb{Z}_2 invariants to be calculated directly in the supercell by deriving a suitable single-point limit. In addition, it is important to assess the convergence with respect to the system size, as different approaches might deliver the same correct answer at very different computational costs. For instance, recent works [46, 47] claimed that the difference between the Chern number and Bott index is within a correction of the order $O(1/L)$, where L is the linear size of the system. Such slow convergence can hinder the study of the system close to a topological phase transition; in fact Huang and Liu empirically added a singular value decomposition to their algorithm to improve an otherwise slow convergence [46].

Here, we take a different approach, that essentially combines the work of Ceresoli and Resta on the single-point Chern number [5] and the insights from Prodan on a generalised spin Chern number [50]. Notably, our single-point invariant is directly derived by its parent formula for crystalline systems, it shows exponential convergence with the supercell size, both in the pristine and disordered case, it is easy to implement in electronic structure codes, and it works well also in presence of strong Rashba spin–orbit coupling (SOC).

2. Methods

In absence of spin-mixing spin–orbit interactions, the spin operator \hat{s}_z commutes with the Hamiltonian and it is possible to discuss the \mathbb{Z}_2 invariant in terms of the spin Chern number. In this case, the occupied states diagonalise \hat{s}_z and can be divided in two subsets, either purely spin-up or spin-down, and the regular Chern

number can be calculated for each spin. As soon as the Hamiltonian does not commute any more with \hat{s}_z , for instance because Rashba SOC is present, such simple-minded spin Chern number cannot be defined any more. Notably, Prodan has shown [50] that it is possible to generalise this definition by projecting the spin operator on the occupied states:

$$P_z = P(\mathbf{k})\hat{s}_zP(\mathbf{k}), \quad (1)$$

where P is the ground-state projector

$$P(\mathbf{k}) = \sum_n |u_{n\mathbf{k}}\rangle\langle u_{n\mathbf{k}}|, \quad (2)$$

$u_{n\mathbf{k}}$ are the periodic part of the Bloch eigenstates and n labels the occupied state at each \mathbf{k} -point in the BZ. Then, we diagonalise P_z :

$$P_z|u_\lambda\rangle = s_\lambda|u_\lambda\rangle. \quad (3)$$

If only diagonal SOC terms are present, the eigenvalue spectrum of P_z consists of two values only $s_\lambda = \pm\frac{1}{2}$ and one can select a single spin component by choosing the eigenstates which correspond to one of the two eigenvalues s_λ . The crucial observation made by Prodan [50] is that, even if Rashba SOC is present, the spectrum of P_z displays two separate bands of eigenvalues symmetric around the origin and one can still introduce a well-defined spin Chern number by selecting the eigenvectors with positive (or negative) eigenvalues. Finally, the spin Chern number can be computed as:

$$C_s = \frac{C_+ - C_-}{2} \bmod 2 \quad (4)$$

where C_\pm are calculated on the u_λ eigenstates with positive and negative eigenvalues respectively; in general it is sufficient to compute either C_+ or C_- only and consider its parity. These results are of paramount practical relevance, as it is typically much simpler to deal with a formulation based on generalised Chern numbers, which can be written as full BZ integrals and do not require taking into account TR symmetry or complex gauge fixing, as required instead by more general \mathbb{Z}_2 formulations [32, 51].

In principle, if the Rashba interaction is strong enough then the gap of the P_z spectrum might close, preventing the spin Chern number to be defined. Remarkably, as we will discuss in full detail in the section 3, this does not seem to occur in practice. As long as the system is insulating, equation (4) is well defined even if the Rashba SOC is several times larger than the diagonal SOC. Hence, we adopt the approach of Prodan [50] and target the derivation of a single-point formula. In order to obtain the correct single-point limit, we follow the approach of Ceresoli and Resta [5] for the derivation of the single-point Chern number in TR-broken systems (the latter admit a \mathbb{Z} topological invariant). Let us start with the formula for the generalised spin Chern number in 2D periodic systems:

$$C_\sigma = \frac{1}{2\pi} \int_{BZ} \text{Tr} \Omega_{xy}^\sigma(\mathbf{k}) d\mathbf{k} \quad (5)$$

$$= \frac{1}{2\pi} \sum_{\lambda \in \{\lambda_\sigma\}} \int_{BZ} -2\text{Im} \langle \partial_{k_x} u_\lambda(k_x, k_y) | \partial_{k_y} u_\lambda(k_x, k_y) \rangle dk_x dk_y, \quad (6)$$

where u_λ are the eigenvectors of P_z with eigenvalues s_λ (see equation (3)) and the set $\{\lambda_\sigma\}$ includes only the ones having either positive ($\sigma = +$) or negative ($\sigma = -$) eigenvalues. Now we consider the parallelogram Brillouin zone and change coordinate system to have a rectangular integration domain:

$$C_\sigma = -\frac{1}{\pi} \text{Im} \sum_{\lambda \in \{\lambda_\sigma\}} \int_0^{\mathbf{b}_1} dk_1 \int_0^{\mathbf{b}_2} dk_2 \langle \partial_{k_1} u_\lambda(k_1, k_2) | \partial_{k_2} u_\lambda(k_1, k_2) \rangle \quad (7)$$

$$\simeq -\frac{|\mathbf{b}_1||\mathbf{b}_2|}{\pi} \text{Im} \sum_{\lambda \in \{\lambda_\sigma\}} \langle \partial_{k_1} u_\lambda(k_1, k_2) | \partial_{k_2} u_\lambda(k_1, k_2) \rangle |_{\mathbf{k}=\Gamma}, \quad (8)$$

where $\mathbf{b}_{1,2}$ are the two reciprocal lattice vectors and the last step is performed in the limit of a very large supercell. In the same limit, we can calculate derivatives through finite differences:

$$\partial_{k_j} |u_\lambda(k_1, k_2)\rangle |_{\mathbf{k}=\Gamma} = \lim_{\eta \rightarrow 0} \frac{|u_\lambda(\eta \mathbf{b}_j)\rangle - |u_\lambda(\Gamma)\rangle}{\eta |\mathbf{b}_j|}, \quad (9)$$

where, for a large supercell, we can drop the limit and just consider the difference $|u_\lambda(\mathbf{b}_j)\rangle - |u_\lambda(\Gamma)\rangle$. Equation (9) requires a differentiable function, which is not guaranteed in numerical diagonalisations. Hence, we fix the gauge by adopting a discretised version of the covariant derivative [52, 53] as successfully performed for the Chern number by Ceresoli and Resta [5]. One replaces the states with their ‘duals’:

$$|\tilde{u}_\lambda(\mathbf{b}_j)\rangle = \sum_\mu S_{\mu\lambda}^{-1}(\mathbf{b}_j) |u_\mu(\mathbf{b}_j)\rangle \quad (10)$$

where we define the overlap matrix $S_{\lambda\mu}(\mathbf{b}_j) = \langle u_\lambda(\Gamma) | u_\mu(\mathbf{b}_j) \rangle$ and the dual states satisfy $\langle u_\lambda(\Gamma) | \tilde{u}_\mu(\mathbf{b}_j) \rangle = \delta_{\lambda\mu}$. Next, we construct the states $u_\lambda(\mathbf{b}_j)$ by imposing the periodic gauge, which allows us to perform a single diagonalisation at Γ :

$$|u_\lambda(\mathbf{b}_j)\rangle = e^{-i\mathbf{b}_j \cdot \boldsymbol{\Gamma}} |u_\lambda(\Gamma)\rangle. \quad (11)$$

The states in equation (11) are P_z eigenstates at \mathbf{b}_j , but they might correspond to different eigenvalues with respect to the ones at Γ ; the ordering is anyway fixed by the covariant derivative. We note in passing, that while a non-trivial Chern number would prevent the adoption of a periodic gauge for the wavefunction, here the periodic gauge is only temporarily imposed to build each $|u_\lambda(\mathbf{b}_j)\rangle$ from the knowledge of the $|u_\lambda(\Gamma)\rangle$, but it is effectively replaced by the parallel transport gauge enforced by the covariant derivative. The final single-point formula for the spin Chern number is

$$C_\sigma^{(\text{asym})} = -\frac{|\mathbf{b}_1||\mathbf{b}_2|}{\pi} \text{Im} \sum_{\lambda \in \{\lambda_\sigma\}} \langle \tilde{u}_\lambda(\mathbf{b}_1) | \tilde{u}_\lambda(\mathbf{b}_2) \rangle. \quad (12)$$

In equation (12), we emphasise with the superscript ‘asym’ the implicit choice made in equation (9), which corresponds to the right-hand derivative. In fact, an alternative choice is the symmetric derivative

$$\partial_{k_j} |u_\lambda(k_1, k_2)\rangle|_{\mathbf{k}=\Gamma} \simeq \frac{|u_\lambda(\mathbf{b}_j)\rangle - |u_\lambda(-\mathbf{b}_j)\rangle}{2|\mathbf{b}_j|}, \quad (13)$$

which can also be computed with a single Γ -only diagonalisation and leads to the following formula for the spin Chern number:

$$C_\sigma^{(\text{sym})} = -\frac{|\mathbf{b}_1||\mathbf{b}_2|}{4\pi} \text{Im} \sum_{\lambda \in \{\lambda_\sigma\}} (\langle \tilde{u}_\lambda(\mathbf{b}_1) | - \langle \tilde{u}_\lambda(-\mathbf{b}_1) |) (|\tilde{u}_\lambda(\mathbf{b}_2)\rangle - |\tilde{u}_\lambda(-\mathbf{b}_2)\rangle). \quad (14)$$

In section 3, we will show how the symmetric formula converges much faster than the asymmetric version, at essentially the same computational cost.

We have implemented the single-point formulas in a dedicated Python package, available on GitHub¹. The code provides user-friendly interfaces to two popular tight-binding (TB) packages such as PythTB² and TBmodels [54], and it can be easily interfaced to other codes.

3. Numerical results and discussion

We validate our approach on the paradigmatic Kane–Mele (KM) model [32, 55], both pristine and in presence of Anderson disorder (see figure 1). It is a model of spin $s = \frac{1}{2}$ fermions on the honeycomb lattice, whose TB Hamiltonian reads

$$H_{\text{KM}} = t \sum_{\langle i,j \rangle} c_i^\dagger c_j + \Delta \sum_i \xi_i c_i^\dagger c_i + i\lambda_{\text{SO}} \sum_{\langle\langle i,j \rangle\rangle} \nu_{ij} c_i^\dagger \sigma^z c_j + i\lambda_R \sum_{\langle i,j \rangle} c_i^\dagger (\boldsymbol{\sigma} \times \hat{\mathbf{d}}_{ij})_z c_j, \quad (15)$$

where i and j run over all sites in the lattice and the creation and annihilation operators are expressed in the contracted form $c_i^\dagger = (c_{i\uparrow}^\dagger, c_{i\downarrow}^\dagger)$. The first term is a real nearest-neighbour hopping (denoted by $\langle \cdot, \cdot \rangle$), if taken alone that would yield four (pair-degenerate) bands with gapless Dirac cones centred at the high-symmetry points \mathbf{K} and \mathbf{K}' in the BZ. The second term is a staggered on-site potential ($\xi_i = \pm 1$ is the sublattice index of the i th site) while the third term is the KM SOC [32, 55] which involves a complex next-nearest neighbour hopping (denoted by $\langle\langle \cdot, \cdot \rangle\rangle$) with a spin-dependent amplitude proportional to the Pauli matrix σ^z . The factor $\nu_{ij} = \text{sign}(\mathbf{d}_1 \times \mathbf{d}_2)_z$ depends on the orientation of the vectors \mathbf{d}_1 and \mathbf{d}_2 along the two bonds

¹ <https://github.com/roberta-favata/spinv>

² www.physics.rutgers.edu/pythtb/index.html

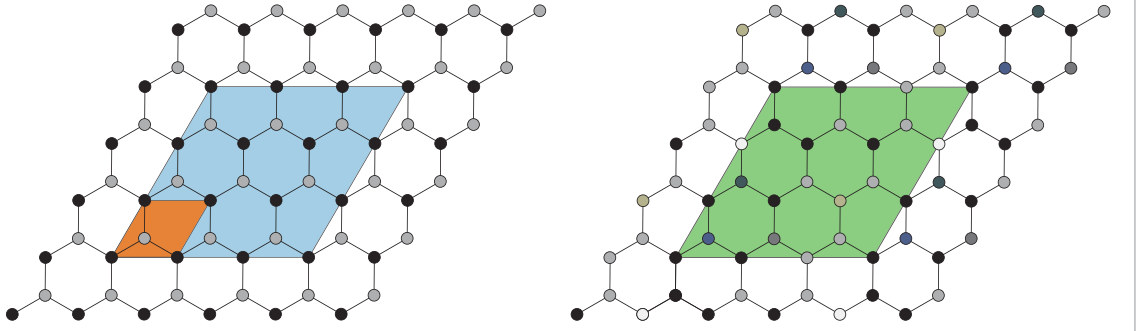


Figure 1. The Kane–Mele model in the supercell approach. Left panel: pristine Kane–Mele model, the primitive cell is shown in orange while a 3×3 supercell is marked in blue. Right panel: random realisation of a disordered Kane–Mele model in a 3×3 supercell (green) with periodic boundary conditions, where different colours are used to represent the on-site terms. In the following, supercells are labelled by their integer size $L \times L$ (in units of the pristine primitive cell) and the corresponding number of sites $N = 2L^2$.

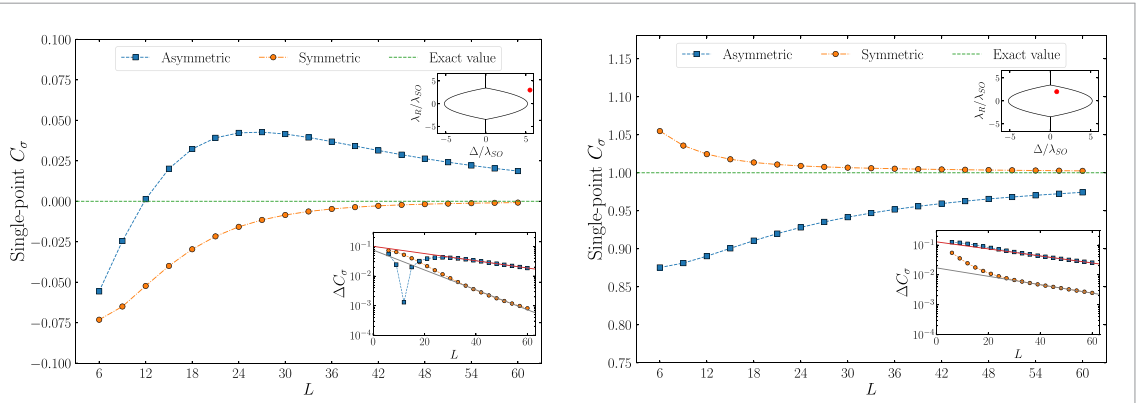


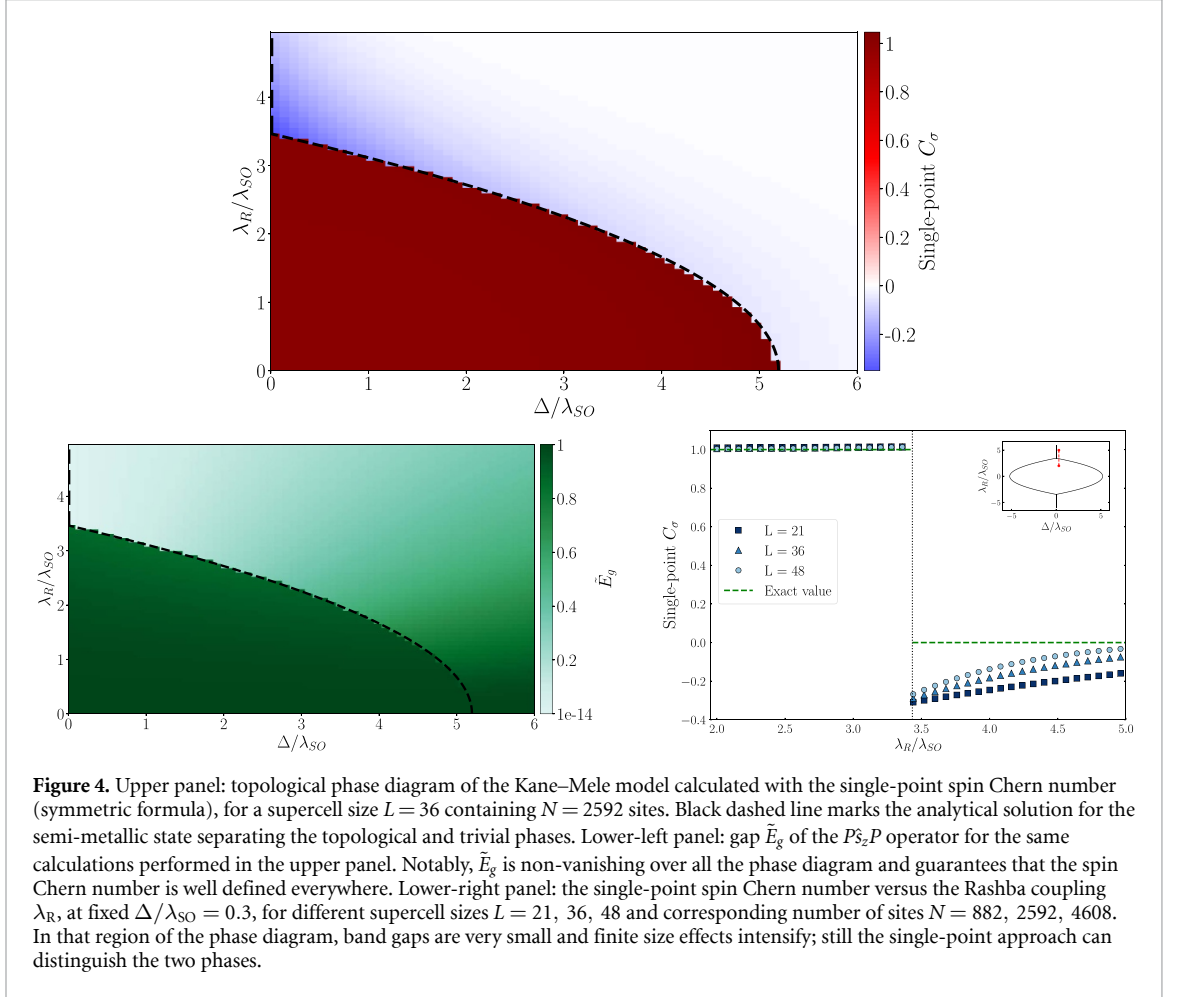
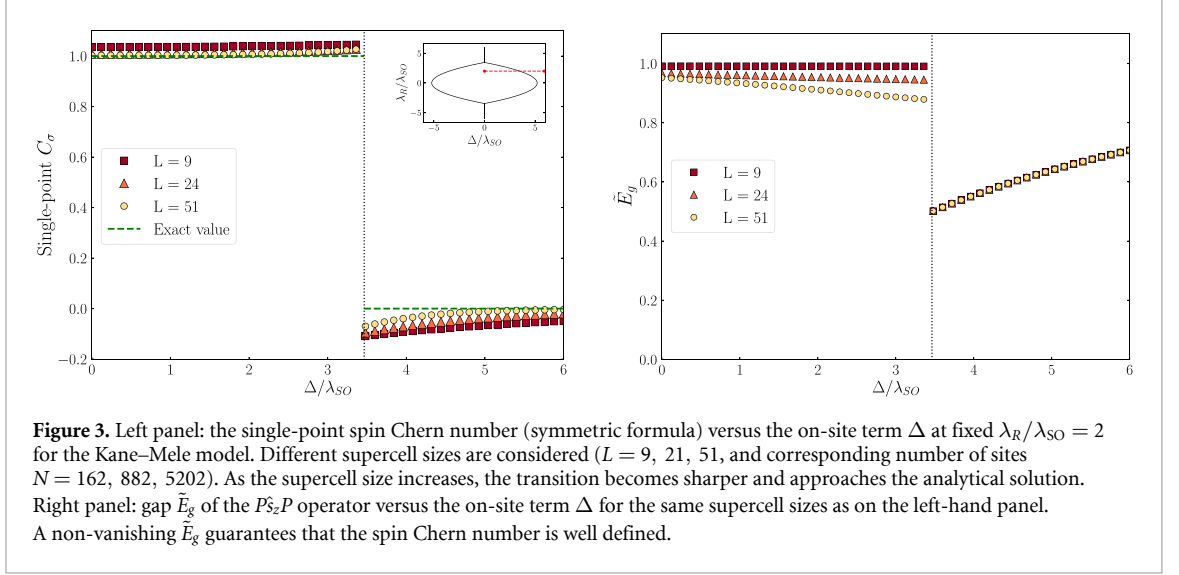
Figure 2. Convergence of the single-point spin Chern number, in its symmetric and asymmetric implementation, with respect to supercell size for the Kane–Mele model, where the Hamiltonian is diagonalised at the Γ -point only. In the upper insets, a sketch of the corresponding point in the pristine phase diagram. The lowest insets show the difference between the single-point calculations of the spin Chern number and the thermodynamic limit. Left panel: the spin Chern number converges to zero in the trivial phase ($\Delta/\lambda_{\text{SO}} = 5.5$, $\lambda_{\text{R}}/\lambda_{\text{SO}} = 3$). Right panel: in the topological phase ($\Delta/\lambda_{\text{SO}} = 0.8$, $\lambda_{\text{R}}/\lambda_{\text{SO}} = 2$) the spin Chern number converges to one. In all cases, the asymptotic convergence is exponential, but the symmetric formula converges visibly faster than its asymmetric counterpart.

connecting i to the next-nearest neighbour site j . The fourth term is the Rashba SOC and is a complex nearest-neighbour hopping with off-diagonal spin components, where $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ is the vector of Pauli matrices and $\hat{\mathbf{d}}_{ij}$ is the unit vector between sites j and i . In the following, we consider a KM Hamiltonian at fixed parameters $t = 1$ and $\lambda_{\text{SO}} = 0.03 t$, which ensure that the energy gap is insulating all over the entire phase diagram [32, 55].

3.1. Validation and convergence tests for crystalline systems

In the single-point approach, the topological invariants become exact integer numbers only in the thermodynamic limit of an infinite supercell. First, we test the convergence properties of the single-point spin Chern number (SPSCN) on the pristine KM model, in both asymmetric (equation (12)) and symmetric (equation (14)) formulation. We inspect the SPSCN as a function of the supercell size L , here defined as the number of primitive cells along each lattice vector that makes the supercell $L \times L$ (see figure 1); the number of sites inside the supercell is $N = 2L^2$. A representation of a supercell 3×3 is given in the left-hand panel of figure 1. In our calculations only values of L which are multiple of 3 are considered, to always include the special points \mathbf{K} and \mathbf{K}' folded at Γ . We benchmark the accuracy of the formulas inside the \mathbb{Z}_2 -even and \mathbb{Z}_2 -odd domains in figure 2.

The symmetric formula converges faster than the asymmetric one in both trivial and topological phases. Remarkably, the quantity $\Delta C_\sigma = |C_\sigma(L) - C_\sigma(\infty)|$, which is the difference between the spin Chern number given by the single-point formulas at finite sizes and the exact value obtained in the thermodynamic limit, decreases exponentially in both formulations. However, the global prefactor in the symmetric case is an order of magnitude smaller than the one of the asymmetric formula, leading to more accurate results at significantly smaller sizes L . Hence, in the following we adopt the symmetric formula only and study the



topological phase transition as a function of the on-site Δ , results are reported in figure 3. Our SPSCN is able to reproduce the sharp topological transition already at relatively small supercell sizes, as shown in the left-hand panel of figure 3. The band gap vanishes on the boundary of the phase transition and in the corresponding neighbourhood of parameters convergence is slower and larger supercell sizes must be employed. In the right-hand panel of figure 3, we show how the gap \tilde{E}_g of the P_z operator varies across the topological phase transition, but always remains finite, ensuring that our single-point invariant is everywhere well defined. Then, we validate the SPSCN by calculating the entire topological phase diagram of the KM model, which is reported in the upper panel of figure 4. Notably, the method can distinguish topological and

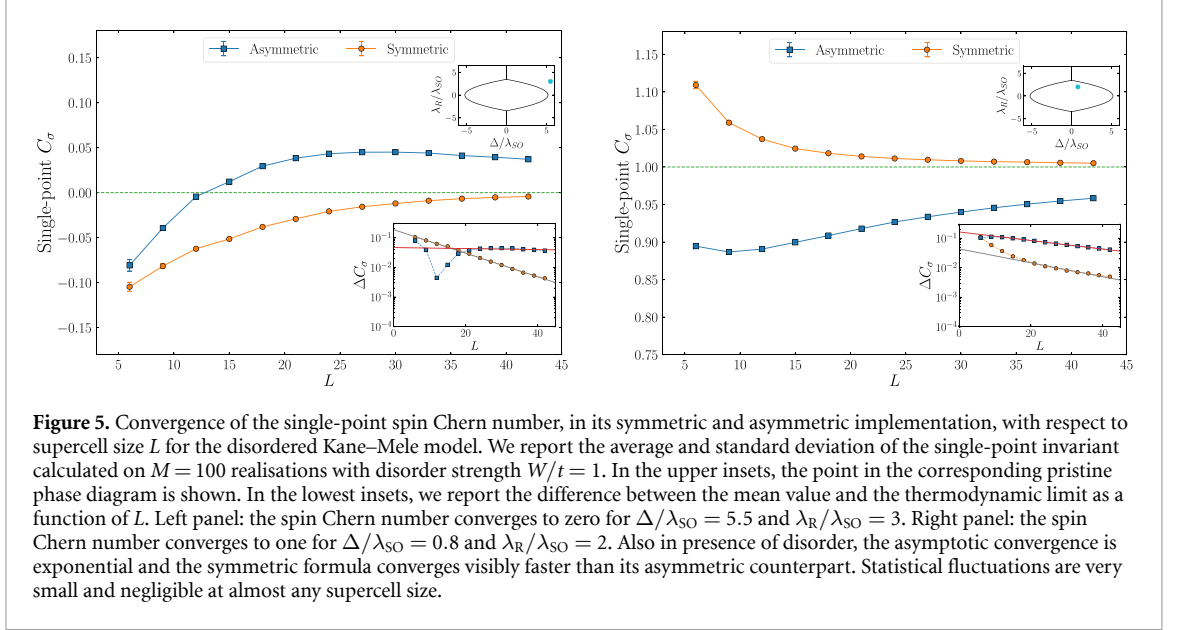


Figure 5. Convergence of the single-point spin Chern number, in its symmetric and asymmetric implementation, with respect to supercell size L for the disordered Kane–Mele model. We report the average and standard deviation of the single-point invariant calculated on $M = 100$ realisations with disorder strength $W/t = 1$. In the upper insets, the point in the corresponding pristine phase diagram is shown. In the lowest insets, we report the difference between the mean value and the thermodynamic limit as a function of L . Left panel: the spin Chern number converges to zero for $\Delta/\lambda_{\text{SO}} = 5.5$ and $\lambda_{\text{R}}/\lambda_{\text{SO}} = 3$. Right panel: the spin Chern number converges to one for $\Delta/\lambda_{\text{SO}} = 0.8$ and $\lambda_{\text{R}}/\lambda_{\text{SO}} = 2$. Also in presence of disorder, the asymptotic convergence is exponential and the symmetric formula converges visibly faster than its asymmetric counterpart. Statistical fluctuations are very small and negligible at almost any supercell size.

trivial phases even for small, but still finite, values of both the gap of the Hamiltonian and the gap of P_z (the latter is reported in the lower left-hand panel in figure 4). Larger differences between the SPSCN and the exact value (zero), which are visible in the upper-left side of the topological phase diagram (marked in blue), are finite-size effects and are reduced for large supercells, as highlighted in the lower right-hand panel in figure 4: in that region both the Hamiltonian and the P_z operator gaps are indeed very small. Therefore, our formulas work well also in presence of very strong Rashba SOC and small band gaps.

3.2. Convergence tests and topological phase transitions in presence of disorder

The presence of disorder is often modelled by means of an ensemble of large supercells, each representing a specific random realisation as schematically represented in the right-hand panel of figure 1. In electronic structure simulations, defect calculations are performed by considering large supercells, to suppress the spurious interactions due to the periodic replicas. Alloys are often simulated through the so-called special quasi-random structures [56]. In addition, a non-perturbative treatment of temperature effects always requires working with supercells, being a single structure with special atomic displacements [57] or a collection of snapshots obtained from *ab initio* molecular dynamics.

The SPSCN particularly suits this framework, and we now assess the accuracy and convergence properties of our formula on the KM model supplemented by an Anderson disorder term [58], where we highlight its capability to detect topological phase transitions due to the presence of disorder. We emphasise that the simple KM model in presence of rather strong Anderson disorder is used as a prototype and a proxy for testing, although our approach targets the more general scenario mentioned above, of supercell calculations, either for model Hamiltonians or first-principles simulations.

The Hamiltonian of the disordered KM model reads

$$H_{\text{dis}} = H_{\text{KM}} + \sum_i w_i c_i^\dagger c_i, \quad (16)$$

where $w_i \in [-\frac{W}{2}, \frac{W}{2}]$ is a randomly distributed on-site potential and W is the disorder strength which, in the following, is reported in units of the nearest-neighbour hopping amplitude t . In figure 5 we test the convergence of the single-point formulas (equations (12) and (14)) with increasing supercell size L for the disorder strength $W/t = 1$, which is weak enough not to destroy the topological phases of the corresponding pristine KM model. The SPSCN is evaluated as the mean value over M realisations of random disorder with supercells of size $L \times L$. Also in presence of disorder, the convergence of the formulas is exponential and the symmetric version converges faster than the asymmetric one.

In addition, we consider increasing disorder strengths and study the robustness of the topological phase, results are reported in figure 6. For sufficiently strong disorder, the topological phase is destroyed and the SPSCN becomes trivial. As expected, the width of the phase transition becomes smaller with increasing supercell sizes.

Furthermore, as investigated in [59], for a certain range of parameters, the disordered KM model given by equation (16) displays a topological state called topological Anderson insulator (TAI). It is a phase of

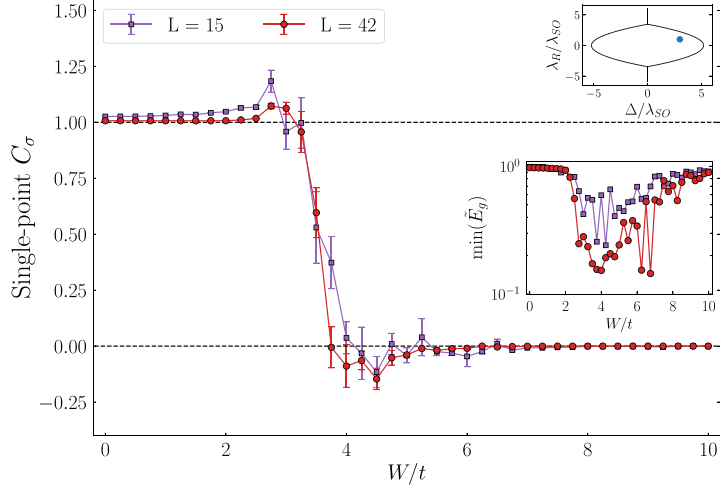


Figure 6. Robustness of the topological phase with respect to disorder. The symmetric single-point spin Chern number is calculated as a function of the disorder strength W/t , starting from the system in the topological phase ($\Delta/\lambda_{SO} = 3$, $\lambda_R/\lambda_{SO} = 1$). For each W , we report the mean and standard deviation over $M = 50$ realisations of Anderson disorder for supercells of sizes $L = 15, 42$ and number of sites $N = 450, 3528$ respectively. Upper inset: a sketch of the point where the calculations are computed reported on the pristine phase diagram ($W/t = 0$). Lower inset: minimum value, over the disorder realizations, of the gap \tilde{E}_g of the $P\tilde{s}_zP$ operator as a function of W/t . With increasing supercell size L , the transition becomes sharper. \tilde{E}_g does not vanish with Anderson disorder and the approach performs well also in the strong-disorder regime.

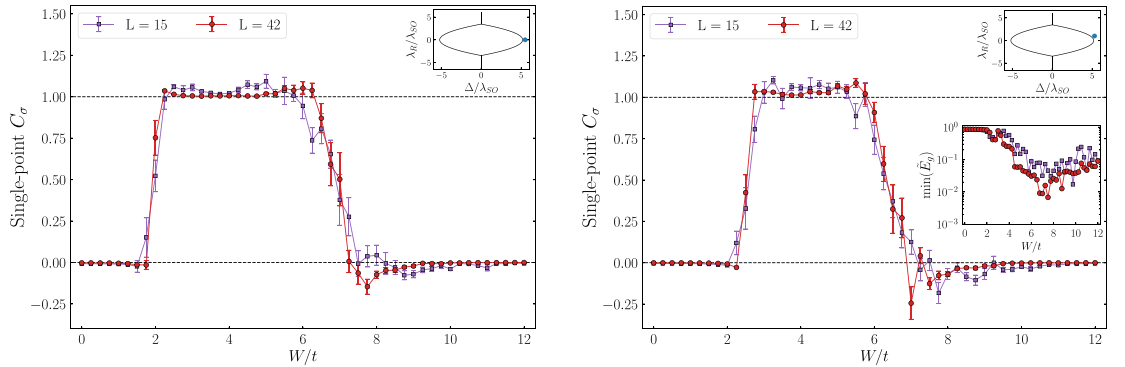


Figure 7. Topological Anderson insulator (TAI). The symmetric single-point spin Chern number is calculated as a function of the disorder strength W/t , starting from the system in a trivial state close to the phase transition. We report the mean value and the standard deviation over $M = 50$ realisations of Anderson disorder for supercells of sizes $L = 15, 42$ and corresponding numbers of sites $N = 450, 3528$ respectively. For $5 \leq W/t \leq 10$ the number of random realisations is purposely increased to $M = 100$ to reduce the standard deviation. Left panel: TAI state in absence of Rashba coupling ($\Delta/\lambda_{SO} = 5.5$, $\lambda_R = 0$). Right panel: TAI at finite Rashba coupling ($\Delta/\lambda_{SO} = 5.3$, $\lambda_R/\lambda_{SO} = 1$). Here, the minimum value of the gap \tilde{E}_g (over M disorder realizations) is reported versus W/t in the lower inset (the same plot is not present in the right-hand left panel since \tilde{E}_g is constantly equal to one for $\lambda_R = 0$).

quantized conductance which is obtained adding Anderson disorder to a trivial insulator or metal which are relatively close to a topological phase transition [60–62]. The mechanism for this disorder-induced transition has been discussed in terms of a renormalization of the model parameters such as the on-site term [61]. The weak-disorder boundary of a TAI can be studied within an effective-medium theory and the self-consistent Born approximation [59, 61], but these perturbative approach might fail in the strong-disorder regime, where the TAI phase is destroyed in favour of a trivial insulating phase, as we show next. In figure 7 we use the SPSCN to inspect the topological invariant as a function of the disorder strength. In order to compare with previous work on the disordered KM model [59] and for the sake of clarity, we consider a value of $\lambda_{SO} = 0.3 t$ which is an order of magnitude greater than the one used for the previous examples. First, we fix $\lambda_R = 0$ (left-hand panel) and observe that the TAI appears at about $W/t = 2$, in agreement with the conductance calculation in [59] and the spin Bott index results in [46] (note the factor of two with respect to the W defined therein). Then, we consider finite Rashba SOC and show the results in the right-hand panel of figure 7, where we note that the TAI region has become narrower, in agreement with [59]. A check on the gap

\tilde{E}_g of the operator P_z is performed for every SPSCN calculation in presence of disorder: Anderson disorder never fully closes the gap and the invariant can always be computed.

4. Conclusions

In this work, we have introduced a robust and efficient single-point formula to calculate the \mathbb{Z}_2 topological invariant in non-crystalline 2D materials. We have validated our method with supercell numerical simulations on the KM model, both pristine and disordered. Our approach can reproduce the entire phase diagram of the KM model, where each calculation requires only a single-point diagonalisation in the supercell framework, even in presence of strong Rashba SOC. In addition, we have extensively tested our method in presence of Anderson disorder, and we have shown how the single-point formula can correctly describe topological phase transitions appearing in the presence of strong disorder. In particular, we have discussed both the process where disorder destroys the topological phase and where disorder actually promotes it, as for the TAI phase; that is in agreement with calculations of the conductance [34, 59] and spin Bott index [46] reported in the literature. Our single-point approach converges exponentially with size, so it is typically sufficient to work with relatively small supercells, which is critical for applications in *ab initio* modelling. One of the side benefits of adopting Prodan's approach is that the formula can, at least in principle, be meaningful also in presence of weak TR-breaking perturbations [50, 63]. This feature could be useful to study how the bulk topology is affected by the presence of magnetic impurities, or of a magnetic substrate through the proximity effect; even though the absence of TR symmetry would allow backscattering between the two helical edge states. To encourage the use of our approach, we release a dedicated Python package that allows to seamlessly calculate the single-point Chern (\mathbb{Z}) and spin-Chern (\mathbb{Z}_2) invariants of any TB model thanks to dedicated interfaces to PythTB and TBmodels, two very popular TB codes. Notably, these two packages also allow working with Wannier Hamiltonians, which are read in the format produced by Wannier90 [64, 65]; that provides a simple way to apply our work in the context of first-principles calculations. Then, it would be interesting to explore the effect of the TB approximation where the real-space position operator is taken to be diagonal, versus considering all off-diagonal elements, essentially taking into account the overlap between Wannier functions. Nonetheless, the formalism is rather simple and it could be implemented with limited effort directly into plane-wave first-principles codes, such as Quantum ESPRESSO [66, 67]. In short, our approach allows studying 2D topological insulators in a supercell framework, which is crucial to investigate very relevant phenomena such as disorder, defects, alloying, and to study dynamical and temperature effects through *ab initio* molecular dynamics simulations.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.24435/materialscloud:tx-ke> [68].

Code availability

The code used in this work, implementing the single-point formulas for the spin Chern number, is available on GitHub at <https://github.com/roberta-favata/spinv> through a dedicated Python package.

ORCID iDs

Roberta Favata  <https://orcid.org/0000-0003-3042-3664>

Antimo Marrazzo  <https://orcid.org/0000-0003-2053-9962>

References

- [1] Bernevig B and Hughes T 2013 *Topological Insulators and Topological Superconductors* (Princeton, NJ: Princeton University Press)
- [2] Vanderbilt D 2018 *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge: Cambridge University Press)
- [3] Baldereschi A 1973 *Phys. Rev. B* **7** 5212–5
- [4] Resta R 1994 *Rev. Mod. Phys.* **66** 899–915
- [5] Ceresoli D and Resta R 2007 *Phys. Rev. B* **76** 012405
- [6] Mounet N *et al* 2018 *Nat. Nanotechnol.* **13** 246–52
- [7] Fu L and Kane C L 2007 *Phys. Rev. B* **76** 045302
- [8] Sgiaravello C, Peressi M and Resta R 2001 *Phys. Rev. B* **64** 115202
- [9] Soluyanov A A and Vanderbilt D 2011 *Phys. Rev. B* **83** 235401
- [10] Soluyanov A A and Vanderbilt D 2012 *Phys. Rev. B* **85** 115415
- [11] Gresch D, Autés G, Yazyev O V, Troyer M, Vanderbilt D, Bernevig B A and Soluyanov A A 2017 *Phys. Rev. B* **95** 075146

- [12] Yu R, Qi X L, Bernevig A, Fang Z and Dai X 2011 *Phys. Rev. B* **84** 075119
- [13] Alexandradinata A, Dai X and Bernevig B A 2014 *Phys. Rev. B* **89** 155114
- [14] Alexandradinata A and Bernevig B A 2016 *Phys. Rev. B* **93** 205104
- [15] Zak J 1982 *Phys. Rev. B* **26** 3010–23
- [16] Michel L and Zak J 1999 *Phys. Rev. B* **59** 5998–6001
- [17] Slager R-J, Mesaros A, Juričić V and Zaanen J 2012 *Nat. Phys.* **9** 98–102
- [18] Fang C, Gilbert M J and Bernevig B A 2012 *Phys. Rev. B* **86** 115112
- [19] Po H C, Vishwanath A and Watanabe H 2017 *Nat. Commun.* **8** 50
- [20] Khalaf E, Po H C, Vishwanath A and Watanabe H 2018 *Phys. Rev. X* **8** 031070
- [21] Song Z, Zhang T, Fang Z and Fang C 2018 *Nat. Commun.* **9** 3530
- [22] Kruthoff J, de Boer J, van Wezel J, Kane C L and Slager R J 2017 *Phys. Rev. X* **7** 041069
- [23] Bradlyn B, Elcoro L, Cano J, Vergniory M G, Wang Z, Felser C, Aroyo M I and Bernevig B A 2017 *Nature* **547** 298–305
- [24] Vergniory M G, Elcoro L, Wang Z, Cano J, Felser C, Aroyo M I, Bernevig B A and Bradlyn B 2017 *Phys. Rev. E* **96** 023310
- [25] Cano J, Bradlyn B, Wang Z, Elcoro L, Vergniory M G, Felser C, Aroyo M I and Bernevig B A 2018 *Phys. Rev. B* **97** 035139
- [26] Vergniory M, Elcoro L, Felser C, Regnault N, Bernevig B A and Wang Z 2019 *Nature* **566** 480–5
- [27] Zhang T, Jiang Y, Song Z, Huang H, He Y, Fang Z, Weng H and Fang C 2019 *Nature* **566** 475–9
- [28] Tang F, Po H, Vishwanath A and Wan X 2019 *Nature* **566** 486–9
- [29] Essin A M, Moore J E and Vanderbilt D 2009 *Phys. Rev. Lett.* **102** 146805
- [30] Leung B and Prodan E 2012 *Phys. Rev. B* **85** 205136
- [31] Guo H-M 2010 *Phys. Rev. B* **82** 115122
- [32] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 146802
- [33] Fulga I C, Hassler F and Akhmerov A R 2012 *Phys. Rev. B* **85** 165409
- [34] Fulga I C, Hassler F, Akhmerov A R and Beenakker C W J 2011 *Phys. Rev. B* **83** 155429
- [35] Avron J E, Seiler R and Simon B 1994 *Commun. Math. Phys.* **159** 399–422
- [36] Avron J, Seiler R and Simon B 1994 *J. Funct. Anal.* **120** 220–37
- [37] Katsura H and Koma T 2016 *J. Math. Phys.* **57** 021903
- [38] Katsura H and Koma T 2018 *J. Math. Phys.* **59** 031903
- [39] Akagi Y, Katsura H and Koma T 2017 *J. Phys. Soc. Japan* **86** 123710
- [40] Hastings M B and Loring T A 2010 *J. Math. Phys.* **51** 015214
- [41] Hastings M B and Loring T A 2011 *Ann. Phys., NY* **326** 1699–759
- [42] Loring T A and Hastings M B 2011 *Europhys. Lett.* **92** 67004
- [43] Liu J and Vanderbilt D 2014 *Phys. Rev. B* **90** 125133
- [44] Munoz-Segovia D, Corbae P, Varjas D, Hellman F, Griffin S M and Grushin A G 2023 arXiv:2301.02686
- [45] Huang H and Liu F 2018 *Phys. Rev. Lett.* **121** 126401
- [46] Huang H and Liu F 2018 *Phys. Rev. B* **98** 125130
- [47] Toniolo D 2022 *Lett. Math. Phys.* **112** 126
- [48] Agarwala A and Shenoy V B 2017 *Phys. Rev. Lett.* **118** 236402
- [49] Bandres M A, Rechtsman M C and Segev M 2016 *Phys. Rev. X* **6** 011016
- [50] Prodan E 2009 *Phys. Rev. B* **80** 125327
- [51] Fu L and Kane C L 2006 *Phys. Rev. B* **74** 195312
- [52] Sai N, Rabe K M and Vanderbilt D 2002 *Phys. Rev. B* **66** 104108
- [53] Souza I, Íñiguez J and Vanderbilt D 2004 *Phys. Rev. B* **69** 085106
- [54] Gresch D, Wu Q, Winkler G W, Häuselmann R, Troyer M and Soluyanov A A 2018 *Phys. Rev. Mater.* **2** 103805
- [55] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 226801
- [56] Zunger A, Wei S-H, Ferreira L G and Bernard J E 1990 *Phys. Rev. Lett.* **65** 353–6
- [57] Zacharias M and Giustino F 2020 *Phys. Rev. Res.* **2** 013357
- [58] Anderson P W 1958 *Phys. Rev.* **109** 1492–505
- [59] Orth C P, Sekera T, Bruder C and Schmidt T L 2016 *Sci. Rep.* **6** 24007
- [60] Li J, Chu R-L, Jain J K and Shen S-Q 2009 *Phys. Rev. Lett.* **102** 136806
- [61] Groth C W, Wimmer M, Akhmerov A R, Tworzydło J and Beenakker C W J 2009 *Phys. Rev. Lett.* **103** 196805
- [62] Xing Y, Zhang L and Wang J 2011 *Phys. Rev. B* **84** 035110
- [63] Saha S, Nag T and Mandal S 2021 Eightfold quantum Hall phases in a time reversal symmetry broken tight binding model *Phys. Rev. B* **103** 235154
- [64] Mostofi A A, Yates J R, Lee Y-S, Souza I, Vanderbilt D and Marzari N 2008 *Comput. Phys. Commun.* **178** 685–99
- [65] Pizzi G *et al* 2020 *J. Phys.: Condens. Matter* **32** 165902
- [66] Giannozzi P *et al* 2009 *J. Phys.: Condens. Matter* **21** 395502
- [67] Giannozzi P *et al* 2017 *J. Phys.: Condens. Matter* **29** 465901
- [68] Favata Roberta and Marrazzo Antimo 2023 Materials Cloud Archive 2023.22 (<https://doi.org/10.24435/materialscloud:tx-ke>)