Topology of Bi$_2$Se$_3$ nanosheets

L. Maisel Licerán [1,*] S. J. H. Koerhuis,[1] D. Vanmaekelbergh [2] and H. T. C. Stoof [1]

[1]*Institute for Theoretical Physics and Center for Extreme Matter and Emergent Phenomena, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands*

[2]*Debye Institute for Nanomaterials Science, Leonard S. Ornsteinlaboratorium, Utrecht University, Princetonplein 1, 3584 CC Utrecht, The Netherlands*

Recently, the quantum spin-Hall edge channels of two-dimensional colloidal nanocrystals of the topological insulator Bi$_2$Se$_3$ were observed directly. Motivated by this development, we reconsider the four-band effective model which has been traditionally employed in the past to describe thin nanosheets of this material. Derived from a three-dimensional $\mathbf{k} \cdot \mathbf{p}$ model, it physically describes the top and bottom electronic surface states at the Γ point that become gapped due to the material's small thickness. However, we find that the four-band model for the surface states alone, as derived directly from the three-dimensional theory, is inadequate for the description of thin films of a few quintuple layers and even yields an incorrect topological invariant within a significant range of thicknesses. To address this limitation we propose an eight-band model which, in addition to the surface states, also incorporates the set of bulk states closest to the Fermi level. We find that the eight-band model not only captures most of the experimental observations, but also agrees with previous first-principles calculations of the $\mathbb{Z}_2$ invariant in thin films of varying thickness. The band inversion around the Γ point, which endows the surfacelike bands with topology, is shown to be enabled by the presence of the additional bulklike states without requiring any reparametrization of the resulting effective Hamiltonian.

I. INTRODUCTION

Topological insulators (TIs), and more generally topological materials, have experienced a massive surge in interest in the past decade due to their excellent prospects for energy-efficient electronics, spintronics, and transport applications [1–15]. A common feature shared by most conventional TIs is the appearance of protected states at the boundaries of a finite sample, which are typically perfectly conducting and whose dispersions are linear and cross the semiconducting band gap. In three dimensions (3D) these take the form of surface states, while in two dimensions (2D) they are realized as edge states. The properties of these boundary modes strongly depend on which symmetries are present in a given system. Arguably, the best known examples are the integer quantum Hall state [16–21], where chiral electrons flow without dissipation along the edge due to breaking of the time-reversal symmetry (TRS), and the quantum spin-Hall (QSH) state [22–29], which preserves the TRS and can be seen as two spin-reversed copies of the quantum Hall state with opposite spins flowing in opposite directions.

Within the vast landscape of topological materials, bismuth selenide (Bi$_2$Se$_3$) is often quoted as a prototypical three-dimensional TI [30–35]. Its atomic structure comes in the form of quintuple layers (QLs) stacked on top of one another and bound together by van der Waals forces [30,36–38]. The topological nature of Bi$_2$Se$_3$ has been extensively investigated both theoretically and experimentally and can be traced back to the existence of a large inverted topological gap arising from the spin-orbit interaction. This leads to topologically protected surface states at the top and bottom surfaces of planar slabs. These have been fully characterized theoretically [39–42] by means of a continuum model for three-dimensional Bi$_2$Se$_3$ around the Γ point derived from $\mathbf{k} \cdot \mathbf{p}$ theory [30,36]. A particularly interesting topic in this regard is the transition from 3D to 2D samples of Bi$_2$Se$_3$, as the dispersion of the surface states becomes gapped due to the hybridization between the modes at both surfaces, and the corresponding gap may or may not be of topological nature. The gapping of the top and bottom surface states has been observed experimentally via angle-resolved photoemission spectroscopy [43–45] and addressed theoretically in some of the aforementioned studies [39–41]. In the latter, an effective nanosheet Hamiltonian is obtained for each number of QLs by projecting the full 3D bulk Hamiltonian onto the subspace of four Γ-point surface states that appear due to the nontrivial topology and which are closest to the Fermi level. For samples of Bi$_2$Se$_3$ with a thickness of a few QLs, these theoretical studies predict a series of topological phase transitions between the 3D and 2D models, in which the material repeatedly oscillates between a trivial phase and a QSH phase. Focusing on the ultrathin regime between 1 and 6 QLs, as for larger thicknesses the surface-state gap becomes negligible, nanosheets of 6 QLs are predicted to show QSH behavior, while those with 5 and 4 QLs are found to be trivial, followed by another QSH phase at 3 QLs, and finally remaining in a trivial phase for 2 and 1 QLs. However, this

*l.maiselliceran@uu.nl

is at odds with state-of-the-art first-principles calculations, which predict a nonzero $\mathbb{Z}_2$ invariant for 3 to 5 QLs and a trivial phase for 1 and 2 QLs [39]. It also disagrees with a recent experiment on colloidal Bi_2Se_3 nanosheets of finite lateral dimensions, where QSH edge states were observed directly in the regime between 4 and 6 QLs [46]. Thus it may seem as though the applicability of the continuum model in the regime between 4 and 6 QLs can be put into question even from the point of view of qualitative predictions in Bi_2Se_3.

It thus appears that the effective four-band model obtained in this way does not provide a complete understanding of the physics in the entire regime between 1 and 6 QLs. Here, we demonstrate that some of the oscillations predicted by the aforementioned theoretical studies are spurious and disappear when one takes into account not only the gapped surface states closest to the Fermi level, but also the first set of bulk conduction and valence states. The resulting eight-band Hamiltonian is not only capable of describing the physics in a broader range of momenta, but also reproduces most of the experimental findings and features of *ab initio* calculations. Our results show that the three-dimensional continuum model for topological Bi_2Se_3 can be successfully used even in the ultrathin limit of 1 QL and also elucidates the crucial role of the bands further away from the Fermi level, which are usually neglected.

This article is organized as follows. In Sec. II we present a general description of Bi_2Se_3 nanosheets by starting from the 3D bulk model and compare the features of the four-band model for the surface states with our eight-band model including also the bulk states closest to the Fermi level. In Sec. III we analyze the topology of the eight-band model by studying the interplay between the different bands. In Sec. IV we validate our model by comparing its topological properties with those calculated from first principles as well as with the most recent experimental results. Finally, in Sec. V we summarize our work and give an outlook on potential future research.

II. EFFECTIVE MODEL FOR Bi_2Se_3 NANOSHEETS

Our starting point is the $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian for three-dimensional Bi_2Se_3 derived in Refs. [30,36]. This model is expressed in the combined orbital-spin basis $(|Bi^+, \uparrow\rangle, |Se^-, \uparrow\rangle, |Bi^+, \downarrow\rangle, |Se^-, \downarrow\rangle)$ that is closest to the Fermi surface, where Bi^+ and Se^- are hybridized Bi and Se p_z orbitals of even and odd parity, respectively. The effective Hamiltonian reads

$$H(\boldsymbol{k}, k_z) = \epsilon_0(\boldsymbol{k}, k_z)\mathbb{I}_s \otimes \mathbb{I}_\tau + \mathcal{M}(\boldsymbol{k}, k_z)\mathbb{I}_s \otimes \tau_z$$
$$+ A_1 k_z s_z \otimes \tau_x + A_2(\boldsymbol{k} \cdot \boldsymbol{s}) \otimes \tau_x. \quad (1)$$

Here, $\epsilon_0(\boldsymbol{k}, k_z) = C + D_1 k_z^2 + D_2 k^2$, $\mathcal{M}(\boldsymbol{k}, k_z) = M - B_1 k_z^2 - B_2 k^2$, $\boldsymbol{s}$ and $\boldsymbol{\tau}$ are the Pauli matrices in the spin and orbital spaces, respectively, and $\mathbb{I}_s$ and $\mathbb{I}_\tau$ are the identity matrices in these respective subspaces. Here and below, $\boldsymbol{k} \equiv (k_x, k_y)$ denotes the in-plane momentum, with $k \equiv |\boldsymbol{k}|$. We employ the parameters of the paper by Zhang *et al.* [30], which have been fitted

to their *ab initio* calculation. Their numerical values are $C = -0.0068\,\mathrm{eV}$, $M = 0.28\,\mathrm{eV}$, $A_1 = 0.22\,\mathrm{eVnm}$, $A_2 = 0.41\,\mathrm{eVnm}$, $B_1 = 0.10\,\mathrm{eVnm}^2$, $B_2 = 0.566\,\mathrm{eVnm}^2$, $D_1 = 0.013\,\mathrm{eVnm}^2$, and $D_2 = 0.196\,\mathrm{eVnm}^2$.

We are interested in thin nanosheets of only a few QLs. This geometry breaks the translational invariance in the z direction and thus we must solve the model of Eq. (1) after substituting $k_z \to -i\partial_z$. It is customary to employ hard-wall boundary conditions for the wave functions at both surfaces, $\Psi(z = \pm L_z/2) = 0$, with L_z the nanosheet thickness. To obtain a low-energy effective model, one first solves the Hamiltonian at the 2D Γ point $k = 0$. The solutions to this problem are described in detail in Refs. [40–42]. One then projects the full Hamiltonian at nonzero $\boldsymbol{k}$ onto a subset spanned by these solutions. The size of this subspace essentially determines the validity of the effective model around $k = 0$. Only when projecting on all states is full equivalence with the higher-dimensional Hamiltonian recovered.

Note that, of course, one can in principle also solve the full 3D model above at arbitrary $\boldsymbol{k}$, without needing to resort to an effective model (cf. Fig. 5 below). However, this requires solving a boundary eigenproblem for each $\boldsymbol{k}$, which is undesirable for many purposes, e.g., computing many-body properties involving interaction matrix elements at different momenta or computing Chern numbers or other topological invariants where the wave functions must be known over the entire Brillouin zone. By contrast, the use of an effective model allows one to solve the boundary problem at a single momentum and thereafter simply diagonalize a (typically small) $\boldsymbol{k}$-dependent matrix. For this reason, it is often useful to work with a reliable effective Hamiltonian whose z dependence has been integrated out.

As 3D Bi_2Se_3 is a TI, the spectrum obtained from the Hamiltonian in Eq. (1) for large thicknesses contains states that are localized at the surfaces of the nanosheet under consideration. Their dispersions around the Γ point form a gapless Dirac cone that crosses the semiconductor band gap. As one decreases L_z to a few QLs, the Dirac cone becomes gapped at Γ as a result of the hybridization between the states localized at opposite surfaces. We are interested in thicknesses small enough for this gap to be of observable magnitude. It is not until a thickness of around $L_z \simeq 6\,\mathrm{nm}$ that this gap becomes of the order of $1\,\mathrm{meV}$, so we will focus on the regime $L_z \lesssim 6\,\mathrm{nm}$. Note that the thickness of a single QL is approximately $1\,\mathrm{nm}$ [47,48], so that the thickness of the nanosheet in nanometers is in good approximation the number of QLs. We also mention that, in this ultrathin limit, what is meant by "surface states" are those states whose dispersions evolve from the gapless Dirac point at large L_z, even though their wave functions are no longer strongly localized due to the aforementioned hybridization. The remaining spectrum of the Hamiltonian (1) at $k = 0$ consists of states whose energies are gapped for all L_z and whose wave functions never show strong localization; we call these the bulk states. Finally, we clarify that, because our starting model is derived from a low-energy expansion, it does not contain the surface state replicas described by Kung *et al.* located deep inside the bulk energy region [49]. However, it is safe to assume that their effects can be neglected for the purposes of this article.

A. Four-band model

Previous works on the model considered here have always assumed that projecting on the topological surface states alone is enough to obtain an accurate low-energy model for any thickness. However, as we show here, this introduces some issues and in particular the topology of the resulting model does not align with experimental findings and density-functional theory (DFT) calculations.

At the 2D Γ point, the Hamiltonian decouples into two subspaces of opposite spin and its spectrum contains four electronic surface states close to the Fermi level. These states have well-defined spin and parity, so we denote them by $|S^{\pm}, \uparrow(\downarrow)\rangle$, where the label S indicates that they are surface states [50].

We define the basis $(|S^+, \uparrow\rangle, |S^-, \downarrow\rangle, |S^-, \uparrow\rangle, |S^+, \downarrow\rangle)$, in which the effective four-band Hamiltonian for the surface states takes the form

$$H^{\text{eff}}_{\text{4-band}, \xi}(\boldsymbol{k}) = (E_0 - Dk^2)\mathbb{I}_{\sigma} + v_{\text{F}}(\boldsymbol{\sigma} \times \boldsymbol{k}) \cdot \hat{\boldsymbol{z}}$$
$$- \xi\left(\frac{\Delta}{2} - Bk^2\right)\sigma_z. \quad (2)$$

There are two subblocks with $\xi = \pm 1$, sometimes called the hyperbola index, as the Hamiltonians $H^{\text{eff}}_{\text{4-band}, \pm}(\boldsymbol{k})$ resemble those at the K and K' points of graphene but shifted to the Γ point. Here, however, we choose to call ξ the *spin-orbit parity* (SOP), as the product of spin times orbital parity in each subspace is precisely ξ if we identify $\uparrow(\downarrow)$ with $+1(-1)$. Furthermore, $\boldsymbol{\sigma}$ are Pauli matrices that couple the basis states $(|S^+, \uparrow\rangle, |S^-, \downarrow\rangle)$ for $\xi = +1$ and $(|S^-, \uparrow\rangle, |S^+, \downarrow\rangle)$ for $\xi = -1$, and $\mathbb{I}_{\sigma}$ stands for the identity in each of these subspaces of fixed SOP. We note that, even though the states in the doublets $(|S^+, \uparrow\rangle, |S^-, \downarrow\rangle)$ and $(|S^-, \uparrow\rangle, |S^+, \downarrow\rangle)$ have opposite spin, $\boldsymbol{\sigma}$ is *not* directly the physical spin operator, as the wave functions of the up and down spin states in each subspace are not equal. Instead, they correspond to two distinct hybridized orbitals, so that $\boldsymbol{\sigma}$ is more appropriately understood as a pseudospin that mixes the orbital and spin degrees of freedom.

Each 2×2 subblock has two bands with dispersions $\varepsilon_{c,v}(k) = E_0 - Dk^2 \pm \sqrt{(\Delta/2 - Bk^2)^2 + (v_{\text{F}}k)^2}$. The corresponding eigenstates are

$$\psi^{c,v}_{\xi\boldsymbol{k}} = \mathcal{N}^{c,v}_{\xi\boldsymbol{k}}\left[\begin{array}{c} -\xi\left(\frac{\Delta}{2} - Bk^2\right) \pm \sqrt{\left(\frac{\Delta}{2} - Bk^2\right)^2 + (v_{\text{F}}k)^2} \\ -iv_{\text{F}}k_+ \end{array}\right],$$
$$(3)$$

where $k_{\pm} = k_x \pm ik_y$ and $\mathcal{N}^{c,v}_{\xi\boldsymbol{k}}$ is readily found by normalizing the eigenstates to unity. The Chern number of these states is given by $\mathcal{C}^{\xi}_{c,v} = \mp\frac{\xi}{2}(\text{sgn}\,\Delta + \text{sgn}\,B)$. This is nonzero only when Δ and B have the same sign. Physically, this may be understood from the fact that the topology arises from a band inversion, which in the four-band model is driven by the combined action of Δ and B. The fact that the band inversion leads to a topologically nontrivial phase is rooted in the parity flip that takes place as one goes from the Γ point to large momenta. Due to the inversion, the valence band around Γ has opposite parity to that when $k \to \infty$. This argument was first formalized by Fu and Kane [51] and in systems with inversion symmetry allows one to determine the $\mathbb{Z}_2$ invariant

by evaluating the parity of the eigenstates at the time-reversal-invariant momenta. This procedure will be paramount in our description of the topology in the eight-band model below.

The Hall conductivity of each subspace is then $\sigma^{xy}_{\xi} = (e^2/h)\mathcal{C}^{\xi}_v$ provided that the Fermi level stays within the gap with decreasing thickness. The superposition of two opposite Hall conductivities causes the total σ^{xy} to vanish, but if the individual conductivities are nonzero the system is in a QSH phase. However, strictly speaking this QSH effect is in terms of the pseudospin of the underlying basis and *not* in terms of the z-direction component of the real electronic spin. This can be easily seen from the fact that each 2×2 subblock mixes up and down spins, contrary to the prototypical Bernevig-Hughes-Zhang model for the QSH effect, in which the two subspaces separately describe spin-up and spin-down electrons [23,24]. As a result, there is a nontrivial spin texture along the z direction given by $\langle S_i \rangle^{c,v}_{\xi\boldsymbol{k}}(z) = \psi^{c,v}_{\xi\boldsymbol{k}}(z)^{\dagger} S_i \psi^{c,v}_{\xi\boldsymbol{k}}(z)$, with $\psi^{c,v}_{\xi\boldsymbol{k}}(z) = \langle z|S^{\xi}, \uparrow\rangle\langle S^{\xi}, \uparrow|\psi^{c,v}_{\xi\boldsymbol{k}}\rangle + \langle z|S^{-\xi}, \downarrow\rangle\langle S^{-\xi}, \downarrow|\psi^{c,v}_{\xi\boldsymbol{k}}\rangle$ and $S_i = \frac{1}{2}s_i \otimes \mathbb{I}_{\tau}$ the spin operator. Note that $\langle z|S^{\pm}, \uparrow(\downarrow)\rangle$ are simply the wave functions corresponding to the surface states at the Γ point introduced in the previous section, which are z-dependent four-component vectors in the orbital-spin basis.

Similarly, the edge states of this Hamiltonian at the boundaries of a finite sample also present a nontrivial spin texture along the vertical direction. An analysis of these edge modes reveals that the physical spin is always perpendicular to the momentum along the edge. More precisely, there is a nonvanishing projection in the z direction, whose average over the nanosheet thickness is in general nonzero. Furthermore, the spin in the direction perpendicular to the edge and parallel to the nanosheet has a nontrivial texture along the vertical direction, but its average over z vanishes. Finally, the spin in the direction parallel to the edge is identically zero for all z. Consequently, we recover the well-known QSH picture, realized now in terms of the z-averaged vertical component of the real electronic spin. We emphasize that this is not necessarily clear a priori, given that the underlying basis mixes the up and down components as explained above. A sketch of the situation is shown in Fig. 1, where the edge is taken along the x direction and thus the spin lies entirely in the yz plane. For nonzero k_x, the edge states always follow a linear dispersion, given by

$$\varepsilon^{\pm}_{\text{edge}}(k_x) = E_{\Gamma} \pm \tilde{v}_{\text{F}}k_x, \quad (4)$$

where $\tilde{v}_{\text{F}} = v_{\text{F}}\sqrt{1 - D^2/B^2}$ is a renormalized Fermi velocity which is lower than that of the surface states and E_{Γ} is the energy at the one-dimensional Γ point $k_x = 0$. We note that all of this is valid only if $|D| < |B|$, as otherwise the global energy gap disappears and no edge states are found.

The form of Eq. (1) is indeed capable of describing the surface states of electrons in Bi$_2$Se$_3$ nanosheets, as previous experiments have successfully matched the observed band structure with that arising from the effective model [43]. However, there is a caveat: it is crucial to realize that a faithful description of the system in terms of the four-band model requires adjusting the numerical values of the parameters in the low-energy Hamiltonian to a set of experimentally
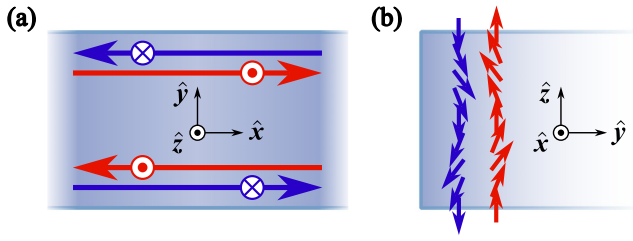
FIG. 1. Sketch of the spin behavior in the QSH regime of a thin Bi_2Se_3 nanoribbon along the x direction. (a) Top view of the nanoribbon. The long arrows along the edges show the velocity of the edge states and the out-of-plane arrows indicate the direction of the z-averaged electron spin. Note that the average spin polarization is in general not 100%, even though its direction of polarization is perpendicular to the nanosheet. (b) Side view of the nanoribbon. The small arrows show the nontrivial microscopic spin texture along the nanosheet thickness. This texture is such that there is a z-dependent tilt in the yz plane. Its average over the nanosheet thickness vanishes in the y direction, but is nonzero in the z direction. Thus the usual QSH picture emerges in terms of this nonvanishing vertical component, as shown in (a).

determined values. This is not only completely *ad hoc*, but also only possible if experiments in the 2D limit are available. Since this may not be the case in other similar materials, it is desirable to have a reliable scheme which allows 2D topological properties to be inferred solely from the 3D bulk model. In practice, the actual theoretical prediction of the projected four-band model for Bi_2Se_3 is far from satisfactory, as the parameters obtained directly through the projection procedure give a band structure that disagrees with the experimental findings. The most important aspect in this regard is that the topological properties appear incorrect for a relevant range of thicknesses. The reason is that the parameters Δ and B, which together determine the topology, have opposite signs for 4 and 5 QLs, a regime where both experiments [46] and DFT calculations [39] show the presence of QSH edge states. More precisely, in this picture the model oscillates four times between a trivial and a QSH phase between 1 and 6 QLs. Two of these oscillations are due to the gap closing and two more seemingly take place because, while the gap remains open, the parameter B changes sign and the band inversion disappears. This phenomenon is in fact an artifact of such a continuum model: placing the model on a lattice reveals that a change in Chern number is always accompanied by a gap closing, but depending on Δ and B this may happen at the edges of the Brillouin zone instead of at the Γ point. This is problematic, because it would indicate that the low-energy subspace we wish to study is not located at the Γ point for a certain range of thicknesses. As we will see, however, enhancing the model to an eight-band Hamiltonian eliminates this spurious change in Chern number. Hence, for all thicknesses, the system is still described by the physics around Γ, thus providing the physically expected picture. Moreover, for $L_z \gtrsim 3.23$ nm it is found in the four-band model that $|D| > |B|$, which means that the valence band is not inverted and thus actually *grows* in energy when $k \to \infty$. This leads to the absence of a global gap in the spectrum, which in this system leads to no edge states even if the Chern number of the valence band is nonzero. Once again,

this is undesirable and in contradiction with experiments on Bi_2Se_3 nanosheets and DFT calculations.

B. Eight-band model

We have demonstrated that projection on the surface states alone is not enough to obtain accurate results for the low-energy physics of Bi_2Se_3 nanosheets. We now proceed to include in this projection also the first set of bulk states that arises from solving the model at the two-dimensional Γ point, which we denote by $|B^{\pm}, \uparrow (\downarrow)\rangle$. As we explain below, this is enough to solve all issues present in the previous model.

One can show that the effective Hamiltonian always decouples into two separate subspaces, which are related by TRS. This is a consequence of the TRS in combination with the mirror symmetry with respect to the xy plane, as the latter enforces a definite parity for the individual components of the Γ-point wave functions under the operation $z \to -z$. In other words, the spin-orbit parity ξ is always a well-defined quantum number in the presence of both TRS and planar mirror symmetry. For our combined eight-band Hamiltonian we choose the basis

$$\underbrace{(|S^+, \uparrow\rangle, |S^-, \downarrow\rangle, |B^+, \uparrow\rangle, |B^-, \downarrow\rangle,}_{\xi=+1}$$
$$\underbrace{|S^-, \uparrow\rangle, |S^+, \downarrow\rangle, |B^-, \uparrow\rangle, |B^+, \downarrow\rangle).}_{\xi=-1} \quad (5)$$

Due to the TRS, we focus on the analysis of the first 4×4 subblock, with SOP $\xi = +1$. Its Hamiltonian reads

$$H_{8\text{-band},\xi=+1}^{\text{eff}}(\boldsymbol{k}) = \begin{bmatrix} H_{SS} & H_{SB} \\ H_{SB}^\dagger & H_{BB} \end{bmatrix}, \quad (6)$$

where

$$H_{II} = \epsilon_0^I(\boldsymbol{k})\mathbb{I}_{2\times2} + \begin{bmatrix} \mathcal{M}^I(\boldsymbol{k}) & (A^I)^* k_- \\ A^I k_+ & -\mathcal{M}^I(\boldsymbol{k}) \end{bmatrix}, \quad (7a)$$

$$H_{SB} = \begin{bmatrix} ak^2 & bk_- \\ ck_+ & dk^2 \end{bmatrix}, \quad (7b)$$

with $I \in \{S, B\}$, $\epsilon_0^I(\boldsymbol{k}) = C^I + D^I k^2$, and $\mathcal{M}^I(\boldsymbol{k}) = M^I - B^I k^2$. All parameters $(M^I, A^I, B^I, C^I, D^I, a, b, c, d)$ depend on the thickness L_z and together determine the topology of the corresponding nanosheet. It is important to realize that their numerical values are unambiguously determined from the initial set of parameters of the 3D bulk Hamiltonian, thus not requiring any further adjustments. In Fig. 2 we show the band structure of the eight-band model for 4 QLs in the subspace $\xi = +1$. There are four bands which we call the upper and lower valence or conduction band (from top to bottom: UCB, LCB, UVB, and LVB). The topological properties can be understood by tracking the spin projection of the different bands as a function of momentum while taking care of some subtleties detailed in the next section.

We have computed the spectrum of each subspace separately on a ribbon along the x direction. The width in the y direction is taken as $L_y = 100$ nm and we solve the continuum model by substituting $k_y \to -i\partial_y$ and employing hard-wall boundary conditions for the wave function at the edges. The
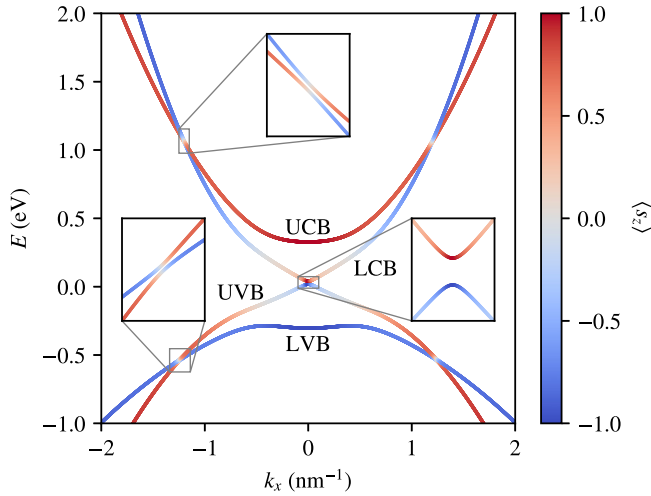
FIG. 2. Band structure of a spatially extended nanosheet of 4 QLs in thickness for the subspace with SOP $\xi = +1$, as given by the eight-band model. The band-name abbreviations are defined in the main text. Here we have set $k_y = 0$, as the dispersion is rotationally symmetric in our $\boldsymbol{k} \cdot \boldsymbol{p}$ model. The energy bands of the $\xi = -1$ subspace are degenerate with these ones, but their spin is opposite due to the TRS. Avoided crossings are found at $|k_x| \simeq 1.2\,\text{nm}^{-1}$, at which the two valence or conduction bands exchange their spin. As explained in detail in the main text, this gives a nontrivial twist to the LVB and the UCB, which become topological in the effective model, while effectively undoing the twist of the UVB and the LCB visible around $|k_x| \simeq 0.5\,\text{nm}^{-1}$ and making them trivial. Contrary to the four-band model for 4 QLs, the eight-band model shows a global gap, meaning that when $k \to \infty$ all conduction and valence bands go to large positive and negative energies, respectively. The insets show magnifications of the avoided crossings and the surface gap.

results for 1–5 QLs are plotted in Fig. 3. In all cases the valence bands are fully inverted, i.e., they go down in energy when $k \to \infty$, meaning that for $L_z \gtrsim 3.23\,\text{nm}$ the bulk states are required to obtain a fully gapped spectrum. For 1 and 2 QLs the spectrum is devoid of edge states, as already happens in the four-band model. However, we now find edge states not only for 3, but also for 4 and 5 QLs, in contrast to the situation in the four-band model. In the first row of the figure we have colored the states according to their surface or bulk character and the edge states are shown according to their localization along the width of the ribbon. In the second row we show the states colored according to their average spin in the z direction. It is worth noting that the picture of Fig. 1 remains valid, except for the fact that the precise spin texture is now also influenced by the wave functions of the newly added bulk states. We comment on the case of 5 QLs, where there seems to be no Dirac point. This is due to the fact that the gap closing takes place at a thickness of 5.02 nm, so that the last panel of Fig. 3 is extremely close to the transition. As such, the gap is so small that the upper half of the Dirac cone, which is slightly gapped due to finite-size effects, has already merged with the lowest conduction band. Although not shown here, this can be verified by tracking the evolution of the edge conduction band at intermediate thicknesses between 4 and 5 QLs, which poses no complications in a continuous model.

III. ANALYSIS OF THE TOPOLOGY

To determine the topological protection of the aforementioned edge states it is necessary to compute the $\mathbb{Z}_2$ invariant of the system, as the total 8×8 Hamiltonian lies in class AII [13]. However, the system decouples into two subspaces that get interchanged under time reversal, so in this case the $\mathbb{Z}_2$ classification is equivalent to a double Chern-number classification [52]. The topology can thus be equally determined from the Chern numbers of each subblock, as they individually break the TRS. A distinction will have to be made between *physical* Chern numbers (PCNs) and *effective-model* Chern numbers (EMCNs). The latter refer to those obtained by directly integrating the Berry curvatures of the eight-band model in the entire range of k, which is unbounded in our continuum model. We will argue that they generally differ from the PCNs, i.e., the Chern numbers which should be considered physical, albeit the Hall conductivity stays the same. These PCNs will be defined later and can be easily inferred once we have gained an intuitive understanding of the features of Fig. 2.

We have numerically computed the EMCNs of the four bands of each subspace. The total Hall conductivity is then $(e^2/h) \sum_i \mathcal{C}_i^{\xi}$, where the sum runs over the occupied bands. In the range between the gap closings at $L_z = 2.51\,\text{nm}$ and $L_z = 5.02\,\text{nm}$, we find that the EMCN of the UVB is zero, whereas that of the LVB band is nontrivial and equal to ± 1 in each subblock. The Hall conductivity is thus nonzero and we have accordingly found that edge states are present. For $L_z < 2.51\,\text{nm}$ and $5.02\,\text{nm} < L_z \leqslant 6\,\text{nm}$, we actually find that *both* occupied bands of each subblock have opposite unit EMCNs. Thus the total Hall conductivity vanishes and there is no spin-Hall current.

For completeness, we have also explicitly calculated the $\mathbb{Z}_2$ invariant directly by analyzing the Pfaffian of the matrix

$$A_{ij}(\boldsymbol{k}) = \langle \psi_{i\boldsymbol{k}} | \Theta | \psi_{j\boldsymbol{k}} \rangle, \qquad (8)$$

which we denote by $P(\boldsymbol{k}) \equiv \text{Pf}[A(\boldsymbol{k})]$. Here, Θ is the time-reversal operator and one takes only the eigenstates $|\psi_{i\boldsymbol{k}}\rangle$ whose energy lies below the Fermi level. In our eight-band model, the time-reversal operator is represented by the matrix $\Theta = -i(\sigma_x^{\text{SOP}} \otimes \sigma_z^{\text{SB}} \otimes \sigma_y)\mathcal{K}$, where σ_x^{SOP} acts on the two uncoupled subspaces with opposite SOP, σ_z^{SB} acts on the surface-bulk degree of freedom, σ_z acts on the pseudospin defined in Sec. II A (generalized to surface and bulk orbitals), and $\mathcal{K}$ is the complex conjugation operator. The topological invariant is determined by analyzing the phase of $P(\boldsymbol{k})$ along a contour C that encloses precisely *half* of the Brillouin zone or, in the case of our continuum model, half of the infinite momentum plane. The $\mathbb{Z}_2$ invariant is then given by [10,53]

$$\nu = \tfrac{1}{2} \times [\text{sign changes of } P(\boldsymbol{k}) \text{ along } C] \bmod 2. \qquad (9)$$

Choosing C to be formed by a line integral over k_x and an irrelevant semicircular path around the upper half plane with $k \to \infty$, we can simply analyze $P(k_x, k_y = 0)$. This is plotted in Fig. 4 for 1–6 QLs and it is found that it remains negative for 1, 2, and 6 QLs, but changes sign twice for 3 to 5 QLs.

The results for the EMCNs determined above can now be explained in an intuitive manner. We refer to the picture of Fig. 2, which shows the case of 4 QLs, and focus on the valence bands. At the Γ point, the gap defined by the surface
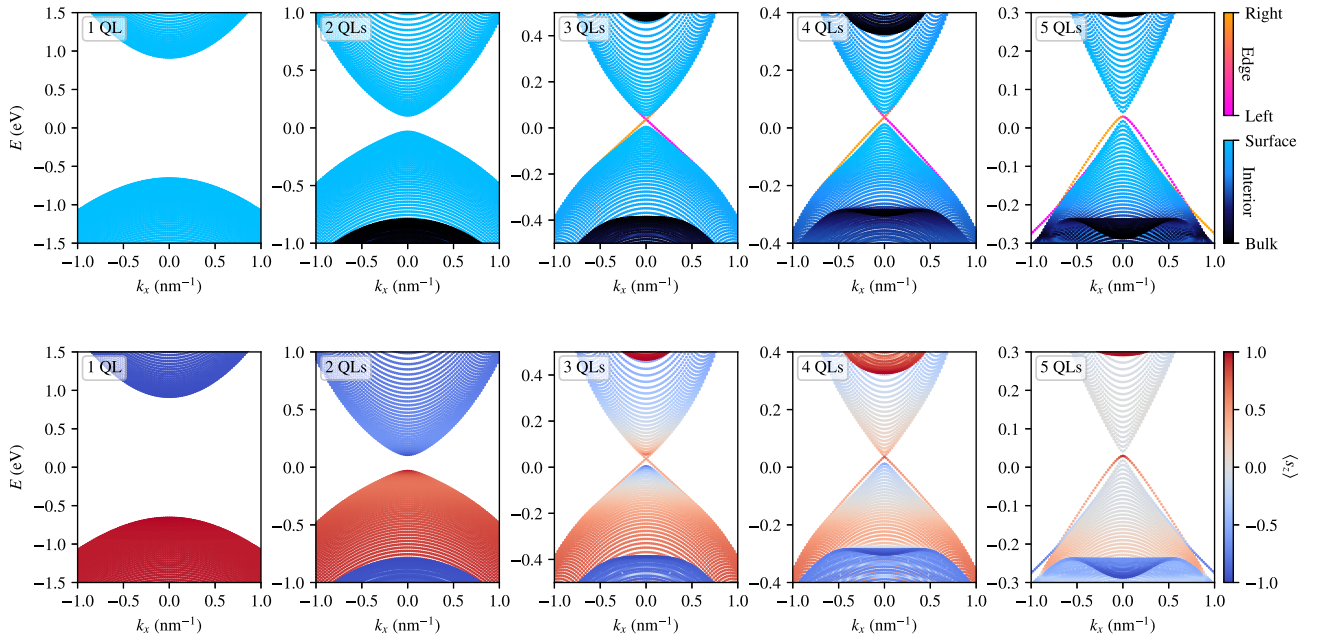
FIG. 3. Band structure of a ribbon of width $L_y = 100$ nm and 1 to 5 QLs in thickness for the subspace with SOP $\xi = +1$, as given by the eight-band model. All parameters used arise directly from the projection of the 3D states at the Γ point. The top row depicts the surface or bulk character of the bands for 1 to 5 QLs, that is, whether the corresponding states predominantly live in the outer or inner QLs, respectively. The edge states, only present between 3 and 5 QLs, are colored according to their localization on the ribbon. The bottom row depicts the spin polarization along the z direction for 1 to 5 QLs. Despite not being fully polarized, the edge states still have a nonzero average spin in the z direction, as explained in detail in Sec. II A. One must keep in mind that there is another subblock whose band structure is related by TRS to the one shown here. Hence the spins and velocities of electrons in the time-reversed subspace are opposite, giving the usual QSH picture but with a partial spin polarization.

states is inverted, as the spin reverses when moving away from it. Thus the UVB exchanges its character with the LCB. However, at larger momenta there is now an avoided crossing (AC) between the UVB and the LVB. This causes the spin of
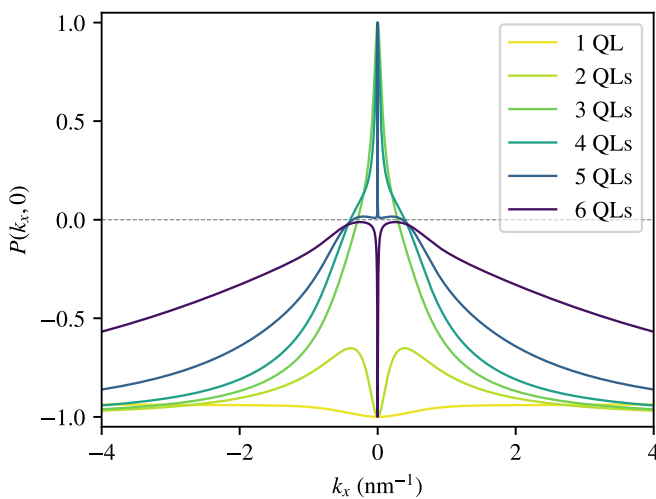

FIG. 4. Pfaffian of the matrix $A(\boldsymbol{k})$ for 1 to 6 QLs in the eight-band model. In the cases of 3, 4, and 5 QLs, the Pfaffian changes sign twice along the contour, giving $\nu = 1$, while the other cases are trivial. Note that our model has rotational symmetry and thus the Pfaffian's zeros, if any, form a circle enclosing the origin of the (k_x, k_y) plane.

the UVB to flip again, while also reversing that of the LVB. Thus the LVB ultimately acquires a nontrivial twist when going from $k = 0$ to $k \to \infty$, while the UVB remains trivial because its twist is undone at the AC. The same happens for 3 and 5 QLs. As a result, there is no net twist of the UVB and its EMCN is accordingly zero, while the net spin flip of the LVB results in a nontrivial EMCN. In contrast, for 1, 2, and 6 QLs, the closing of the gap undoes the band inversion at the origin. However, the AC remains, resulting in the spin twisting once for every band, but oppositely between the UVB and the LVB. The result is that the occupied bands have nonzero but opposite EMCNs, leading to a vanishing Hall conductivity, in agreement with our calculations.

In fact, this intuitive understanding becomes rigorous by virtue of the parity argument by Fu and Kane [51], because in addition to the TRS our $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian possesses the inversion symmetry of the underlying crystal lattice. In our continuum model there are only two time-reversal-invariant momenta, namely $k = 0$ and $k = \infty$, as the momentum space now has the topology of a sphere with all points at infinity identified. We find that the eigenstates of the effective Hamiltonian with a nonzero EMCN indeed change parity as one goes from the origin of $\boldsymbol{k}$ to infinity. This is ultimately the same as comparing the spin direction at these points, because the product of spin and parity is well defined in each subspace due to our definition of the SOP in Sec. II A.

We have now understood how the direction of the spin can be used to read off the topology of the bands in the effective model. However, one has to be careful regarding the

nature of the ACs involved in the process. While these are sources of Berry curvature and thus affect the EMCNs, they are actually an artifact of the effective model, which employs a finite number of bands. This can be seen by including even more bands: as we increase the size of the effective model, the ACs remain very localized but shift to larger values of k. In the limit of an infinite number of bands these will have disappeared and one recovers the smooth band structure arising from the solution to the 3D model at arbitrary $\boldsymbol{k}$. However, the band inversion of the LCB and the UVB close to Γ (at around $k_x \sim 0.4\,\text{nm}^{-1}$ in Fig. 2) remains localized in the same region no matter how many bands we include. One must therefore conclude that this inversion is really physical. Indeed, our $\boldsymbol{k} \cdot \boldsymbol{p}$ model is based on the premise that the topological properties of Bi₂Se₃ arise from the band inversion around the Γ point due to the spin-orbit coupling [30]. Thus the nontrivial Berry curvature should be localized around this point only, which is precisely expressed by the spin twists in this region.

In view of the above, the effective model is physical only in the region before the ACs take place. This means that the EMCNs do not actually correspond to the Chern numbers of the model with infinitely many bands, as they include the effect of the unphysical ACs. To remedy this, we introduce the PCNs. These must be calculated via the above spin-flip counting procedure, but only in the region before the first AC. Then only the physical band inversion contributes, if present. Note that, in any model with a finite band number, the PCNs cannot be calculated by integration of the Berry curvature as the result will not be quantized due to the truncation of the momentum space. The PCNs, as calculated via spin flips, are guaranteed to be the same as the Chern numbers of the infinite-band model. It then follows that for 3, 4, and 5 QLs the UVB has a unit PCN while that of the LVB is trivial, that is, the effective-model picture is essentially reversed. Furthermore, for 1, 2, and 6 QLs all bands have trivial PCNs. However, note that the physical Hall conductivity remains unchanged with respect to that of the effective model because it is given by a *sum* of Chern numbers. Thus it does not matter if it is calculated with the EMCNs or the PCNs, because each AC oppositely affects both involved bands. The $\mathbb{Z}_2$ invariant calculated via the Pfaffian is similarly correct in both cases, since it simultaneously includes the effect of all occupied bands.

We stress that the inclusion of the bulk states at the Γ point is essential to reach our conclusions regarding the PCNs. As said before, there is *no* band inversion around Γ in the effective four-band model for 4 and 5 QLs, i.e., all Chern numbers are trivial in this case. The eight-band model is thus the minimal model that contains all the topology, which can now be understood as follows. First, there is a band inversion between the LCB and the UVB that endows these bands with nontrivial PCNs. In the region where this inversion takes place, both of these bands are essentially fully surfacelike, as seen from Fig. 3. However, this inversion is only observed in the presence of the UCB and the LVB, which in this region are almost fully bulklike. We thus conclude that the nontrivial topology resides in states on the surfaces and is due to a band inversion enabled by the presence of deeper-lying states with bulk character.

TABLE I. Overview of the features of the four- and eight-band models discussed in the main text. Here, ν is the topological $\mathbb{Z}_2$ invariant and edge states are found only when it is nontrivial *and* there is a global energy gap, that is, all valence bands go to negative energies when k grows large.

	Thickness (QLs)	ν	Global gap	Edge states
4-band model	1	0	✓	✗
	2	0	✓	✗
	3	1	✓	✓
	4	0	✗	✗
	5	0	✗	✗
	6	1	✗	✗
8-band model	1	0	✓	✗
	2	0	✓	✗
	3	1	✓	✓
	4	1	✓	✓
	5	1	✓	✓
	6	0	✗	✗

Finally, in Table I we present a comparison summary between the four-band and the eight-band models. It is apparent that the four extra bands drastically modify the properties of the model for more than 3 QLs and must therefore be included. We note that for 6 QLs the UVB is not inverted, leading still to the absence of a global gap, as seen in Fig. 5. We have checked that this can be fixed by including yet another set of states, but this does not change the PCNs.
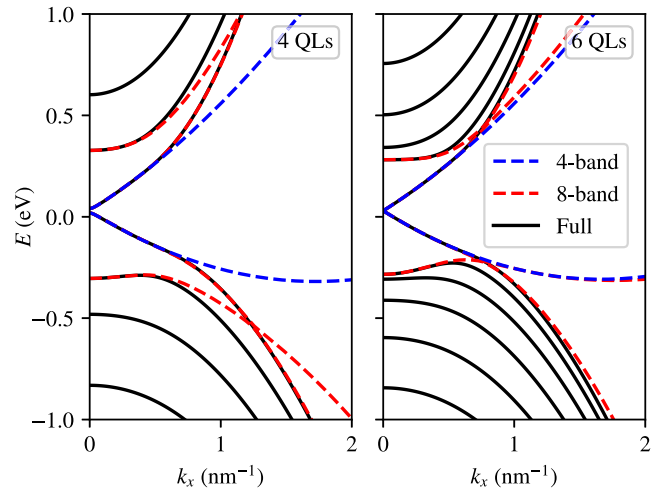

FIG. 5. Comparison of the four-band and eight-band models with the exact solution of the 3D model. The panels show a distinctive artifact of the four-band model, namely that the valence band is not inverted for 4 QLs and higher and thus goes up in energy for large k. In the case of 4 and 5 QLs, this is fixed by the eight-band model, whose valence bands go down in energy as desired. For 6 QLs and thicker, the eight-band model also presents this artifact. However, the topological invariant of this larger model remains the same as that of the eight-band model, which cannot be said when going from the four-band to the eight-band model.
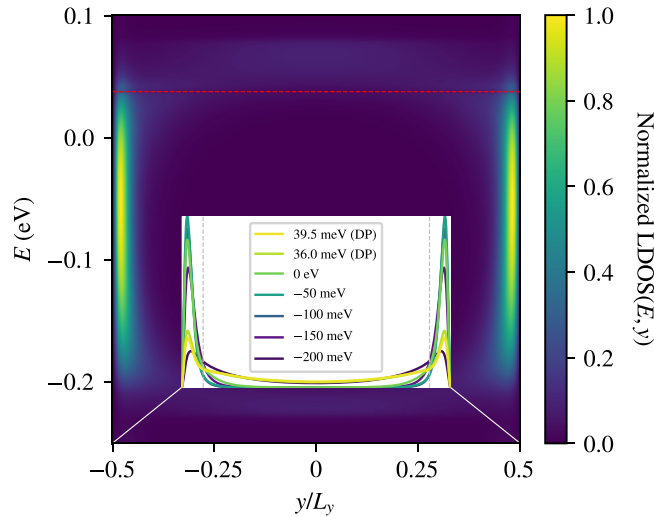
FIG. 6. Local density of states of the edge modes on a ribbon of width $L_y = 100\,\text{nm}$. The states span an energy range of roughly 200 meV and penetrate about 8 nm into the interior at both edges. The red dashed line marks the position of the Dirac point, where we find two states with a tiny gap. The inset shows slices of the density of states at various energies, with the ones corresponding to the Dirac point labeled as DP. The vertical dashed lines mark a distance of 8 nm from the edges, showing the spatial extent of the edge states that agrees with the experiment on colloidal nanosheets.

IV. COMPARISON WITH A RECENT EXPERIMENT

In a recent experiment [46], the presence or absence of QSH edge states was studied in colloidal nanocrystals of Bi_2Se_3 with thicknesses between 1 and 6 QLs. The findings clearly show that platelets of 1 to 3 QLs do not sustain edge modes, while platelets from 4 to 6 QLs do. A potential explanation for the 3 and 6 QL cases, which do not conform to our eight-band model, is given in the next section. Except for this issue, our eight-band model agrees with the experimental findings, while the four-band model does not if one simply takes the parameters arising from the projection of the 3D bulk Hamiltonian. In the experiment, the edge states are found to span a large energy window of around 500 meV and penetrate around 8 nm into the interior of the crystals. In Fig. 6 we have plotted the local density of edge states in the case of 4 QLs as calculated via the eight-band model. Their penetration is about 8 nm, in excellent agreement with the experimental value. Furthermore, they span an energy range of approximately 250 meV. While this is roughly only half of that observed experimentally, we note that the gap measured in the laboratory is also larger than that of our effective eight-band model [43]. Hence it is natural that the energy range spanned by the edge states is also larger than that predicted by the latter. Nevertheless, our eight-band model still provides an insight on the underlying mechanism behind this behavior. In Fig. 3 we see that the edge states merge with the conduction band very close to the Dirac point, but that they leak remarkably deep below the top of the valence band, especially for 4 and 5 QLs. This is attributed to a shift in the position of the Dirac point in combination with a renormalized Fermi velocity of the edge states, which around the Γ point is slightly lower

than that of the surface states. This enables the edge modes to live well separated from the gapped surfacelike bands for an energy range much larger than the zero-momentum gap. In fact, the same conclusion can be reached via the four-band model with experimentally adjusted parameters. For $D > 0$ (which is the case according to Ref. [43]) the Dirac point shifts upwards and the Fermi velocity in both cases becomes smaller than that of the surface states, as mentioned already after Eq. (4). As a result, the edge modes of the four-band model never actually touch the valence band; the eight-band model is needed to observe the merging that takes place in Fig. 3 and predict a finite energy range.

V. SUMMARY AND OUTLOOK

In this article we have investigated the effect of the nanosheet thickness on thin films of the 3D TI Bi_2Se_3. Whereas previous studies had already shown a dimensional transition from a 3D to a 2D TI, the theoretical description in terms of an underlying $\boldsymbol{k} \cdot \boldsymbol{p}$ Hamiltonian was unsatisfactory. Indeed, the emergent effective 2D model for the surface states can only describe the experimentally observed QSH phase at 4 and 5 QLs if its parameters are adjusted appropriately *a posteriori*. This amounts to disregarding the values that naturally arise from the projection in favor of a different set of parameters taken from experiment. In this work we have shown that this becomes unnecessary upon the inclusion of bulk states further away from the Fermi level. Our extended eight-band model precisely captures the topological properties computed via *ab initio* calculations in the whole range of 1 to 6 QLs and is close to the experimental findings except for 3 and 6 QLs. We find that the band inversion of the surfacelike bands, which determines the presence or absence of topology, can only be observed upon the inclusion of additional bulklike bands. This provides a more complete picture that is missing if one simply reparametrizes the surface bands of the four-band model. This observation constitutes one of the main points of this article.

Despite the very good agreement of our eight-band model with a recent experiment and previous DFT calculations, two brief comments are in place. First, we have found that the UVB is not inverted in the regime of intermediate thicknesses of 6 QLs and beyond. Hence a quantitative analysis of this regime would require projection onto a higher number of states. Note, however, that their inclusion does not further change the topology, i.e., the *topological* properties can be fully explained by the eight-band model. Secondly, our results for 3 and 6 QLs seem to clash with the recent experimental efforts involving Bi_2Se_3 nanocrystals in a way which indicates that both gap closings take place at a slightly larger thickness. It is important to note that in this article we have employed the 3D bulk parameters as fitted from *ab initio* calculations. To further test our eight-band model it would thus be desirable to directly input the parameters from experimental studies of bulk Bi_2Se_3. We expect that more precise knowledge of the bulk parameters will account for this discrepancy, as the values of L_z at which the gap closes are quite sensitive to the former.

Our model is in principle not restricted to Bi_2Se_3, but also applicable to Bi_2Te_3 and Sb_2Te_3 if we use the appropriate

parameter values. Experimental research in all of these materials would thus be a useful benchmark for our eight-band model as well. In the future we want to test our eight-band model in more general settings geared towards potential practical applications. One interesting possibility in this regard is heterostructures, in particular those involving a superconducting substrate which induces a proximity effect on the Bi$_2$Se$_3$ slab. Another intriguing avenue for which our combined surface-bulk Hamiltonian is especially well suited involves hybrid excitons, where an electron on the surface is coupled to a hole in the bulk, or vice versa. This system has been the subject of a recent experiment [54] and can be studied theoretically by combining our eight-band model with our approach to excitons in Ref. [55]. We hope that the considerations presented in this article will be useful for research in these and other fascinating areas of topological phases of matter.

ACKNOWLEDGMENTS

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

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

[3] J. E. Moore, The birth of topological insulators, Nature (London) **464**, 194 (2010).

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

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

[6] D. Kong and Y. Cui, Opportunities in chemistry and materials science for topological insulators and their nanostructures, Nat. Chem. **3**, 845 (2011).

[7] M. Z. Hasan and J. E. Moore, Three-dimensional topological insulators, Annu. Rev. Condens. Matter Phys. **2**, 55 (2011).

[8] D. Culcer, Transport in three-dimensional topological insulators: Theory and experiment, Physica E **44**, 860 (2012).

[9] D. Pesin and A. H. MacDonald, Spintronics and pseudospintronics in graphene and topological insulators, Nat. Mater. **11**, 409 (2012).

[10] B. A. Bernevig, *Topological Insulators and Topological Superconductors* (Princeton University Press, Princeton, NJ, 2013).

[11] Y. Ando, Topological insulator materials, J. Phys. Soc. Jpn. **82**, 102001 (2013).

[12] F. Ortmann, S. Roche, and S. O. Valenzuela, *Topological Insulators: Fundamentals and Perspectives* (John Wiley & Sons, New York, 2015).

[13] C.-K. Chiu, J. C. Y. Teo, A. P. Schnyder, and S. Ryu, Classification of topological quantum matter with symmetries, Rev. Mod. Phys. **88**, 035005 (2016).

[14] L. Kou, Y. Ma, Z. Sun, T. Heine, and C. Chen, Two-dimensional topological insulators: Progress and prospects, J. Phys. Chem. Lett. **8**, 1905 (2017).

[15] N. Xu, Y. Xu, and J. Zhu, Topological insulators for thermoelectrics, npj Quantum Mater. **2**, 51 (2017).

[16] K. v. Klitzing, G. Dorda, and M. Pepper, New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance, Phys. Rev. Lett. **45**, 494 (1980).

[17] 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 (1988).

[18] M. Stone, *Quantum Hall Effect* (World Scientific, Singapore, 1992).

[19] D. Yoshioka, *The Quantum Hall Effect* (Springer Science & Business Media, New York, 2002), Vol. 133.

[20] K. S. Novoselov, Z.-F. Jiang, Y.-S. Zhang, S. Morozov, H. L. Stormer, U. Zeitler, J. Maan, G. Boebinger, P. Kim, and A. K. Geim, Room-temperature quantum Hall effect in graphene, Science **315**, 1379 (2007).

[21] M. E. Cage, K. v. Klitzing, A. Chang, F. Duncan, M. Haldane, R. B. Laughlin, A. Pruisken, and D. Thouless, *The Quantum Hall Effect* (Springer Science & Business Media, New York, 2012).

[22] C. L. Kane and E. J. Mele, Quantum spin Hall effect in graphene, Phys. Rev. Lett. **95**, 226801 (2005).

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

[24] B. A. Bernevig and S.-C. Zhang, Quantum spin Hall effect, Phys. Rev. Lett. **96**, 106802 (2006).

[25] M. König, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Quantum spin Hall insulator state in HgTe quantum wells, Science **318**, 766 (2007).

[26] M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, The quantum spin Hall effect: theory and experiment, J. Phys. Soc. Jpn. **77**, 031007 (2008).

[27] X.-L. Qi and S.-C. Zhang, The quantum spin Hall effect and topological insulators, Phys. Today **63**, 33 (2010).

[28] J. Maciejko, T. L. Hughes, and S.-C. Zhang, The quantum spin Hall effect, Annu. Rev. Condens. Matter Phys. **2**, 31 (2011).

[29] X. Qian, J. Liu, L. Fu, and J. Li, Quantum spin Hall effect in two-dimensional transition metal dichalcogenides, Science **346**, 1344 (2014).

[30] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Topological insulators in Bi$_2$Se$_3$, Bi$_2$Te$_3$ and Sb$_2$Te$_3$ with a single Dirac cone on the surface, Nat. Phys. **5**, 438 (2009).

[31] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava *et al.*, Observation of a large-gap topological-insulator class with a single Dirac cone on the surface, Nat. Phys. **5**, 398 (2009).

[32] D. Hsieh, Y. Xia, D. Qian, L. Wray, J. Dil, F. Meier, J. Osterwalder, L. Patthey, J. Checkelsky, N. P. Ong *et al.*, A tunable topological insulator in the spin helical Dirac transport regime, Nature (London) **460**, 1101 (2009).

[33] W. Zhang, R. Yu, H.-J. Zhang, X. Dai, and Z. Fang, First-principles studies of the three-dimensional strong topological insulators Bi$_2$Te$_3$, Bi$_2$Se$_3$ and Sb$_2$Te$_3$, New J. Phys. **12**, 065013 (2010).

[34] F. Manjón, R. Vilaplana, O. Gomis, E. Pérez-González, D. Santamaría-Pérez, V. Marín-Borrás, A. Segura, J. González, P. Rodríguez-Hernández, A. Muñoz *et al.*, High-pressure studies of topological insulators Bi$_2$Se$_3$, Bi$_2$Te$_3$ and Sb$_2$Te$_3$, Phys. Status Solidi B **250**, 669 (2013).

[35] K. Mazumder and P. M. Shirage, A brief review of Bi$_2$Se$_3$ based topological insulator: from fundamentals to applications, J. Alloys Compd. **888**, 161492 (2021).

[36] C.-X. Liu, X.-L. Qi, H.-J. Zhang, X. Dai, Z. Fang, and S.-C. Zhang, Model Hamiltonian for topological insulators, Phys. Rev. B **82**, 045122 (2010).

[37] J. Zhang, Z. Peng, A. Soni, Y. Zhao, Y. Xiong, B. Peng, J. Wang, M. S. Dresselhaus, and Q. Xiong, Raman spectroscopy of few-quintuple layer topological insulator Bi$_2$Se$_3$ nanoplatelets, Nano Lett. **11**, 2407 (2011).

[38] H. Cao, S. Xu, I. Miotkowski, J. Tian, D. Pandey, M. Z. Hasan, and Y. P. Chen, Structural and electronic properties of highly doped topological insulator Bi$_2$Se$_3$ crystals, Phys. Status Solidi RRL **7**, 133 (2013).

[39] C.-X. Liu, H.-J. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, Oscillatory crossover from two-dimensional to three-dimensional topological insulators, Phys. Rev. B **81**, 041307(R) (2010).

[40] H.-Z. Lu, W.-Y. Shan, W. Yao, Q. Niu, and S.-Q. Shen, Massive Dirac fermions and spin physics in an ultrathin film of topological insulator, Phys. Rev. B **81**, 115407 (2010).

[41] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, Effective continuous model for surface states and thin films of three-dimensional topological insulators, New J. Phys. **12**, 043048 (2010).

[42] J. Linder, T. Yokoyama, and A. Sudbø, Anomalous finite size effects on surface states in the topological insulator Bi$_2$Se$_3$, Phys. Rev. B **80**, 205401 (2009).

[43] Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, X. Dai, W.-Y. Shan *et al.*, Crossover of the three-dimensional topological insulator Bi$_2$Se$_3$ to the two-dimensional limit, Nat. Phys. **6**, 584 (2010).

[44] Y. Sakamoto, T. Hirahara, H. Miyazaki, S.-I. Kimura, and S. Hasegawa, Spectroscopic evidence of a topological quantum phase transition in ultrathin Bi$_2$Se$_3$ films, Phys. Rev. B **81**, 165432 (2010).

[45] M. Neupane, A. Richardella, J. Sánchez-Barriga, S. Xu, N. Alidoust, I. Belopolski, C. Liu, G. Bian, D. Zhang, D. Marchenko *et al.*, Observation of quantum-tunnelling-modulated spin texture in ultrathin topological insulator Bi$_2$Se$_3$ films, Nat. Commun. **5**, 3841 (2014).

[46] J. Moes, J. F. Vliem, P. M. Melo, T. C. Wigmans, R. G. Mendes, E. van Brenk, I. Swart, L. Maisel Licerán, H. T. C. Stoof, C. Delerue, Z. Zanolli, and D. Vanmaekelbergh, Characterization of the edge states in colloidal Bi$_2$Se$_3$ platelets, Nano Lett. **24**, 5110 (2024).

[47] H. Okamoto, The Bi-Se (Bismuth-Selenium) system, J. Phase Equilib. **15**, 195 (1994).

[48] H. Lind, S. Lidin, and U. Häussermann, Structure and bonding properties of (Bi$_2$Se$_3$)$_m$(Bi$_2$)$_n$ stacks by first-principles density functional theory, Phys. Rev. B **72**, 184101 (2005).

[49] H.-H. Kung, A. Goyal, D. Maslov, X. Wang, A. Lee, A. Kemper, S.-W. Cheong, and G. Blumberg, Observation of chiral surface excitons in a topological insulator Bi$_2$Se$_3$, Proc. Natl. Acad. Sci. USA **116**, 4006 (2019).

[50] The states $|S^+, \uparrow (\downarrow)\rangle$ and $|S^-, \uparrow (\downarrow)\rangle$ correspond to $|\chi^{\uparrow(\downarrow)}\rangle$ and $|\varphi^{\uparrow(\downarrow)}\rangle$ in Refs. [40,41], respectively.

[51] L. Fu and C. L. Kane, Topological insulators with inversion symmetry, Phys. Rev. B **76**, 045302 (2007).

[52] H. Li, L. Sheng, D. N. Sheng, and D. Y. Xing, Chern number of thin films of the topological insulator Bi$_2$Se$_3$, Phys. Rev. B **82**, 165104 (2010).

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

[54] R. Mori, S. Ciocys, K. Takasan, P. Ai, K. Currier, T. Morimoto, J. E. Moore, and A. Lanzara, Spin-polarized spatially indirect excitons in a topological insulator, Nature (London) **614**, 249 (2023).

[55] L. Maisel Licerán, F. García Flórez, L. D. Siebbeles, and H. T. Stoof, Single-particle properties of topological Wannier excitons in bismuth chalcogenide nanosheets, Sci. Rep. **13**, 6337 (2023).