First-principles modeling of binary layered topological insulators: Structural optimization and exchange-correlation functionals

Thomas K. Reid¹, S. Pamir Alpay^{1,2} Alexander V. Balatsky^{1,3} and Sanjeev K. Nayak^{1,*}

¹Department of Materials Science and Engineering and Institute of Materials Science, University of Connecticut, Storrs, Connecticut 06269, USA

²Department of Physics, University of Connecticut, Storrs, Connecticut 06269, USA

³NORDITA, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, Stockholm SE-106 91, Sweden

(Received 23 July 2019; revised manuscript received 12 December 2019; accepted 11 February 2020; published 28 February 2020)

Topological insulators (TIs) are materials that are insulating in the bulk but have zero band-gap surface states with linear dispersion and are protected by time-reversal symmetry. These unique characteristics could pave the way for many promising applications that include spintronic devices and quantum computations. It is important to understand and theoretically describe TIs as accurately as possible to predict properties. Quantum mechanical approaches, specifically first-principles density-functional-theory (DFT)-based methods, have been used extensively to model electronic properties of TIs. Here we provide a comprehensive assessment of a variety of DFT formalisms and how these capture the electronic structure of TIs. We concentrate on Bi₂Se₃ and Bi₂Te₃ as examples of prototypical TI materials. We find that the generalized gradient approximation (GGA) and kinetic density functional (metaGGA) increase the thickness of the TI slab, whereas we see the opposite behavior in DFT computations using LDA. Accounting for van der Waals (vdW) interactions overcomes the apparent over-relaxations and retraces the atomic positions toward the bulk. Based on a systematic computational study, we show that GGA with vdW treatment is an appropriate method for structural optimization. However, the vdW corrections recover the experimental bulk parameters, and do not influence the charge density implicitly. Thus, electronic structures derived from the base GGA functional, employing experimental lattice parameters, is sufficient.

DOI: 10.1103/PhysRevB.101.085140

I. INTRODUCTION

Within the last decade, it was reported that the pnictogen chalcogenides Bi_2Se_3 and Bi_2Te_3 exhibit the properties of a topological insulator (TI)—small energy gap at the Fermi level, inverted parity of the band edge states leading to a Dirac band dispersion, and quantum oscillations for topological surface states (TSSs) [1–8]. This discovery has resulted in significant interest in these two materials, as their relatively large bulk band gaps and chemical simplicity make the study of their topological electronic physics readily accessible to theory and experimental work alike. TIs are projected to be the basis of many cutting-edge device applications, including spintronics and quantum computers [9–11]. TIs also offer fertile ground for fundamental studies of exotic electronic phenomena arising from wave-function topology [4,12–20].

Structurally, Bi₂Se₃ and Bi₂Te₃ are most often described in terms of quintuple layers (QLs), in which atomic layers are arranged in sets of five sublayers along the out-of-plane *c* axis of Bi₂Se₃ and Bi₂Te₃, a configuration called the quintuple layer (QL). The arrangement of atoms in a QL is B(1)-A-B(2)-A-B(1) (for A₂B₃) along the [0001] direction of the hexagonal crystal system with space group $R\overline{3}m$ (No. 166) [21]. The layers in a QL are bounded by stronger chemical bonding but the inter-QL interaction is comparatively weaker and of van der Waals (vdW) type.

The number of QLs (nQL) needed to achieve a robust TSS is of primary interest for the accurate description of the TI behavior. Different factors such as quantum effects due to finite size, undercoordinated surface atoms, and hybridization of orbitals from the surface and bulk atoms also contribute to the emergence or suppression of a TSS. Zhang et al. [6] reported the appearance of a Dirac point in samples with thicknesses 6QLs and greater for Bi₂Se₃, while, for Bi₂Te₃, Liu et al. [22] have reported that four quintuple layers (4QLs) are sufficient. Using a combined model Hamiltonian study and first-principles calculations, Liu et al. [23] reported an oscillatory crossover from a two-dimensional to a three-dimensional TI as a function of film thickness (i.e., nQL). Since both intra-QL and inter-QL interactions are crucial to stabilize the TSS [24,25], it is important to standardize the settings of density-functional theory (DFT) modeling to converge toward a common perspective as applied to this general class of materials.

DFT is one of the most popular theoretical tools to study structural and electronic properties of materials. It relies upon the Kohn-Sham formulation, where the knowledge of electron density is sufficient to calculate several materials properties that depend on the electronic and atomic structure. It is an

^{*}sanjeev.nayak@uconn.edu

efficient means of establishing the ground state, even for relatively large systems. Furthermore, because it is a firstprinciples method, it allows an unambiguous comparison to model experiments. It is important to note that DFT's robustness is not based only in its capacity for verification and explanation of experimental results. It also allows for predicting properties for previously unexplored systems to guide experimental work. While the validity of principles of DFT is established, the deficiencies of exchange-correlation functionals (XCFs) may limit its predictive capabilities. A substantial amount of knowledge has been accumulated to quantify these shortcomings, and ways to circumvent them have been suggested with mixed success of transferability [26–37].

For TI materials like Bi_2Se_3 and Bi_2Te_3 , DFT can be utilized in two major ways: The first is the description of the ground-state atomic positions (relaxation). The second is the description of the converged charge density. Both play a significant role in quantifying surface states (SSs) and the selection of the appropriate XCFs. TI materials have been studied with XCFs including the local density approximation (LDA) and generalized gradient approximation (GGA) [1,2,6,38– 45], but the effect of the kinetic density functional (metaGGA) has not been thoroughly investigated. Together, these three XCFs, in order, are the first three "rungs" on the "ladder" of chemical accuracy. The many-body perturbation theory in the GW approximation, where G is the single particle Green's function and W is the screened Coulomb interaction, has been applied to improve description of the band structure [46–48]. It is also important to realize that atomic structure is intricately related to electronic properties. For example, the GW approximations are often applied to experimental or DFTderived atomic structures. On the other hand, for layered TI materials like Bi₂Se₃ and Bi₂Te₃, treatment of the vdW forces has proven particularly important, and many studies have demonstrated that doing so greatly affects DFT's description of electronic structures. This is because, for a slab or film of a fixed nOL, a relationship exists between the thickness of the slab or film (i.e., due to strain) and the possibility of observing a TSS. VdW forces mediate the interactions between QLs across the inter-QL space-depending on the magnitude of their influence, they create smaller or larger vdW gaps. As a result, the vdW-corrected DFT might predict a different overall thickness and, thus, it might predict a different type of SS [49-53]. However, it must be noted that the vast majority of the thickness of a slab is made up of intra-OL separations. To satisfactorily account for a given XCF's contribution to change in thickness relative to the experimental parameters and those produced by other XCFs, we must assess the role played by the intra-QL space in overall thickness as well, which has not been taken into account by previous studies.



FIG. 1. Schematic pictures of bulk crystal lattice of Bi₂Se₃ and Bi₂Te₃ in hexagonal system with space group $R\overline{3}m$ (166) in (a) and (b). Slab models constructed from the bulk lattice with varied quintuple layers are shown in (a).

In this paper, we carry out a systematic analysis in pursuit of a consensus regarding the role of XCFs in the outcome of structural optimization and its subsequent influence over electronic structure calculations in Bi₂Se₃ and Bi₂Te₃. We specifically asses a variety of XCFs including: LDA, GGA, metaGGA, GGA + vdW, and metaGGA + vdW. We find that the structural optimization is highly sensitive to the type of XCF employed. LDA and GGA produce relaxation trends opposite of each other. The effective role of the vdW treatment is to bring the relaxation toward the experimental bulk positions. While the outcomes of the electronic-structure calculations are dependent on the specific XCF, we used the vdW treatments of Grimme et al. [54], which do not depend on charge density, making it redundant to apply them to the self-consistent charge density (SCCD) and band-structure phases of our calculations. As a result, our vdW-treated XCFs preserve the charge density distribution generated by their root functionals. It is important to note that our structural optimizations with the vdW treatments indicate little deviation (below 2%) from experimental bulk coordinates of Nakajima (Ref. [21]). Therefore, we deduce that the experimental bulk coordinates of Nakajima are sufficient to create structural models of Bi₂Se₃ and Bi₂Te₃ in which SSs that exhibit thickness dependence in agreement with experiment can be expected, so long as the GGA or metaGGA XCFs (the root functionals of our vdW-corrected XCFs) are used. Lastly, mindful of the range of XCFs, nQL (slab thicknesses), and computational strategies pursued in the theoretical literature on SSs of Bi₂Se₃ and Bi₂Te₃, we extend these conclusions to a prescription for a general method for generating models suitable for studying TI physics in binary layered pnictogen chalcogenides, with DFT.

In this way, we seek to evaluate with precision the outcome of structural optimizations with a variety of XCFs and benchmark a procedure for reliably modeling electronic structure in these systems. We expect this study should act as the groundwork for the production of models that might be used to generate high quality data that would serve the development of novel devices.

II. METHODOLOGY

Calculations were performed with DFT using the VI-ENNA AB INITIO SIMULATION PACKAGE (VASP) [55,56] which solves the scalar-relativistic Kohn-Sham Hamiltonian. The spin-orbit coupling is added to the DFT Hamiltonian through the term $H_{\text{SOC}}^{\alpha\beta} \propto \boldsymbol{\sigma} \cdot \boldsymbol{L}$, where α and β are the majority and minority spin components, σ is the Pauli-spin operator, and L is the angular momentum operator [57]. We used the following three root XCFs: LDA with Ceperley-Alder parametrization [58], GGA with Perdew-Burke-Ernzerhof parametrization (PBE) [59], and metaGGA with strongly constrained and appropriately normed (SCAN) approach [60]. Slab models were prepared from the unit bulk hexagonal lattice $(R\overline{3}m)$ with lattice parameters taken from experiments [21]. This set of experimental lattice parameters matches the DFT band structure as reported in Ref. [50]. The schematic picture of the bulk hexagonal unit cell and the slab models are shown in Figs. 1(a)-1(c). The slabs were centered along the vertical direction. A vacuum layer of thickness of ~8.8 Å was added to

TABLE I. Table showing the lattice contants a and c, Wyckoff reference of Bi, Se(1), and Se(2) positions, and optical band gap of bulk Bi₂Se₃ obtained from various treatments of XCFs in DFT and the corresponding experimental values.

a (Å)	c (Å)	z-Bi	z-Se(1)	z-Se(2)	E _g (eV)
4.108 ^a	27.293 ^a	0.4012 ^a	0.0 ^a	0.2093 ^a	0.042 (-0.326) ^a
4.188 ^b	31.429 ^b	0.3953 ^b	0.0^{b}	0.2216 ^b	0.691 (-0.302) ^b
4.174 ^c	28.872 ^c	0.4001 ^c	0.0 ^c	0.2125 ^c	$0.234 (-0.198)^{c}$
4.157 ^d	29.812 ^d	0.3983 ^d	0.0^{d}	0.2160 ^d	$0.353 (-0.058)^{d}$
4.140 ^e	28.657 ^e	0.4004 ^e	0.0 ^e	0.2121 ^e	0.115 (-0.278) ^e
4.143 ^f	28.636 ^f	0.4008 ^f	0.0^{f}	0.2117 ^f	$-0.220^{\text{g}}, -0.335^{\text{h}}$

^aLDA (+SOC), present calculation.

^bPBE (+SOC), present calculation.

^cPBE+vdW (+SOC), present calculation.

^dSCAN (+SOC), present calculation.

eSCAN+vdW (+SOC), present calculation.

^fExperiment, Ref. [21]

^gExperiment, Ref. [64].

^hExperiment, Ref. [65].

either side of the surfaces making a total vacuum thickness of \sim 17.6 Å. This range of vacuum thickness ensures negligible dipole interaction between the surfaces through the vacuum and is regularly used for modeling surfaces and nanomaterials [61–63]. A systematic study from one quintuple-layer (1QL) to eight quintuple-layers (8QLs) models were performed. The calculations were performed using the projector-augmented wave pseudopotentials, a symmetrized $7 \times 7 \times 1$ k-point grid to span the two-dimensional Brillouin zone, the kinetic energy cutoff of the plane waves set to 500 eV, and the precision tag set to "accurate," which generated 162 000, 225 000, 288 000, 352 800, 432 000, 486 000, 540 000, and 604 800 number of plane waves for 1QL, 2QL, 3QL, 4QL, 5QL, 6QL, 7QL, and 8QL models, respectively. An additional support grid was included in the calculation through the tag addgrid, which uses an augmentation grid twice the size of the regular grid for the representation of the pseudo wave functions, accounting for

TABLE II. Table showing the lattice contants a and c, Wyckoff reference of Bi, Te(1), and Te(2) positions, and optical band gap of bulk Bi₂Te₃ obtained from various treatments of XCFs in DFT and the corresponding experimental values.

a (Å)	c (Å)	z-Bi	<i>z</i> -Te(1)	z-Te(2)	E _g (eV)
4.356 ^a	29.903 ^a	0.4016 ^a	0.0 ^a	0.2089 ^a	0.138 (-0.130) ^a
4.445 ^b	32.247 ^b	0.3970 ^b	0.0^{b}	0.2163 ^b	0.555 (-0.031) ^b
4.432 ^c	30.446 ^c	0.3998 ^c	0.0 ^c	0.2105 ^c	$0.198 (-0.244)^{c}$
4.411 ^d	31.474 ^d	0.3981 ^d	0.0^{d}	0.2139 ^d	$0.391 (-0.079)^{d}$
4.400 ^e	30.096 ^e	0.4001 ^e	0.0 ^e	0.2096 ^e	$0.179(-0.169)^{e}$
4.386 ^f	30.497 ^f	0.4000 ^f	0.0^{f}	0.2097 ^f	-0.171^{g}

^aLDA (+SOC), present calculation.

^bPBE (+SOC), present calculation.

^cPBE+vdW (+SOC), present calculation.

^dSCAN (+SOC), present calculation.

^eSCAN+vdW (+SOC), present calculation.

^fExperiment, Ref. [21].

^gExperiment, Ref. [66].



FIG. 2. Change in slab thickness expressed in percent with respect to bulk values obtained after geometrical optimization using different XCFs for (a) Bi_2Se_3 and (b) Bi_2Te_3 .

accurate interatomic forces and hence leads to higher quality geometrical optimizations. The total energy convergence for self-consistent field cycles (SCF) is set to 10^{-7} eV and the force convergence for geometrical optimization is set to 0.001 eV/Å.

Each QL of a multi-QL system is weakly bounded to the others by the vdW force. Past works have demonstrated the importance of accounting for the influence of the vdW forces in layered systems like Bi_2Se_3 and Bi_2Te_3 in DFT calculations by using specially adapted treatments [49,50]. We



FIG. 3. Relaxation of layers in 3QL and 6QL slabs for various treatment of XCFs for Bi_2Se_3 [(a), (b)] and Bi_2Te_3 [(c), (d)]. Left panel is the difference between relaxed and bulk atomic positions Δz . The zero value of layer index (abscissa) is midway between the slab thickness (see Fig. 1). The right panel is the fraction of inter-QL separation (orange color) and intra-QL thickness (turquoise color) contribution to the total thickness. The fraction of inter-QL separation in percentage is displayed in numbers.

incorporated such treatments into another set of calculations with the GGA functional using the zero-damping DFT-D3 method of Grimme *et al.* [54]. We wanted to cross-check the outputs including the vdW interactions, so a set of calculations using the SCAN + vdW treatment was carried out.

III. STRUCTURAL PROPERTIES

The experimental bulk lattice parameters and corresponding results from the DFT calculations for Bi_2Se_3 and Bi_2Te_3 are shown in Tables I and II, respectively. Our calculations are delineated by rows by XCF. The data shows that SCAN + vdW predictions of lattice parameters most closely reflect experimental results, followed in this regard by PBE + vdW. LDA and PBE XCFs deviate more from the experimental values, with LDA predicting structural parameters lower than experiment and PBE predicting structural parameters higher than experiment. Relaxation along *a* resulted in far less change overall compared with relaxations along *c*. The band gaps also vary with the choice of functional (see Sec. IV for further discussions about band gap).

The slab models generated from the bulk experimental coordinates of Ref. [21] show variation in the direction and extent of relaxation with different XCF treatments. The relative change of thickness of the slab is $\Delta t/t = (t_{\text{XCF}} - t_{\text{expt.}})/t_{\text{expt.}}$ where $t_{\rm XCF}$ and $t_{\rm expt.}$ are the optimized thickness and ideal bulk-cut thickness, respectively. This value is presented as a function of nQL, and of XCF, for Bi2Se3 and Bi2Te3, in Figs. 2(a) and 2(b), respectively. The horizontal line passing through zero for $\Delta t/t$ represents the thickness of the structure with experimental lattice parameters. For all calculations conducted with the LDA functional, a negative value of $\Delta t/t$ is observed. Naturally, this implies the LDA functional induces a reduction in the thickness of all slabs. An opposite trend is observed for the PBE functional, implying that PBE predicts the thickness of any slab should increase. The SCAN functional, for both materials, predicts a less severe increase of the slab thickness than PBE. When we apply the vdW treatments to both PBE and SCAN, both functionals predict that the thickness of the slab should be bulk-like.

While it is obvious each functional affects the general thickness of the slab, a focus on the thickness alone obscures the control each XCF exerts over the role of more specific mechanisms in the outcome of a given structural optimization. In the left panels of Fig. 3, we illustrate the change in the position of each atomic layer along c relative to its bulk index position for 3QL and 6QL models for Bi₂Se₃ [Figs. 3(a) and 3(b)] and Bi₂Te₃ [Figs. 3(c) and 3(d)], respectively, taking into account the five discussed XCFs. $\Delta z \ (= z_{\text{relaxed}} - z_{\text{initial}})$ is the amount of change undergone by an atomic layer in comparison to its position in the initial model generated from the experimental bulk lattice parameters. Since the zeroth index in the left panels of Figs. 3(a)-3(d) is set to the center layer of the slab for models with an odd nQL and one of the two center layers for models with an even nQL, an atomic layer's position relative to the horizontal at zero is indicative of the direction of its relaxation. Presented in such a way, the positive value of Δz for the negative layer indices and negative value of Δz for the positive layer indices imply inward relaxation of the two ends of the slab. Similarly, the



FIG. 4. Ratio of the separation of the two topmost atomic layers (d_{OAL}) and the separation of the two topmost QLs (d_{iQL}) optimized with different XCFs for 2–8QL models to their respective ultrathin film experimental values. In both panels, the horizontal and vertical dashed lines intersect at the experimental value (1.0), and the diagonal dashed lines represent a projection to slabs of variable thickness of the ratio obtained from experiments. The experimental values are taken from Ref. [67] for Bi₂Se₃ and Ref. [68] for Bi₂Te₃.

negative value of Δz for the negative layer indices and positive value of Δz for the positive layer indices imply outward relaxation. The trend of the change in the layer positions within a QL-identifiable as the clusters of five data points demarcated by the grey boxes-is generally linear, inward or outward, for all models. A step between each QL represents the vdW gap (not to be confused with the vdW treatments for XCFs) [69]. PBE predicts the largest outward change in the size of the vdW gap relative to the experimental observations. PBE + vdW and SCAN + vdW, on the other hand, predict a far smaller change in the vdW gap, slightly inward or outward depending on the system and the thickness of the slab [refer to the left panel of Figs. 3(a)-3(d)]. The remaining functionals predict absolute changes in the vdW gap larger than the vdWcorrected functionals, but smaller than pure PBE. Comparing to the bulk, which is the horizontal line passing through $\Delta z = 0$ in the left panels of Figs. 3(a)-3(d), it is observed



FIG. 5. A comparison of charge density obtained from SCF calculations of various XCFs using a 6QL slab model with atomic positions of bulk experimental coordinates. The first QL (surface QL) and third QL (inner QL) of the slab model is shown for Bi_2Se_3 in (a) and (b), respectively, and similarly for Bi_2Te_3 in (c) and (d). The intraatomic positions of the QLs are marked with horizontal dashed lines. There are notable differences on the charge density distribution for different XCFs and vdW-corrected DFT preserving the charge density of the corresponding root functional (PBE or SCAN).

that SCAN + vdW and PBE + vdW deviate minimally from the corresponding bulk atomic positions. Roy *et al.* [67,70] performed surface x-ray diffraction studies on (0001) Bi₂Se₃ grown as an ultrathin film using molecular beam epitaxy on Si(111). They found outward relaxation of the top Se-Bi layer by $\sim 2\% - 4\%$ as compared to the bulk coordinates and $\sim 3\%$ contraction of the top vdW gap (d_{iQL}) compared to the same. Our results suggest that the PBE + vdW relaxation matches more closely to these experimental results for the outer Se-Bi layers. However, the contraction of d_{iQL} is not represented in our calculations.

It is clear from the right panels of Figs. 3(a)–3(d) that the majority of the thickness of each relaxed slab is occupied by the intra-QL space and a minority by the inter-QL space. That the thickness of the vdW-corrected functionals maintains approximately the same ratio of intra-QL space to inter-QL space as the other functionals, but exhibits far lower thickness in general, suggests two possibilities: (1) By merit of the large absolute differences between the thicknesses of the non-vdW-corrected and vdW-corrected models, the vdW force is the primary mediating force during structural optimization. (2) The vdW force is involved in the control of the inter-QL distances as well as the intra-QL distances.

For LDA, PBE, and SCAN models, even when they are as large as 8QL, convergence to the experimental bulk charac-

teristics is not obtained. On the other hand, it should again be noted that the relaxations of the vdW-corrected models show strong convergence to the corresponding experimental bulk for any nQL. This suggests that slabs prepared from the experimental coordinates, and not subject to structural optimization, adequately capture the physics accounted for by the vdW treatments, and are reasonable for studying the electronic structure. But a conclusion cannot be drawn without evaluating the electronic structure directly. This is done in Sec. IV.

In Fig. 4, we plot the relationship between the ratio of the optimized d_{iQL} and separation of the two topmost atomic layers d_{OAL} , to their respective ultrathin film experimental values, for Bi₂Se₃ and Bi₂Te₃. The experimental values are taken from Ref. [67] for Bi₂Se₃ and Ref. [68] for Bi₂Te₃. In Figs. 4(a) and 4(b), the horizontal and vertical dashed lines represent the experimental value, which intersect at (1.000, 1.00), and the diagonal dashed lines represent an extrapolation to slabs of variable thickness of the ratio of the d_{iQL} and d_{OAL} in the experimental films. The advantage of this set of plots is that the contributions of the changes in the d_{iQL} and the d_{OAL} to the overall change in thickness after relaxation can be accounted for explicitly. For Bi₂Se₃, PBE induces a large change in the d_{iQL} . Applying the vdW treatment greatly reduces the predicted growth in that variable, while also

inducing a comparable reduction in the d_{OAL} . By contrast, for Bi₂Te₃, PBE greatly increases both d_{OAL} and d_{iQL} relative to bulk, and such that the ratio of the two deviates strongly from the experiment (the dashed diagonal line). Applying the vdW treatment leads to a convergence of the data to the experimental ratio, yet not to the experimental values reported in Ref. [68]. This again implies that the vdW treatment controls the overall thickness by changing both the d_{iQL} and d_{OAL} . For both of the materials, a similar trend can be seen with the SCAN functional and its vdW-treated counterpart. For both materials, LDA predicts a smaller d_{iQL} and d_{OAL} .

IV. ELECTRONIC PROPERTIES

In this section, we seek to compare the selection of the XCF to the relationships between sample thickness and the TI properties observed in experiment using the optimized structures generated in the previous section.

That density-based XCFs underestimate the band gap is generally accepted for wide-band-gap systems, and this is true here, as well [71]. The band gaps for bulk Bi_2Se_3 and Bi_2Te_3 obtained from various treatment of XCFs are summarized in Tables I and II, respectively. The negative band gap signifies an inverted band occupancy between the top of valence band (VB) and bottom of conduction band (CB) compared to a regular semiconductor, where the VB is composed of anion states and CB composed of cation states [72,73]. The band structures of Bi2Se3 and Bi2Te3 for each functional (without and with SOC) are provided in Fig. 10. The SOC is the main interaction that results in negative band gap in the bulk electronic structure. The inclusion of SOC affects the band structures in two salient ways. First, it leads to mixing of majority and minority spin states, which reduces the band gap to the extent that the band inversion appears, given a forbidden gap exists. The second is the shift of VB and CB extrema along k, imparting indirect character to the band structure. The electronic states of TI materials need SOC treatment for a comparison to experimental results. Hereon, our discussion on bandgap only includes SOC.

We find an overestimation of bulk band-gap magnitude as compared to experiments for both PBE and SCAN functionals. In particular, the SCAN performs the worst for band gap prediction [74]. This is more prominent in Bi_2Te_3 . The limitations of DFT in expressing the electronic band structure is partially overcome by the many-body perturbation theory, i.e., the GW method. Förster et al. have shown that the character of the bands at the Fermi level at the Γ -point in Bi₂Se₃ can be changed from indirect to direct by applying the GW treatment over LDA-based DFT [46,47]. By comparison, although a noticeable change in the band energies appears in Bi₂Te₃, there is no qualititative shift in the band character toward a direct band gap. The different response of Bi₂Te₃ to GW can be plausibly attributed to the greater mass of tellurium. A recent study has highlighted the use of a hybridfunctional (B3PW91) to deliver band structures matching those of G_0W_0 calculations [75]. The key point in applications of these advanced treatments is to recover the direct band gap at the Γ -point for Bi₂Se₃. Similar results are also accounted for in metaGGA functional, with the shortcoming that the magnitude of the band gap is still underestimated. From



FIG. 6. Optical band gap as a function of nQL for Bi₂Se₃ and Bi₂Te₃ shown in (a) and (b), respectively. Some data is compared to literature with (a), (b), and (c) for data sources referring to Ref. [38], Ref. [6], and Ref. [22], respectively.

our study, we can conclude that vdW corrections to DFT functionals are still important for accurate prediction of bulk lattice parameters and electronic structure.

A 6QL slab model generated from experimental bulk lattice parameters is chosen to show the xy-integrated charge density as a function of z. Further, the charge density is compared with the LDA-obtained charge density as the reference by the construct $(\rho_z - \rho_z(LDA))/\rho_z(LDA)$. The comparison of charge density is done for the outer QL and one of the inner QLs (third QL from the outer QL) of Bi₂Se₃ in Figs. 5(a) and 5(b), respectively, and similarly for Bi_2Te_3 in Figs. 5(c) and 5(d). The results demonstrate that the charge densities of PBE and SCAN show different trends as compared to LDA, and do not change with respect to where the QL is located relative to the center of the model. These 6QL unrelaxed models give band-gap values of 0.0002 eV, 0.0027 eV, 0.0027 eV, 0.0050 eV, and 0.0050 eV for LDA, PBE, PBE + vdW, SCAN, and SCAN + vdW XCFs for Bi_2Se_3 . Similarly for Bi₂Te₃, the band gaps are 0.0007 eV, 0.0012 eV, 0.0012 eV,



FIG. 7. Band structure of 6QL Bi₂Se₃ [(a), (b)] and 6QL Bi₂Te₃ [(c), (d)] compared for PBE (left) and PBE + vdW (right). The band gap in surface states is sensitive to the thickness of the slab and the relaxation effects, which is a consequence of XCFs chosen in the calculations. \overline{M} , $\overline{\Gamma}$, and \overline{K} are the high-symmetry points of middle of the side, center, and vertex, respectively, of the surface Brillouin zone, which is a two-dimensional hexagon.

0.0008 eV, and 0.0008 eV for LDA, PBE, PBE + vdW, SCAN and SCAN + vdW, respectively. The data suggests that the band gaps are controlled by the root functionals, with no change appearing with the addition of vdW treatments.

Figure 6 shows the evolution of the band gap as a function of relaxed slab thickness for Bi_2Se_3 and Bi_2Te_3 with all XCFs. It is evident that the best agreement to experimental trends is achieved for Bi_2Se_3 by the SCAN functional, followed closely by the vdW-corrected functionals and the *GW* method. For Bi_2Te_3 , best agreement to experimental trends is achieved by the SCAN + vdW functional, followed closely by LDA and PBE + vdW. Given issues of computational efficiency and convergence we confronted with the SCAN and SCAN + vdW functional [76], we conclude PBE + vdW strikes the best balance between computability and accuracy of all the functionals tested for both systems.

Using a slightly different flavor of PBE (optPBE) and vdW treatments, the results in the band gap obtained for Bi_2Se_3 are shown to go to zero smoothly as a function of thickness [49], which is similar to our results obtained with vdW treatment, both PBE + vdW and SCAN + vdW. In the same study, the band gap of Bi_2Te_3 appears to converge toward zero from 2QL onward, which in our case (PBE + vdW) goes to zero for 3QL and larger, with all XCFs converged to zero from 4QLs. This is consistent with the experimental report of Ref. [22]. Comparing the trends that can be observed in Fig. 6, in which band gap as a function of nQL for both experimental and vdW-treated XCFs data is illustrated, we project that six QLs for Bi_2Se_3 and four QLs for Bi_2Te_3 are minimum models that capture a stable SS.

In Fig. 7, we compare the band structures produced from models optimized with PBE and PBE + vdW for Bi₂Se₃ and Bi₂Te₃. As can be seen, the vdW-corrected models produce the gapless Γ -point states and linear dispersion expected at 6QLs for both systems. In contrast, the PBE-optimized structures do not. Given that both models are derived from the same charge-density predictions, this figure demonstrates the importance of the vdW treatments in producing results that agree with experiment purely as a matter of atomic structure.

We now attempt to theoretically evaluate the SSs of our models by examining other TI characteristics besides the band gap: namely, linearity of the bands. In a regular semiconductor, the effective mass $[m^*$ as defined by Eq. (1)] captures the parabolic nature of bands, which is also a representation of deviation from linear character:

$$m^* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1}.$$
 (1)



FIG. 8. Hexagonal band warping factor λ [refer to Eq. (4)] as a function of nQL, of two-dimensional Brillouin zone for slab models of Bi₂Se₃ and Bi₂Te₃ in (a) and (b), respectively. The horizontal dashed lines are λ values obtained from experimental analysis as reported in Ref. [78].

In the limit of the ideal TI SS,

$$\frac{\partial^2 E}{\partial k^2} \to 0$$
, and band gap $\to 0$, (2)

which does not allow us to apply Eq. (1) for a direct calculation of m^* , as at the Dirac point one encounters divergence.

The theoretical band structure of TIs can be compared to experiment by calculating the hexagonal bandwarping factor, which represents the snowflake-like shape of Fermi surface for pnictogen chalcogenides, as proposed by Fu [77].

The hexagonal warping term is described by the Hamiltonian

$$H_w = \frac{\lambda}{2} (k_+^3 + k_-^3) \sigma_z,$$
 (3)

where, k_+ and k_- are the raising and lowering operators of k, σ_z is the z component of the Pauli spin matrix, and λ is the warping factor. The H_w is invariant under threefold rotation and imparts hexagonal distortion to the otherwise circular Fermi surface. Symmetry arguments hold that H_w is zero along $\overline{\Gamma} \overline{M}$ direction, but is maximum along the $\overline{\Gamma} \overline{K}$ direction. Using the dispersion of the SS obtained from the $k \cdot p$ model,

$$E(k) = E_0(k) + \sqrt{v^2 k^2 + \lambda^2 k^6 \cos^2(3\phi)},$$
 (4)

with $E_0(k) = \hbar^2 k^2 / 2m^*$, we use the recipe presented in Ref. [78], where the dispersion along $\overline{\Gamma} \overline{M}$ is fitted to Eq. (4) (not including the Γ point) to extract the band parameters. These are then applied to the energies along $\overline{\Gamma} \overline{K}$ to obtain the warping factor for each model. In terms of the band structures we present in this paper, the hexagonal band warping along $\overline{\Gamma} \overline{K}$ is distinguished by an increase in the slope of the band relative to the one measured along $\overline{\Gamma} \overline{M}$, which would otherwise be the same. The values of λ as a function of the number of QLs is shown in Fig. 8. While, as a function of nQL, there is an initial phase of growth, most of the slabs for each XCF exhibit approximately the same warping factor. Our results vary around the value of $\sim 80 \text{ eV}\text{\AA}^3$ for Bi₂Se₃ and $\sim 275 \text{ eV}\text{\AA}^3$ for Bi₂Te₃. This is consistent with the results presented in Table I of Ref. [78], where the warping factor of Bi₂Te₃ was found to be about four times that of Bi₂Se₃. The warping factor of Sn-doped Bi₂Te₃ is found to be 250 eV Å³ [77], which is of a similar range to Bi₂Te₃. Overall, the λ values obtained from DFT band structure are larger for all the models considered for both systems. An experimental study on Bi₂Se₃ reports a larger value for $\lambda = 128 \text{ eV}\text{\AA}^3$ [79].

Thus far, we have presented a picture of the SSs in Bi₂Se₃ and Bi₂Te₃ that includes only the band properties. Given the importance of spin characteristics to the TI physics, we present a detailed analysis of the evolution of the spin vector with slab thickness for the SSs. In Fig. 9(a), we have illustrated the spatial projection of the electronic spin, for the magnitude (|*S*|), the azimuthal angle (θ), and the inclination angle (ϕ). Figures 9(b) and 9(c) demonstrate the dependence of these dimensions on the slab thickness and XCFs, for Bi₂Se₃ and Bi₂Te₃, respectively. In the left panel of these figures, the dimensional values are measured near the Γ point ($k_{\Gamma\overline{K}} = 0.005 \text{ Å}^{-1}$), and, in the right panel, they are measured at $k_{\overline{\Gamma\overline{K}}} = 0.1 \text{ Å}^{-1}$. From this data, important observations can



FIG. 9. (a) A schematic illustration of the polar axes on which the spin vector is projected. (b), (c) The change in the polar coordinates of the spin vector as a function of nQL, at $k_{\overline{\Gamma K}} = 0.005 \text{ Å}^{-1}$ (left panel) and $k_{\overline{\Gamma K}} = 0.1 \text{ Å}^{-1}$ (right panel), for Bi₂Se₃ and Bi₂Te₃, respectively. The behavior of spin orientation near the Γ point $(k_{\overline{\Gamma K}} = 0.005 \text{ Å}^{-1})$ is highly coherent for all XCFs and nQLs.

be made. Notably, at the Γ point, all three dimensions exhibit convergence as a function of slab thickness. θ and ϕ converge to 90°, which implies that the spin lies in the hexagonal plane and is perpendicular to the $\overline{\Gamma K}$ direction. This is consistent with the behavior expected of the spin vector for the TSS [4,80]. At a *k* distant from the Γ point, the magnitude of spin tends to increase. But, the coherence of the orientation tends



FIG. 10. Band structures of rhombohedral Bi_2Se_3 (left double column) and Bi_2Te_3 (right double column) for LDA, PBE, SCAN, PBE + vdW, and SCAN + vdW functionals, without and with SOC, from top to bottom. The energy is scaled with respect to the valence band maximum. Tables I and II show the corresponding lattice parameters and band gaps.

to decrease, resulting in massive variance of the orientation as a function of nQL. This can be attributed to interference effects from the bulk band states as k increases.

V. CONCLUSIONS

Using prototypical TIs Bi_2Se_3 and Bi_2Te_3 , we have here systematically assessed the role of several XCFs and relaxation effects, particularly as they relate to inter- and intra-QL separations, for reliable prediction of TI properties. We show here that there is an inextricable relationship between nQL, effects of structural optimization, and the electronic structure when different versions of XCFs are employed in first-principles calculations carried out via DFT. A detailed analysis of the inter- and intra-layer relaxations reveal that GGA overestimates inter- and intra-QL separation, contradicting experimental findings [67]. On the other hand, the relatively simple LDA functional produces an opposite trend in relaxation effects, also in contradiction of experimental results. When more complicated XCFs such as metaGGA are employed, the relaxed structural parameters are found to be between the values obtained from experimental work and GGA. With vdW treatments applied, GGA and metaGGA predict bulklike crystal structures. In addition to structural parameters, we also investigate the thickness dependence of the band gap for Bi_2Se_3 and Bi_2Te_3 for all XCFs. We ultimately determine that GGA + vdW offers the most reliable results, for both accuracy—when compared to experimental studies-and computability (computational resource requirements and electronic convergence). We also determine that the GGA + vdW functional produces a structure that does not deviate much from the bulk, and generates a thickness-bandgap dependence that agrees well with experimental results. Given the comparable atomic numbers of the high-Z (and thus, strong SOC) pnictogens and chalcogenides, it seems reasonable to expect that, in other binary combinations with layered structures of these two sets of elements, that the vdW forces would be approximately as strong as they are in Bi₂Se₃ and Bi₂Te₃. We thus conclude that use of the experimental bulk parameters as the basis for a slab might be sufficient to model TI physics in binary layered pnictogen chalcogenides of high-Z components, and that the base GGA functional is sufficient for the SCCD and band-structure calculation phases. It must be noted that modeling TI systems of this type with complex nonstoichiometric chemistry, doping effects, mechanical deformations (such as uniaxial strain or biaxial strain in order to account for Poisson's effect) would require structural optimization, where the GGA + vdW is most appropriate. Comparing the trends in band gap and spin texture near the Γ -point as a function of nQL for vdW-treated XCFs

- [1] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. 5, 438 (2009).
- [2] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nat. Phys. 5, 398 (2009).
- [3] D. Hsieh, Y. Xia, D. Qian, L. Wray, J. H. Dil, F. Meier, J. Osterwalder, L. Patthey, J. G. Checkelsky, N. P. Ong, A. V. Fedorov, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature 460, 1101 (2009).
- [4] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. 82, 3045 (2010).
- [5] J. E. Moore, Nature 464, 194 (2010).
- [6] 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, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, Nat. Phys. 6, 584 (2010).
- [7] A. A. Taskin, S. Sasaki, K. Segawa, and Y. Ando, Adv. Mater. 24, 5581 (2012).
- [8] P. Tsipas, E. Xenogiannopoulou, S. Kassavetis, D. Tsoutsou, E. Golias, C. Bazioti, G. P. Dimitrakopulos, P. Komninou, H. Liang, M. Caymax, and A. Dimoulas, ACS Nano 8, 6614 (2014).
- [9] S. Das Sarma, M. Freedman, and C. Nayak, Phys. Today 59(7), 32 (2006).
- [10] K. L. Wang, M. Lang, and X. Kou, Spintronics of topological insulators, in *Handbook of Spintronics*, edited by Y. Xu, D. D. Awschalom, and J. Nitta (Springer Netherlands, Dordrecht, 2016), pp. 431–462.
- [11] Y. Fan and K. L. Wang, SPIN 06, 1640001 (2016).
- [12] F. Zhang, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. 110, 046404 (2013).
- [13] F. Zhang, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. 111, 056403 (2013).
- [14] C. L. Kane and F. Zhang, Phys. Scr. T164, 014011 (2015).

and experimental results, we conclude that the six QL model for Bi_2Se_3 and the four QL models for Bi_2Te_3 possess the minimum necessary thicknesses for capturing TI properties.

ACKNOWLEDGMENTS

The authors would like to acknowledge useful discussions with M. Geilhufe and G. Fernando. A.V.B.'s work was supported by the VILLUM FONDEN via the Centre of Excellence for Dirac Materials (Grant No. 11744) and the European Research Council under the European Unions Seventh Framework Program Synergy HERO.

The computational resources employed in this study were provided by High Performance Computing facility of University of Connecticut.

APPENDIX

Figure 10 shows the band dispersion obtained for bulk Bi_2Se_3 and bulk Bi_2Te_3 with and without spin-orbit coupling for the XCFs LDA, PBE, SCAN, PBE+vdW, and SCAN+vdW.

- [15] M. A. Metlitski, C. L. Kane, and M. P. A. Fisher, Phys. Rev. B 92, 125111 (2015).
- [16] J. C. Y. Teo, L. Fu, and C. L. Kane, Phys. Rev. B 78, 045426 (2008).
- [17] D. Hsieh, Y. Xia, L. Wray, D. Qian, A. Pal, J. H. Dil, J. Osterwalder, F. Meier, G. Bihlmayer, C. L. Kane, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Science 323, 919 (2009).
- [18] L. Fu and C. L. Kane, Phys. Rev. Lett. 100, 096407 (2008).
- [19] C. L. Kane and E. J. Mele, Phys. Rev. Lett. 95, 146802 (2005).
- [20] T. Wehling, A. Black-Schaffer, and A. Balatsky, Adv. Phys. 63, 1 (2014).
- [21] S. Nakajima, J. Phys. Chem. Solids 24, 479 (1963).
- [22] Y. Liu, G. Bian, T. Miller, M. Bissen, and T.-C. Chiang, Phys. Rev. B 85, 195442 (2012).
- [23] C.-X. Liu, H. J. Zhang, B. Yan, X.-L. Qi, T. Frauenheim, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B 81, 041307(R) (2010).
- [24] S. M. Young, S. Chowdhury, E. J. Walter, E. J. Mele, C. L. Kane, and A. M. Rappe, Phys. Rev. B 84, 085106 (2011).
- [25] W. Liu, X. Peng, C. Tang, L. Sun, K. Zhang, and J. Zhong, Phys. Rev. B 84, 245105 (2011).
- [26] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science 321, 792 (2008).
- [27] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Chem. Rev. 112, 289 (2012).
- [28] B. Santra and J. P. Perdew, J. Chem. Phys. 150, 174106 (2019).
- [29] A. Zunger, S. Lany, and H. Raebiger, Physics 3, 53 (2010).
- [30] W. A. Adeagbo, S. Thomas, S. K. Nayak, A. Ernst, and W. Hergert, Phys. Rev. B 89, 195135 (2014).
- [31] S. K. Nayak, H. T. Langhammer, W. A. Adeagbo, W. Hergert, T. Müller, and R. Böttcher, Phys. Rev. B 91, 155105 (2015).
- [32] W. A. Adeagbo, H. Ben Hamed, S. K. Nayak, R. Böttcher, H. T. Langhammer, and W. Hergert, Phys. Rev. B 100, 184108 (2019).

- [33] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. 92, 246401 (2004).
- [34] K. Berland, D. Chakraborty, and T. Thonhauser, Phys. Rev. B 99, 195418 (2019).
- [35] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, Phys. Rev. B 52, R5467 (1995).
- [36] A. D. Becke and E. R. Johnson, J. Chem. Phys. 124, 221101 (2006).
- [37] Y. Zhao and D. G. Truhlar, J. Chem. Phys. 125, 194101 (2006).
- [38] O. V. Yazyev, E. Kioupakis, J. E. Moore, and S. G. Louie, Phys. Rev. B 85, 161101(R) (2012).
- [39] O. V. Yazyev, J. E. Moore, and S. G. Louie, Phys. Rev. Lett. 105, 266806 (2010).
- [40] W. Li, X.-Y. Wei, J.-X. Zhu, C. S. Ting, and Y. Chen, Phys. Rev. B 89, 035101 (2014).
- [41] L. L. Tao and E. Y. Tsymbal, Nat. Commun. 9, 2763 (2018).
- [42] C. Datzer, A. Zumbülte, J. Braun, T. Förster, A. B. Schmidt, J. Mi, B. Iversen, P. Hofmann, J. Minár, H. Ebert, P. Krüger, M. Rohlfing, and M. Donath, Phys. Rev. B 95, 115401 (2017).
- [43] L. Seixas, D. West, A. Fazzio, and S. B. Zhang, Nat. Commun. 6, 7360 (2015).
- [44] C. Seibel, J. Braun, H. Maaß, H. Bentmann, J. Minár, T. V. Kuznetsova, K. A. Kokh, O. E. Tereshchenko, T. Okuda, H. Ebert, and F. Reinert, Phys. Rev. B 93, 245150 (2016).
- [45] I. P. Rusinov, I. A. Nechaev, and E. V. Chulkov, J. Exp. Theor. Phys. 116, 1006 (2013).
- [46] T. Förster, P. Krüger, and M. Rohlfing, Phys. Rev. B 92, 201404(R) (2015).
- [47] T. Förster, P. Krüger, and M. Rohlfing, Phys. Rev. B 93, 205442 (2016).
- [48] I. Aguilera, C. Friedrich, G. Bihlmayer, and S. Blügel, Phys. Rev. B 88, 045206 (2013).
- [49] W. Liu, X. Peng, X. Wei, H. Yang, G. M. Stocks, and J. Zhong, Phys. Rev. B 87, 205315 (2013).
- [50] X. Luo, M. B. Sullivan, and S. Y. Quek, Phys. Rev. B 86, 184111 (2012).
- [51] K. Shirali, W. A. Shelton, and I. Vekhter, arXiv:1905.01269.
- [52] J. Li, Y. Li, S. Du, Z. Wang, B.-L. Gu, S.-C. Zhang, K. He, W. Duan, and Y. Xu, Sci. Adv. 5, eaaw5685 (2019).
- [53] G. Cao, H. Liu, J. Liang, L. Cheng, D. Fan, and Z. Zhang, Phys. Rev. B 97, 075147 (2018).
- [54] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. 132, 154104 (2010).
- [55] G. Kresse and J. Furthmüller, Comput. Mater. Sci. 6, 15 (1996).
- [56] G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996).
- [57] S. Steiner, S. Khmelevskyi, M. Marsmann, and G. Kresse, Phys. Rev. B 93, 224425 (2016).
- [58] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. 45, 566 (1980).
- [59] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996).
- [60] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. 115, 036402 (2015).
- [61] S. Sahoo, S. P. Alpay, and R. J. Hebert, Surf. Sci. 677, 18 (2018).

- [62] S. Sahoo, S. N. Khanna, and P. Entel, Appl. Phys. Lett. 107, 043102 (2015).
- [63] S. Sahoo, A. Hucht, M. E. Gruner, G. Rollmann, P. Entel, A. Postnikov, J. Ferrer, L. Fernández-Seivane, M. Richter, D. Fritsch, and S. Sil, Phys. Rev. B 82, 054418 (2010).
- [64] G. Martinez, B. A. Piot, M. Potemski, Y. S. Hor, A. Materna, S. G. Strzelecka, A. Hruban, O. Caha, J. Novák, A. Dubroka, v. Drašar, and M. Orlita, Sci. Rep. 7, 6891 (2017).
- [65] I. A. Nechaev, R. C. Hatch, M. Bianchi, D. Guan, C. Friedrich, I. Aguilera, J. L. Mi, B. B. Iversen, S. Blügel, P. Hofmann, and E. V. Chulkov, Phys. Rev. B 87, 121111(R) (2013).
- [66] C.-Y. Li, A. L. Ruoff, and C. W. Spencer, J. Appl. Phys. 32, 1733 (1961).
- [67] S. Roy, H. L. Meyerheim, K. Mohseni, A. Ernst, M. M. Otrokov, M. G. Vergniory, G. Mussler, J. Kampmeier, D. Grützmacher, C. Tusche, J. Schneider, E. V. Chulkov, and J. Kirschner, Phys. Rev. B 90, 155456 (2014).
- [68] N. Fukui, T. Hirahara, T. Shirasawa, T. Takahashi, K. Kobayashi, and S. Hasegawa, Phys. Rev. B 85, 115426 (2012).
- [69] S. S. Hong, W. Kundhikanjana, J. J. Cha, K. Lai, D. Kong, S. Meister, M. A. Kelly, Z.-X. Shen, and Y. Cui, Nano Lett. 10, 3118 (2010).
- [70] S. Roy, H. L. Meyerheim, A. Ernst, K. Mohseni, C. Tusche, M. G. Vergniory, T. V. Menshchikova, M. M. Otrokov, A. G. Ryabishchenkova, Z. S. Aliev, M. B. Babanly, K. A. Kokh, O. E. Tereshchenko, E. V. Chulkov, J. Schneider, and J. Kirschner, Phys. Rev. Lett. **113**, 116802 (2014).
- [71] S. K. Nayak, M. Ogura, A. Hucht, H. Akai, and P. Entel, J. Phys.: Condens. Matter 21, 064238 (2009).
- [72] M. Geilhufe, S. K. Nayak, S. Thomas, M. Däne, G. S. Tripathi, P. Entel, W. Hergert, and A. Ernst, Phys. Rev. B 92, 235203 (2015).
- [73] J. Liu and D. Vanderbilt, Phys. Rev. B 90, 125133 (2014).
- [74] The strategy we have adopted is structural relaxation without SOC and then using the relaxed structure for final SCF calculation including SOC.
- [75] J. M. Crowley, J. Tahir-Kheli, and W. A. Goddard, J. Phys. Chem. Lett. 6, 3792 (2015).
- [76] We confronted significant convergence issues in the use of SCAN and SCAN+vdW, which we overcame by first converging in GGA and preserving the CHGCAR and WAVECAR files, then restarting the calculations with the following mixing tags: ISTART = 1, IMIX = 1, AMIX = 0.05, BMIX = 0.0001, AMIX_MAG = 0.1, BMIX_MAG = 0.0001.
- [77] L. Fu, Phys. Rev. Lett. 103, 266801 (2009).
- [78] M. Nomura, S. Souma, A. Takayama, T. Sato, T. Takahashi, K. Eto, K. Segawa, and Y. Ando, Phys. Rev. B 89, 045134 (2014).
- [79] K. Kuroda, M. Arita, K. Miyamoto, M. Ye, J. Jiang, A. Kimura, E. E. Krasovskii, E. V. Chulkov, H. Iwasawa, T. Okuda, K. Shimada, Y. Ueda, H. Namatame, and M. Taniguchi, Phys. Rev. Lett. 105, 076802 (2010).
- [80] A. Bansil, H. Lin, and T. Das, Rev. Mod. Phys. 88, 021004 (2016).