中国物理 B
Chinese Physics
Chinese Physics B

SymTopo: An automa tictool for calculating topological properties of nonmagnetic crystalline materials

Yuqing He(贺雨晴), Yi Jiang(蒋毅), Tiantian Zhang(张田田), He Huang(黄荷), Chen Fang(方辰), and Zhong Jin(金钟)
Citation: Chin. Phys. B . 2019, 28(8): 087102 . **doi:** 10.1088/1674-1056/28/8/087102
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Structural, elastic, and electronic properties of topological semimetal WC-type *MX* family by first-principles calculation

Sami Ullah, Lei Wang(王磊), Jiangxu Li(李江旭), Ronghan Li(李荣汉), Xing-Qiu Chen(陈星秋)
Chin. Phys. B . 2019, 28(7): 077105 . **doi:** 10.1088/1674-1056/28/7/077105

Realization of low-energy type-Ⅱ Dirac fermions in $(Ir_{1-x}Pt_x)Te_2$ superconductors

Bin-Bin Fu(付彬彬), Chang-Jiang Yi(伊长江), Zhi-Jun Wang(王志俊), Meng Yang(杨萌), Bai-Qing Lv(吕佰晴), Xin Gao(高鑫), Man Li(李满), Yao-Bo Huang(黄耀波), Hong-Ming Weng(翁红明), You-Guo Shi(石友国), Tian Qian(钱天), Hong Ding(丁洪)
Chin. Phys. B . 2019, 28(3): 037103 . **doi:** 10.1088/1674-1056/28/3/037103

Band structure engineering and defect control of oxides for energy applications

Hui-Xiong Deng(邓惠雄), Jun-Wei Luo(骆军委), Su-Huai Wei(魏苏淮)
Chin. Phys. B . 2018, 27(11): 117104 . **doi:** 10.1088/1674-1056/27/11/117104

Magnetic interactions in a proposed diluted magnetic semiconductor $(Ba_{1-x}K_x)(Zn_{1-y}Mn_y)_2P_2$

Huan-Cheng Yang(杨焕成), Kai Liu(刘 凯), Zhong-Yi Lu(卢仲毅)
Chin. Phys. B . 2018, 27(6): 067103 . **doi:** 10.1088/1674-1056/27/6/067103

Topological nodal line semimetals

Chen Fang(方辰), Hongming Weng(翁红明), Xi Dai(戴希), Zhong Fang(方忠)
Chin. Phys. B . 2016, 25(11): 117106 . **doi:** 10.1088/1674-1056/25/11/117106

SymTopo: An automatic tool for calculating topological properties of nonmagnetic crystalline materials*

Yuqing He(贺雨晴)[1,2,†], Yi Jiang(蒋毅)[3,2,†], Tiantian Zhang(张田田)[3,2,†], He Huang(黄荷)[1,4,‡], Chen Fang(方辰)[2,3], and Zhong Jin(金钟)[1,4,§]

[1] *Computer Network Information Center, Chinese Academy of Sciences, Beijing* 100190, *China*
[2] *University of Chinese Academy of Sciences, Beijing* 100049, *China*
[3] *Beijing National Laboratory for Condensed Matter Physics, and Institute of Physics, Chinese Academy of Sciences, Beijing* 100190, *China*
[4] *Center of Scientific Computing Applications & Research, Chinese Academy of Sciences, Beijing* 100190, *China*

(Received 17 May 2019; revised manuscript received 24 June 2019; published online 25 July 2019)

Topological materials have novel properties both in their bulk and boundaries, thereby attracting a wide interest in the theoretical and experimental communities. The recent development of the topological quantum chemistry and symmetry-based indicator theory in this field has significantly simplified the procedure for determining the topological properties of nonmagnetic crystalline materials. Accordingly, a large number of new topological materials have been found by scanning large crystal databases. This study provides details on the algorithm used in the Catalogue of Topological Electronic Materials. Moreover, based on the algorithm, we develop an automatic package named SymTopo, which calculates the symmetry representations of any given nonmagnetic crystalline material and predicts its topological properties. This package may facilitate the discovery of more topological materials in the future.

Keywords: topological insulators, topological semimetals, topological invariants, high-throughput research mode

PACS: 71.20.–b, 81.05.Zx **DOI:** 10.1088/1674-1056/28/8/087102

1. Introduction

Topology is a branch of modern mathematics that studies the features of shapes that are invariant under continuous deformation. It has found wide applications in the study of condensed matter physics in the recent years. For example, the entire field of topological insulators derives from the intricate interplay of topology and symmetry[1–4] and gives birth to many topological materials with exotic properties, such as surface states free from back-scattering[5] and Majorana zero modes[6] that may play a key role in fault-tolerant quantum computation.[7] The first step of discovering a new topological material is typically the numerical prediction of its potential topological properties,[8–12] which is a difficult step because of the involved expression of the topological invariants.

Recent theoretical progress, including theory of topological quantum chemistry,[13] symmetry-based indicators,[14] and the following works of mapping the irreducible representations (irreps) of valence bands onto topological invariants[15–17] and topological nodes,[18] has greatly relieved the workload needed for predicting the nontrivial topology in materials. According to these theories, only the irreps of the valence bands at the high-symmetry points (HSPs) in the Brillouin zone (BZ) (at most 8 in 230 space groups) are required to diagnose the topological properties, which significantly reduce the amount of first-principles calculation, making it possible to scan through large material databases to search for topological materials. Three independent works published in Nature[19–21] adopted similar algorithms and found thousands of new topological materials. A brief comparison of these three works is given in Appendix A.

In our article titled "Catalogue of topological electronic materials",[19] we developed a fully automatic algorithm and scanned approximately 4×10^4 crystalline materials, finding more than 8000 of them (approximately 9000 if ignoring the spin–orbital coupling effect) being topologically nontrivial. We classified these topological materials into "high-symmetry point semimetal" (HSPSM), "high-symmetry line semimetal" (HSLSM), "generic-momenta semimetal" (GMSM), "topological insulator" (TI), and "topological crystalline insulator" (TCI). The first three are topological semimetals with topological nodes at HSPs, high-symmetry lines (HSLs), and generic points in the BZ, respectively, whereas the last two are topological gapped systems protected by time-reversal symmetry and crystalline symmetries, respectively. All the results from the sweeping scan can be found at http://materiae.iphy.ac.cn.

Here, we give a thorough description of the algorithm used in the sweeping scan and elucidate important technical details that are missing in the previous article because of a limited article length of *Nature*. Meanwhile, we report the development of an automated package named SymTopo based on

*Project supported by the Information Program of the Chinese Academy of Sciences (Grant No. XXH13506-202).
†These authors contributed equally to this work.
‡Corresponding author. E-mail: huanghe@sccas.cn
§Corresponding author. E-mail: zjin@sccas.cn
© 2019 Chinese Physical Society and IOP Publishing Ltd

the algorithm. The package interfaces with two widely used first-principles calculation software Vienna *ab initio* simulation package (VASP)[22–25] and ABINIT,[26–29] enabling users with or without expertise in the symmetry-based indicators theory to easily diagnose the topological properties of materials using our algorithm.

2. Algorithm

Figure 1 summarizes the automatic algorithm for determining the topological classification of a nonmagnetic crystalline material. Starting from an arbitrary nonmagnetic crystalline material, the main steps of the algorithm are as follows:

(i) Check the number of electrons per unit cell. The material is labeled as a conventional metal if it has an odd number of electrons per unit cell, and no further analysis is performed.

(ii) Identify the space group of the crystal structure and standardize the structure for the convenience of further calculations.

(iii) Perform the density functional theory (DFT) calculation for the standardized structure to obtain the wavefunctions and energy eigenvalues at each HSP in the BZ. The calculation is typically performed using selected first-principles software.

(iv) Check the degeneracy between valence and conduction bands at the HSPs. This step is only performed if the material has a direct band gap smaller than 2 meV. The material is labeled as an HSPSM if such degeneracy exists, and no further analysis is performed.

(v) Calculate the irreps of all valence bands at each HSP.

(vi) Check the band crossings between the valence and conduction bands along the HSLs. To do so, the irreps of all valence bands are checked against the compatibility conditions. The violation of the compatibility conditions indicates the existence of band crossings along HSLs, in which case the material is labeled as an HSLSM, and no further analysis is performed; otherwise, proceed into the next step.

(vii) Calculate the symmetry-based indicators. Materials with nonzero indicators are labeled as GMSM in the nsoc setting and TI or TCI in the soc setting. This concludes the algorithm.

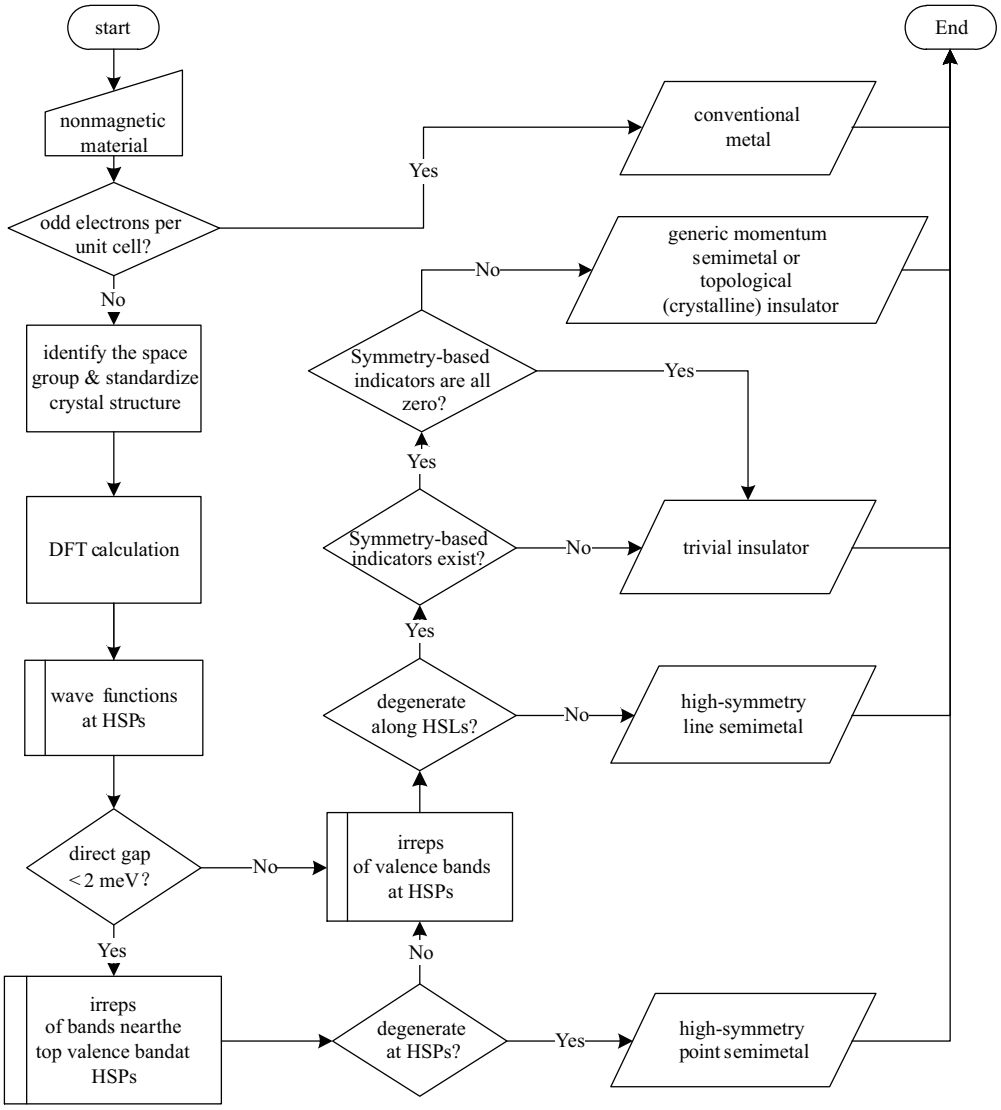


Fig. 1. Flowchart of the algorithm.

In the following subsections, we walk through the diagnosis process with two well-known topological materials, namely SnTe and HgTe, as examples. The technical issues involved in the implementation are described in detail.

2.1. Space group identification and crystal structure standardization

In the preparation step, we first identify the space group of the crystal structure for a nonmagnetic material with an even number of electrons per unit cell and standardize it using an open-source computational material science package, called Phonopy.[30]

The structure data of a material may have numerical errors and different choices of coordinate axes and origin. We use Phonopy to standardize the structure to facilitate high-throughput calculations. Given an input structure and error tolerance, Phonopy will identify its corresponding space group, minimize the error of the atomic positions with respect to the space group, and transform the coordinate system to a consistent primitive cell basis convention. This consistent convention provides a basis to define the matrix forms of the symmetry operations in each space group that are used to calculate the irreps of the wavefunctions later.

We set a default error tolerance of 10^{-4} Å for Phonopy to identify the space group. If the input crystal structure is already associated with a space group and the designation is different from Phonopy's result, we increase the error tolerance until Phonopy finds the designated space group or the error tolerance reaches 0.1 Å, in which case the input data are considered to have large numerical errors and should be replaced by a more accurate structure.

Take SnTe of space group 225 as an example. Table 1 lists its primitive cell basis vectors of the original and standardized structures. Notice that the coordinate system is transformed.

Table 1. Primitive cell basis vectors of SnTe in the soc setting (unit: Å).

Original structure	
$\boldsymbol{a}_1$	(3.8690, 0.0000, 2.2338)
$\boldsymbol{a}_2$	(1.2897, 3.6477, 2.2338)
$\boldsymbol{a}_3$	(0.0000, 0.0000, 4.4675)
Standardized structure	
$\boldsymbol{a}_1$	(0.0000, 3.1590, 3.1590)
$\boldsymbol{a}_2$	(3.1590, 0.0000, 3.1590)
$\boldsymbol{a}_3$	(3.1590, 3.1590, 0.0000)

2.2. DFT calculation

The standardized crystal structure is fed to the first-principles calculation software once obtained. Following the standard DFT calculation procedure, the self-consistent electron density is first calculated and used in the second step to compute the wavefunctions along with the energy eigenvalues at a given list of HSPs in the BZ. The lowest N bands (N being the number of electrons per unit cell) at each HSP are defined as the valence bands.

An HSP is a k-point in the BZ with the highest symmetry in its neighborhood. The positions of the HSPs are determined by the space group with a maximum number of 8 in all 230 space groups. The coordinates of the HSPs for each space group can be found at "*The k-vector types and Brillouin zones of Space Groups*" on the Bilbao Crystallographic Server (BCS).[31,32] Two HSPs related by time-reversal symmetry (with opposite coordinates and one with a suffix "*A*," *e.g.*, "*K*" and "*KA*") are both listed on this page, and we only use one of them.

The HSPs and the symmetry operations of the 230 space groups on the BCS are given in its conventional cell basis convention. They must be transformed into Phonopy's primitive cell basis convention to be consistent with the crystal structure and the DFT calculation results. These two sets of convention differ by a similarity transformation of the coordinate axes and a shift of origin, which are listed in Appendix B. In either case, the coordinates of the HSPs are given as fractions of the reciprocal lattice basis, with the reciprocal lattice basis vector being defined by the standard formulas

$$\boldsymbol{b}_1 = 2\pi \frac{\boldsymbol{a}_2 \times \boldsymbol{a}_3}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)}, \tag{1}$$

$$\boldsymbol{b}_2 = 2\pi \frac{\boldsymbol{a}_3 \times \boldsymbol{a}_1}{\boldsymbol{a}_2 \cdot (\boldsymbol{a}_3 \times \boldsymbol{a}_1)}, \tag{2}$$

$$\boldsymbol{b}_3 = 2\pi \frac{\boldsymbol{a}_1 \times \boldsymbol{a}_2}{\boldsymbol{a}_3 \cdot (\boldsymbol{a}_1 \times \boldsymbol{a}_2)}, \tag{3}$$

where $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ are the basis vectors of the primitive cell.

SnTe of space group 225 only has four HSPs in the BZ, namely $\Gamma = [0,0,0]$, $L = [1/2,1/2,1/2]$, $W = [1/2,1/4,3/4]$, and $X = [1/2,0,1/2]$, with the fractional number being the coordinates under the reciprocal lattice basis in Phonopy's primitive cell convention.

We conducted the calculation with the generalized gradient approximation of the Perdew–Burke–Ernzerhof (PBE)-type exchange-correlation potential using the pseudopotential files of the projector augmented-wave atomic data (PAW_GGA_PBE). Subsection 3.1 lists the other important input parameters. The calculation may be performed with the spin–orbit coupling turned on and off, which are briefly called the "soc setting" and the "nsoc setting," respectively. Note that although spin–orbit coupling always exists in realistic materials, ignoring it is a more physically relevant approximation for systems consisting of small atoms. For materials with a large spin–orbit coupling, the results obtained in the nsoc setting can be used as a reference and may help diagnose the origin of topological properties in the soc setting.

2.3. Irreducible representations extraction

Extracting the irrep of each (multiplet of) valence band(s) is one of the core modules of the algorithm. In the case of one-dimensional irreps (the higher dimensional case will be discussed later), the wavefunction of each band at each HSP has

the corresponding little group symmetry and is a basis function of one of the little group's irreps. For such a wavefunction, its corresponding irrep is extracted in two steps. First, we calculate its character with each symmetry operation in the little group. Second, we compare these characters to the character tables on the BCS to identify the irreps. Our algorithm for extracting irreps currently only supports wavefunctions in a plane wave basis. We discuss the main steps in the following paragraphs.

Assume that the wavefunction in the nsoc setting is expanded in the plane wave basis:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r}}, \tag{4}$$

where $\boldsymbol{k}$ denotes the k-point; $\boldsymbol{G}$ denotes the reciprocal lattice vectors; and $C_{\boldsymbol{G}}^{\boldsymbol{k}}$ stands for the plane wave expansion coefficients. The band index n is omitted here for convenience. The plane wave coefficients can be readily read from the DFT results.

In the case of one-dimensional irreps, the character of $|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle$ under a certain symmetry operation $\{\boldsymbol{R}|\boldsymbol{\tau}\}$ is simply the eigenvalue λ:

$$\{\boldsymbol{R}|\boldsymbol{\tau}\}\,\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \lambda\,|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle, \tag{5}$$

which can be reformulated as follows (assuming that $|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle$ is normalized, and the volume $V = 1$):

$$\begin{aligned}\lambda &= \langle\psi_{\boldsymbol{k}}(\boldsymbol{r})|\,\{\boldsymbol{R}|\boldsymbol{\tau}\}\,|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle \\ &= \sum_{\boldsymbol{G}}\sum_{\boldsymbol{G}'} \left(C_{\boldsymbol{G}'}^{\boldsymbol{k}}\right)^{*} C_{\boldsymbol{G}}^{\boldsymbol{k}} \delta_{\boldsymbol{G}',(\boldsymbol{k}+\boldsymbol{G})\boldsymbol{R}^{-1}-\boldsymbol{k}} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\boldsymbol{R}^{-1}\boldsymbol{\tau}}.\end{aligned} \tag{6}$$

The situation with high-dimensional irreps is similar. Assume that $\left\{\psi_{\boldsymbol{k}},j=1,2,\ldots,n\right\}$ are degenerate and transform under a symmetry operation as follows:

$$\{\boldsymbol{R}|\boldsymbol{\tau}\}\,|\psi_{\boldsymbol{k},i}\rangle = \sum_{j=1}^{n} \left|\psi_{\boldsymbol{k},j}\right\rangle C_{ji}, \tag{7}$$

where C is the representation matrix, $\boldsymbol{R}$ is the O(3) rotation matrix, and $\boldsymbol{\tau}$ is the translation part. The character is the trace

$$\lambda = \sum_{i=1}^{n} C_{ii} = \sum_{i=1}^{n} \langle\psi_{\boldsymbol{k},i}|\,\{\boldsymbol{R}|\boldsymbol{\tau}\}\,|\psi_{\boldsymbol{k},i}\rangle. \tag{8}$$

The wavefunction in the soc setting has two components, one spin-up and one spin-down:

$$|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle = \sum_{\boldsymbol{G}} \begin{pmatrix} C_{\boldsymbol{G}\uparrow}^{\boldsymbol{k}} \\ C_{\boldsymbol{G}\downarrow}^{\boldsymbol{k}} \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+G)\cdot\boldsymbol{r}}. \tag{9}$$

A symmetry operation acts on the wavefunction as

$$\{\boldsymbol{R},\boldsymbol{U},\boldsymbol{\tau}\}\,|\psi_{\boldsymbol{k}}(\boldsymbol{r})\rangle = \sum_{\boldsymbol{G}} \boldsymbol{U}\cdot \begin{pmatrix} C_{\boldsymbol{G}\uparrow}^{\boldsymbol{k}} \\ C_{\boldsymbol{G}\downarrow}^{\boldsymbol{k}} \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\{\boldsymbol{R}|\boldsymbol{\tau}\}^{-1}\cdot\boldsymbol{r}}, \tag{10}$$

where $\boldsymbol{U}$ is the SU(2) matrix corresponding to the O(3) matrix $\boldsymbol{R}$.

However, the representation of the translation subgroup of each space group on BCS takes the convention

$$D^{k}(\boldsymbol{R}_1) = \exp(+\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_1), \tag{11}$$

instead of the usual form

$$D^{k}(\boldsymbol{R}_1) = \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_1) \tag{12}$$

at point $\boldsymbol{k}$. Accordingly, we must modify the acting rule when applying the symmetry operation onto the Bloch wavefunction as

$$\{\boldsymbol{R},\boldsymbol{U},\boldsymbol{\tau}\}\,\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} \boldsymbol{U}^{\dagger}\cdot \begin{pmatrix} C_{\boldsymbol{G}\uparrow}^{\boldsymbol{k}} \\ C_{\boldsymbol{G}\downarrow}^{\boldsymbol{k}} \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\boldsymbol{G})\cdot\{\boldsymbol{R}|\boldsymbol{\tau}\}\cdot\boldsymbol{r}}; \tag{13}$$

the representation matrix transforms the basis functions as

$$\{\boldsymbol{R}|\boldsymbol{\tau}\}\,|\psi_{\boldsymbol{k},i}\rangle = \sum_{j=1}^{n} C_{ij}\left|\psi_{\boldsymbol{k},j}\right\rangle. \tag{14}$$

The wavefunction typically has tens of thousands of plane wave components, and calculating the character can be very time consuming. However, the O(3) matrix is isometric (*i.e.*, it preserves the norm of vectors); hence, in principle, we only need to use one shell of plane wave components with constant $|\boldsymbol{k}+\boldsymbol{G}|$. In real calculations, we use $|\boldsymbol{k}+\boldsymbol{G}| < 5$ Å^{-1} as a truncation condition, which corresponds to several hundreds or thousands of plane waves. Such a truncation may fail in some extreme cases (*e.g.*, when the primitive cell is extremely small). In these cases, we redo the calculation without the truncation.

We identify the irrep of the wavefunction when the characters of all the symmetry operations are calculated. This step uses the orthogonal theorem in group representation theory:

$$\frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} \chi^{i*}(g)\chi^{j}(g) = \delta_{ij}, \tag{15}$$

where $\mathcal{G}$ is a group, and χ^{i} and χ^{j} are two group representations. We calculate the inner product between the obtained characters and those of each irrep of the little group to identify that with a result of 1. The inner product in real calculations is not strictly 1 because of numerical errors, and we set a maximum error of 0.05. Failing to find the irrep within the error range implies a low convergence quality. Hence, increasing the convergence threshold and redoing the DFT calculation are recommended.

Note that although a set of bands belonging to a single high-dimensional irrep is, in principle, degenerate, a small energy difference between such bands may be observed because of numerical errors. As a result, we set a degeneracy error of min (0.5 meV, $0.1\Delta_{\boldsymbol{k}}$), where $\Delta_{\boldsymbol{k}}$ is the direct band gap at the k-point. We split the valence bands into small sets, in which bands have an energy difference smaller than this error, and calculate the irrep for each set of bands separately.

Moreover, time-reversal symmetry always exists in nonmagnetic materials, thereby causing an additional degeneracy between the original irreps of the space groups. An irrep at a high-symmetry momentum is generally mapped under time reversal to either (i) itself or (ii) another different irrep or (iii) another same irrep. As a result, time-reversal symmetry puts different constraints on the occupation numbers of different types of irreps in the valence bands: For type-(i) irreps, the

occupation numbers do not have additional constraints; for type-(ii) irreps, the occupation numbers for the two irreps related by time-reversal symmetry are constrained to equal; and for type-(iii) irreps, the occupation number of a type-(iii) irrep must be even. The compatibility relations are given in terms of the linearly independent occupation numbers; hence, we must extract them after considering these constraints.

The symmetry operations and the character tables of the 230 space groups are obtained from the "*Irreducible representations of the double space groups*" database on the BCS. Note that two tables can be found on the page, and we only use the first one (*i.e.*, representation matrices of the little group). The additional degeneracy caused by the time-reversal symmetry is not included in this page, and we extract this information separately from the "*Band representations and elementary band representations of double space groups*" database on the BCS, in which the irreps written together like "$\Gamma_1\Gamma_2(4)$" are degeneracies caused by time-reversal symmetry (the number 4 in the parentheses denotes the dimension of the whole degenerate irrep).

As mentioned in Subsection 2.2, the symmetry operations on the BCS are given in its conventional cell convention. We transform their coordinate axes and origin to make them consistent with Phonopy's primitive cell convention. A change of the coordinate axes transforms the O(3) matrix of the symmetry operation, whereas a different choice of origin changes the translation part by the formula

$$\boldsymbol{\tau}' = \boldsymbol{s} - \boldsymbol{R}\boldsymbol{s} + \boldsymbol{\tau}, \tag{16}$$

where $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$ denote the translation part in the old and new coordinate systems, respectively; $\boldsymbol{R}$ is the corresponding rotation matrix; and $\boldsymbol{s}$ denotes the shift vector of the origin. The derivation of this formula can be found in Appendix C.

The SU(2) part of the symmetry operations is directly calculated using the standard formula as follows:

$$\boldsymbol{u} = \cos\left(\frac{w}{2}\right)\boldsymbol{\sigma}_0 - \mathrm{i}\sin\left(\frac{w}{2}\right)\boldsymbol{n}\cdot\boldsymbol{\sigma}, \tag{17}$$

where $\boldsymbol{n}$ and $w \in [0, 2\pi)$ are the rotation axis and the angle obtained from the corresponding O(3) matrix in Cartesian coordinates, respectively, and $\{\boldsymbol{\sigma}_0, \boldsymbol{\sigma}\}$ denotes the Pauli matrices. The SU(2) matrices on the BCS use different spin basis convention choices and differ from the SU(2) matrices calculated in Cartesian coordinates by a similarity transformation and a phase factor of ± 1. Note that in double space group representations, each O(3) operator corresponds to two SU(2) matrices differing by ± 1, and these two SU(2) matrices correspond to two rows in the character table that also differ by ± 1. As a result, the calculated SU(2) matrices must be aligned with the SU(2) matrices on the BCS such that one can use the character table of the double space groups on the BCS. Although the similarity transformation is unimportant because it does not affect the irreps, the uncertainty of the ± 1 factor must be eliminated. This step is done by adding ± 1 factors to the calculated SU(2) matrices until identical multiplication tables of the two sets of SU(2) matrices are obtained.

Table 2 lists the calculated irreps of SnTe in the soc setting, with the number in the parentheses denoting the irrep dimension. Notice that SnTe has 20 electrons per unit cell (*i.e.*, it has 20 valence bands); this number of electrons equals the sum of the number of irreps weighted based on their dimensions at each HSP. The time-reversal symmetry will pin $\bar{L}_4$ and $\bar{L}_5$ together. SnTe has the same number of these two irreps.

Table 2. Irreps of SnTe in the soc setting.

Irreps of the Γ	$\bar{\Gamma}_6(2)$	$\bar{\Gamma}_7(2)$	$\bar{\Gamma}_8(2)$	$\bar{\Gamma}_{10}(4)$	$\bar{\Gamma}_{11}(4)$
number	2	1	1	2	1
Irreps of the L	$\bar{L}_4(1)$	$\bar{L}_5(1)$	$\bar{L}_8(2)$	$\bar{L}_9(2)$	
number	3	3	5	2	
Irreps of the W	$\bar{W}_6(2)$	$\bar{W}_7(2)$			
number	5	5			
Irreps of the X	$\bar{X}_6(2)$	$\bar{X}_7(2)$	$\bar{X}_8(2)$	$\bar{X}_9(2)$	
number	4	3	2	1	

2.4. HSP degeneracy examination

The procedure for extracting the band irreps described earlier is applicable to both valence and conduction bands. In our algorithm, we first calculate the irreps of the bands near the top of the valence band after obtaining the wavefunctions and energy eigenvalues at the HSPs from the DFT calculations to determine if the material is an HSPSM.

A material with partially filled irreps at some HSPs is labeled as an HSPSM. In other words, the material has degeneracy between valence and conduction bands at some HSPs, and this degeneracy is protected by the HSP's symmetry, making the material a robust topological semimetal. Note that the degeneracy can be caused by space group symmetry and time-reversal symmetry.

In principle, an HSPSM is gapless at each degenerate HSP, but a small direct gap may exist because of the numerical errors of the first-principle calculations. We perform an HSPSM examination for all materials with a direct gap less than 2 meV. An energy tolerance for degeneracy starting from 0 is set starting from the N-th band (N being the number of electrons per unit cell). The lower and upper bands of the N-th band within this tolerance are found, and their irreps are calculated. If the integer number of the irreps is successfully identified, we check whether the N-th and $(N+1)$-th band belong to the same irrep. If so, the material is labeled as an HSPSM; otherwise, we proceed to the next phase of the algorithm. By contrast, if we fail to obtain the integer number of the irreps, the energy tolerance is increased to include more bands. If this tolerance reaches 2 meV, and the integer number of the irreps still cannot be identified, the material is considered to be poorly converged and should be re-calculated in the DFT phase using a higher-convergence threshold. The highest dimension of irreps in all the space groups is 8; therefore, we only need to examine at most eight bands in this step.

Note that in principle, only the information on the valence bands is required to determine the HSPSMs because the of the

top valence band is always degenerate with at least one conduction band in an HSPSM if (i) the linear decomposition of the valence bands into the irreps has a non-integer coefficient or if (ii) the coefficients of the irreps that are related to each other (*i.e.*, the type-(ii) irreps) are not equal or if (iii) the coefficient of a type-(iii) irrep is odd. These factors can be used as the criteria for determining the HSPSMs (*i.e.*, whether all the valence bands form an integer number of symmetry-enforced degenerate multiplets). However, we also include several low-conduction bands in our algorithm to obtain the degeneracy dimension.

The direct gaps for SnTe at the four HSPs in the soc setting are 4.4196, 0.0939, 1.6635, and 5.6673 eV, with the smallest one being larger than 2 meV, thereby excluding the possibility for SnTe to be an HSPSM.

Take HgTe of space group 216 as another example. It has a zero direct gap at Γ in the soc setting and has 18 electrons per unit cell. Our algorithm confirms that the 17th, 18th, 19th, and 20th bands at Γ belong to the same four-dimensional irrep $\bar{\Gamma}_8$, making HgTe an HSPSM in the soc setting.

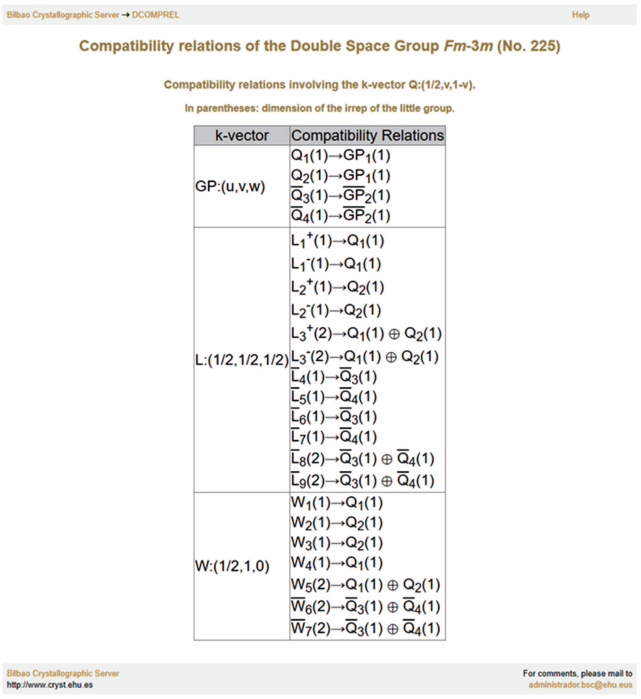
2.5. Degeneracy along the HSL examination

We continue to extract the irreps of all the valence bands and check them against the compatibility conditions on each HSL if a material has no degeneracy between the valence and conduction bands at any HSP (*i.e.*, it is not an HSPSM). The material is considered to be a band insulator and the algorithm proceeds into the next step if all the compatibility conditions are satisfied; otherwise, the material is labeled as an HSLSM, indicating that symmetry-enforced band crossings exist between the conduction and valence bands along the HSLs, where the compatibility conditions are violated.

The little group of an HSL is a subgroup of the little group of any HSP on it; this means that each irrep of the little group at an HSP can be decomposed into the direct sum of the irreps of the little group at a joining HSL. The decomposition coefficients are described by the so-called compatibility relations. For a system to be gapped along an HSL AB that connects two HSPs A and B, the valence bands of A and B must be continuous; thus, their irreps at A and B must be decomposable into the same set of irreps of AB's little group with exactly the same coefficients. We call such a condition the compatibility condition, which can be represented as a set of linear equations involving the total number of each irrep for all valence bands at A and B. The compatibility condition being satisfied along all the HSLs is a necessary condition for a system to have a direct gap. Note that if a material satisfies all the compatibility conditions, band crossings may still exist at some HSLs between the valence and conduction bands; however, the crossing point(s) are not symmetry-enforced and may be gapped by tuning the Hamiltonian without breaking the symmetry. Moreover, note that the HSL mentioned herein is not necessarily a line. It can also be a surface. The generic points in the BZ are also named as an HSL "GP" for convenience, which only gives trivial compatibility conditions (*i.e.*, conservation of the number of bands between two k points).

The compatibility relations for the 230 space groups have been explicitly derived and are available at the "*Compatibility Relations between representations of the Double Space Groups*" database on the BCS. For each space group, the BCS lists these relations between each HSP/HSL and the HSP/HSL connected to it. We use these data to construct two matrices of compatibility conditions for each HSL, that is, for the soc and nsoc settings, respectively. Element (i, j) of the matrix gives the decomposition coefficient of the j-th HSP irrep into the i-th HSL irrep, with a minus sign if the irrep belongs to the second HSP. The irrep vector of a material is generated by collecting the irreps of all valence bands at both HSPs, with each entry being the total number of occurrences of one irrep arranged in the same order matching the columns of the compatibility condition matrix. We obtain a vector, whose nonzero elements indicate the violated compatibility conditions, by multiplying the compatibility condition matrix to the irrep vector.

As an example, figure 2 illustrates one page of the compatibility relations of space group 225 on the BCS, which lists the compatibility relations between HSL $Q = (1/2, v, 1-v)$ and the three HSP/HSLs connected to it, namely $L = (1/2, 1/2, 1/2)$, $W = (1/2, 1, 0)$, and generic points GP.


Bilbao Crystallographic Server → DCOMPREL Help

Compatibility relations of the Double Space Group *Fm-3m* (No. 225)

Compatibility relations involving the k-vector Q:(1/2,v,1-v).

In parentheses: dimension of the irrep of the little group.

k-vector	Compatibility Relations
GP:(u,v,w)	$Q_1(1)\rightarrow GP_1(1)$ $Q_2(1)\rightarrow GP_1(1)$ $\bar{Q}_3(1)\rightarrow \overline{GP}_2(1)$ $\bar{Q}_4(1)\rightarrow \overline{GP}_2(1)$
L:(1/2,1/2,1/2)	$L_1^+(1)\rightarrow Q_1(1)$ $L_1^-(1)\rightarrow Q_1(1)$ $L_2^+(1)\rightarrow Q_2(1)$ $L_2^-(1)\rightarrow Q_2(1)$ $L_3^+(2)\rightarrow Q_1(1)\oplus Q_2(1)$ $L_3^-(2)\rightarrow Q_1(1)\oplus Q_2(1)$ $\bar{L}_4(1)\rightarrow \bar{Q}_3(1)$ $\bar{L}_5(1)\rightarrow \bar{Q}_4(1)$ $\bar{L}_6(1)\rightarrow \bar{Q}_3(1)$ $\bar{L}_7(1)\rightarrow \bar{Q}_4(1)$ $\bar{L}_8(2)\rightarrow \bar{Q}_3(1)\oplus \bar{Q}_4(1)$ $\bar{L}_9(2)\rightarrow \bar{Q}_3(1)\oplus \bar{Q}_4(1)$
W:(1/2,1,0)	$W_1(1)\rightarrow Q_1(1)$ $W_2(1)\rightarrow Q_2(1)$ $W_3(1)\rightarrow Q_2(1)$ $W_4(1)\rightarrow Q_1(1)$ $W_5(2)\rightarrow Q_1(1)\oplus Q_2(1)$ $\bar{W}_6(2)\rightarrow \bar{Q}_3(1)\oplus \bar{Q}_4(1)$ $\bar{W}_7(2)\rightarrow \bar{Q}_3(1)\oplus \bar{Q}_4(1)$

Bilbao Crystallographic Server http://www.cryst.ehu.es — For comments, please mail to administrador.bsc@ehu.eus

Fig. 2. Compatibility relations involving the k-vector Q (No. 225) on the BCS.

The irreps L_1^+, L_1^-, L_3^+, L_3^- and W_1, W_4, W_5 for L and W are connected to Q_1, whereas L_2^+, L_2^-, L_3^+, L_3^- and W_2, W_3, W_5 are connected to Q_2. These two relations can be transformed into two linear equations, namely

$$(L_1^+)+(L_1^-)+(L_3^+)+(L_3^-)-(W_1)-(W_4)-(W_5)=0, \quad (18)$$

$$(L_2^+)+(L_2^-)+(L_3^+)+(L_3^-)-(W_2)-(W_3)-(W_5)=0. \quad (19)$$

They can be written in a matrix form, as presented in Table 3.

Table 3. Compatibility condition matrix between L and W.

Irreps	$L_1^+(1)$	$L_1^-(1)$	$L_2^+(1)$	$L_2^-(1)$	$L_3^+(2)$	$L_3^-(2)$	$W_1(1)$	$W_2(1)$	$W_3(1)$	$W_4(1)$	$W_5(2)$
$L-Q_1-W$	1	1	0	0	1	1	−1	0	0	−1	−1
$L-Q_2-W$	0	0	1	1	1	1	0	−1	−1	0	−1

The whole compatibility condition matrix can be constructed similarly, with each row being one such linear equation where coefficients for all irrelevant irreps are set to 0.

The irreps of SnTe satisfy all the compatibility conditions in the soc setting but not in the nsoc setting. Table 4 lists the irreps of SnTe at L and W in the nsoc setting.

Table 4. Irreps of SnTe in the nsoc setting.

Irreps	$L_1^+(1)$	$L_1^-(1)$	$L_2^+(1)$	$L_2^-(1)$	$L_3^+(2)$	$L_3^-(2)$	$W_1(1)$	$W_2(1)$	$W_3(1)$	$W_4(1)$	$W_5(2)$
Number	2	0	0	2	3	0	3	2	0	1	2

They violate both the L–Q_1–W and L–Q_2–W compatibility conditions, making SnTe an HSLSM in the nsoc setting.

2.6. Symmetry-based indicator calculation

We calculate the symmetry-based indicators for a material if its band structure satisfies all the compatibility conditions of the space group. The nonzero calculated indicators are labeled as a GMSM in the nsoc setting or a TI or TCI in the soc setting. The parity of the last number in the indicators in the soc setting determines whether the material is a TI or a TCI, with an odd number corresponding to a TI and an even one to a TCI. Note that not all space groups have symmetry-based indicators (117 space groups in the soc setting and 53 in the nsoc setting). A material is labeled as a *trivial* insulator if it has zero-valued indicators or belongs to a space group with no indicator.

The group structures of the symmetry-based indicators are given in Ref. [14], and their explicit formulas in terms of the corresponding space group irreps used in our algorithm are derived in Refs. [15] and [18]. As demonstrated by Ref. [18], the nonzero indicators in the nsoc setting correspond to the topological semimetals with topological nodes at the non-high-symmetry momenta. The possible configurations of these nodes can be found in the same article. The mapping between the symmetry indicators and the possible sets of topological invariants is tabulated in Ref. [15]. The mapping is generally one-to-many, with one set of indicators mapping to several possible nonequivalent sets of topological invariants.

Note that materials with zero indicators or belonging to the space groups with no indicators are not necessarily topologically trivial. This only indicates that their band topology, if existing, cannot be detected using any symmetry eigenvalue diagnosis approach because considering only the symmetry information at the HSPs is naturally incomplete.

The symmetry-based indicator formula of space group 225 in the soc setting is[15]

$$\mathrm{Z}_8 = \frac{3}{2}n^+_{(3/2)} - \frac{3}{2}n^-_{(3/2)} - \frac{1}{2}n^+_{(1/2)} + \frac{1}{2}n^-_{(1/2)}, \tag{20}$$

where

$$n^+_{(1/2)} = n(\overline{\Gamma}_6) + n(\overline{\Gamma}_{10}) + 2n(\overline{X}_7) + n(\overline{X}_6) + 2n(\overline{L}_8) + 2n(\overline{L}_4\overline{L}_5) + n(\overline{W}_7), \tag{21}$$

$$n^-_{(1/2)} = n(\overline{\Gamma}_8) + n(\overline{\Gamma}_{11}) + 2n(\overline{X}_9) + n(\overline{X}_8) + 2n(\overline{L}_9) + 2n(\overline{L}_6\overline{L}_7) + n(\overline{W}_6), \tag{22}$$

$$n^+_{(3/2)} = n(\overline{\Gamma}_{10}) + n(\overline{\Gamma}_7) + 2n(\overline{X}_6) + n(\overline{X}_7) + 2n(\overline{L}_8) + 2n(\overline{L}_4\overline{L}_5) + n(\overline{W}_6), \tag{23}$$

$$n^-_{(3/2)} = n(\overline{\Gamma}_{11}) + n(\overline{\Gamma}_9) + 2n(\overline{X}_8) + n(\overline{X}_9) + 2n(\overline{L}_9) + 2n(\overline{L}_6\overline{L}_7) + n(\overline{W}_7). \tag{24}$$

Notice that the notation of irreps here follows the notation on the BCS, which is different from the original notation.[15] The correspondence tables of these two notations of irreps are given in Appendix D.

The SnTe indicator is calculated as $Z_8 = 4$, indicating that SnTe is a TCI, by substituting the number of irreps of the valence bands of SnTe listed in Subsection 2.3.

3. The SymTopo package

Based on the algorithm described in Section 2, we developed a package named SymTopo to automatically predict the topological property of an arbitrary nonmagnetic crystalline material. The package only requires the user to provide the structure data and the spin–orbital coupling setting of the material. It currently interfaces with two first-principle calculation software, namely VASP and ABINIT. The package is freely available with installation and usage instructions detailed at http://materiae.iphy.ac.cn/symtopo.

The package includes two main modules. The first module generates the standard input files for the structure data and the spin–orbital coupling setting provided by the user for the designated first-principle code, with which the band structure of the material can be readily calculated. After the wavefunctions and the energy eigenvalues are obtained by running the DFT calculations, the second module of the package analyzes these results and reports the material's topological properties. The implementation details of these two modules are given in the following subsections.

3.1. Input file preparation for the first-principle calculation

The input generation module standardizes the structure file provided by the user, as mentioned in Subsection 2.1, and prepares all the necessary input files for VASP or ABINIT to run the DFT calculation. Several computational material science packages, including Pymatgen,[33] AbiPy[28,34] (only for ABINIT), and Phonopy, are used by the module to accomplish the task.

The DFT calculation involves a self-consistent electron density calculation and a subsequent step to compute the wavefunctions and eigenvalues at all relevant HSPs. For VASP, the input files for these two steps are stored separately in two folders named "scf" and "wave," with each folder containing four files "POSCAR," "INCAR," "KPOINTS," and "POTCAR," which store the structure, parameter, *k* points, and pseudopotential information in a VASP-native format, respectively.

For ABINIT, both steps are defined in the same set of input files, including a ".in" file that summarizes the structure, parameter, and *k*-point information and a ".files" file that designates the name and the location of the pseudopotential files, the ".in" file, and several output files.

Table 5 presents a summary of the important default input parameters for VASP and ABINIT assumed by the input generation module. These default setup yields satisfying results in most cases. Users can also manually adjust these input files to tune the parameters for the DFT calculation when necessary. For example, in case the first-principle calculations do not converge, users could retry the calculation with a different integral method for the k points or a different convergence scheme.

Table 5. Key input parameter template for VASP and ABINIT.

Software		VASP	ABINIT
Common parameters	Cut-off energy	PREC = high[a]	ecut and pawecutdg are set according to the pseudopotential file
	Convergence criterion	NELM = 300 EDIFF = 1e-06	nstep = 300 toldfe = 1.0d − 6 eV tolwfr2 = 1.0d − 10 eV
	Soc-setting	LSORBIT = True MAGMOM = 3*N_atom*0.0	pawspnorb = 1 magconon = 1
	Nsoc-setting	LSORBIT = False MAGMOM = N_atom*0.0	pawspnorb = 0 magconon = 1 istwfk2 = num_k*1
	Other parameters	ISTART = 0 LMAXMIX = 4	iomode = 3 fband = 1
Parameters for self-consistent calculation	Occupation option	ISMEAR = −5	occopt1 = 3
	Charge density	ICHARG = 2	getden1 = 0 prtden1 = 1
	Wave function	LWAVE = False	prtwf1 = −1
	K-points	Auto 30 (in KPOINTS file)	nshiftk1 = 1 shiftk1 = 0.0 0.0 0.0 ngkpt1 = 3 3 3 kptopt1 = 1(nsoc)/4(soc)
Parameters for band calculation	Occupation option	ISMEAR = 0	occopt2 = 3
	Charge density	ICHARG = 11	getden2 = −1
	Wave function	LWAVE = TRUE	prtwf2 = 1 prteig2 = 1 enunit2 = 1
	K-points	List the coordinates of HSPs (in KPOINTS file)	nkpt2 = number of k points kpt2 = coordinates of HSPs

[a]This setting automatically sets ENCUT = 125% ENMAX, where ENMAX is read from the pseudopotential file.

The module currently works with the PAW_GGA_PBE-type pseudopotential files. Users need to configure the package with the location of the proper pseudopotential files before using the module. In using the other types of pseudopotential files, the user could use the structure and *k*-point information automatically generated by the module and manually edit the input files to fill in the correct values for the other parameters.

3.2. Topological property analysis

The analysis module of the package takes in the output files of the first-principle calculation and reports the topological classification of the material. For VASP, the output files from the band calculation, including "OUTCAR," "WAVECAR," and "EIGENVAL," are analyzed. For ABINIT, the "_WFC.nc" file that stores the wavefunction in NetCDF for-

mat is analyzed.

Aside from a classification, the module also reports relevant details for the topological class. For an HSPSM, the module lists the HSPs and the corresponding irreps, where degeneracy exists between the valence and conduction bands. For an HSLSM, the module lists the pairs of HSPs, where the band crossings between the valence and conduction bands exist at some HSLs connecting these two HSPs, and the irreps of all valence bands at each HSP. For a GMSM/TI/TCI, the symmetry-based indicators and the irreps of all valence bands are given. Lastly, only the triviality and band irreps are listed for a trivial insulator.

4. Conclusion

We have given herein a thorough description of the automated algorithm for calculating the symmetry representations and topological properties used in the sweeping scan of nonmagnetic crystalline materials that were previously reported.[19] The technical issues involved in the implementation were explained in detail. Based on the algorithm, we developed a software tool named SymTopo, which can be used to automatically generate input files for first-principle codes and analyze the output to report the topological properties. The tool is freely available at http://materiae.iphy.ac.cn/symtopo. We hope this package can help discover more topological materials in the future.

Appendix A: Brief comparison of the three topological materials search articles

The three articles,[19–21] published together in *Nature*, all performed a high-throughput calculation in search of novel topological materials. The authors of these papers adopted similar methods and achieved generally consistent results. The crystal structures of the materials they used all came from the Inorganic Crystal Structure Database (ICSD)[35] with different filtering methods. Zhang, *et al.*[19] and Vergniory, *et al.*[20] used VASP for the first-principle calculations, whereas Tang, *et al.*[21] used WIEN2K[36] and performed a modified Becke–Johnson calculation[37] for crosschecking. The small differences in the structural data and the different choice of first-principle calculation software may account for the inconsistencies between the topological material predictions in the three articles.

Zhang, *et al.* and Tang, *et al.* used symmetry-based indicator theory, whereas Vergniory, *et al.* used topological quantum chemistry to predict topological materials. Zhang, *et al.* and Vergniory, *et al.* calculated more than 2.5×10^4 nonmagnetic materials and found that approximately 30% of them were actually topological, whereas Tang, *et al.* selected several hundreds of topological materials with relatively clean Fermi surfaces. Zhang, *et al.* and Vergniory, *et al.* also calculated the indicators based on research by Song, *et al.*,[15] with which the topological invariants can be easily extracted. What is special about Zhang, *et al.* is that they also did a calculation with the nsoc setting and calculated the indicators based on research by Song, *et al.*,[18] whereas Vergniory, *et al.* calculated the topological properties of a material when the symmetry was lowered into subgroups.

Appendix B: Coordinate transformation matrix and origin point shift vectors

We list herein the similarity transformation matrices of the coordinate axes and the shift vector of the origin point between the BCS's conventional cell convention and Phonopy's primitive cell convention.

Table B1. Crystal system and Bravais lattice of the 230 space groups.

Crystal system	Bravais lattice	Space group number
Triclinic	Primitive	1, 2
Monoclinic	Primitive	3, 4, 6, 7, 10, 11, 13, 14
	Base-centered	5, 8, 9, 12, 15
Orthorhombic	Primitive	16–19, 25–34, 47–62
	Base-centered1	20, 21, 35, 36, 37, 63–68
	Base-centered2	38, 39, 40, 41
	Face-centered	22, 42, 43, 69, 70
	Body-centered	23, 24, 44, 45, 46, 71–74
Tetragonal	Primitive	75–78, 81, 83–86, 89–96, 99–106, 111–118, 123–138
	Body-centered	79, 80, 82, 87, 88, 97, 98, 107–110, 119–122, 139, 140, 141, 142
Trigonal	Primitive	146, 148, 155, 160, 161, 166, 167
Hexagonal	Primitive	143, 144, 145, 147, 149, 150–154, 156–159, 162–165, 168–194
Cubic	Primitive	195, 198, 200, 201, 205, 207, 208, 212, 213, 215, 218, 221–224
	Face-centered	196, 202, 203, 209, 210, 216, 219, 225–228
	Body-centered	197, 199, 204, 206, 211, 214, 217, 220, 229, 230

Table B2. Similarity transformation matrices of the Bravais lattices.

Bravais lattice	Similarity transformation matrix
Triclinic primitive Monoclinic primitive Orthorhombic primitive Tetragonal primitive Hexagonal primitive Cubic primitive	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Monoclinic base-centered Orthorhombic base-centered1	$\begin{pmatrix} 1/2 & -1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Orthorhombic base-centered2	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix}$
Orthorhombic body-centered Tetragonal body-centered Cubic body-centered	$\begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{pmatrix}$
Orthorhombic face-centered Cubic face-centered	$\begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$
Trigonal primitive	$\begin{pmatrix} 2/3 & 1/3 & 1/3 \\ -1/3 & 1/3 & 1/3 \\ -1/3 & -2/3 & 1/3 \end{pmatrix}$

Table B3. Shift vector of the space groups.

Space group number	Shift vector
48, 86, 126, 210	(−1/4, −1/4, −1/4)
70, 201	(−3/8, −3/8, −3/8)
85, 129, 130	(1/4, −1/4, 0)
50, 59, 125	(−1/4, −1/4, 0)
133, 134, 137	(1/4, −1/4, 1/4)
68	(1/4, −1/4, 1/4)
88	(3/8, 1/8, 1/4)
141, 142	(1/2, 1/4, 1/8)
138	(1/4, −1/4, −1/4)
222, 224	(1/4, 1/4, 1/4)
227	(1/8, 1/8, 1/8)

Appendix C: Derivation of the shift formula

We derive herein how the translation part of a symmetry operation will change when two coordinate systems differ only by a shift of the origin.

Assume that two coordinate systems O and O_1 differ by an origin shift vector $\boldsymbol{s}$ (*i.e.*, $\boldsymbol{r}_1 = \boldsymbol{r} + \boldsymbol{s}$). Assume that a symmetry operation has forms $\{\boldsymbol{R}|\boldsymbol{\tau}\}$ and $\{\boldsymbol{R}|\boldsymbol{\tau}_1\}$ in O and O_1, respectively, where $\boldsymbol{R}$ is the same rotation matrix, while $\boldsymbol{\tau}$ and $\boldsymbol{\tau_1}$ are the translation parts. We have the following four relations:

$$\boldsymbol{R}\cdot\boldsymbol{r}+\boldsymbol{\tau}=\boldsymbol{r}', \tag{C1}$$

$$\boldsymbol{R}\cdot\boldsymbol{r}_1+\boldsymbol{\tau}_1=\boldsymbol{r}_1', \tag{C2}$$

$$\boldsymbol{r}_1=\boldsymbol{r}+\boldsymbol{s}, \tag{C3}$$

$$\boldsymbol{r}_1'=\boldsymbol{r}'+\boldsymbol{s}, \tag{C4}$$

where $\boldsymbol{r}'$ and $\boldsymbol{r}_1'$ are the transformed points of $\boldsymbol{r}$ and $\boldsymbol{r}_1$ respectively. We arrive at the following shift formula by eliminating these four temporary variables:

$$\boldsymbol{\tau}_1=\boldsymbol{s}-\boldsymbol{R}\cdot\boldsymbol{s}+\boldsymbol{\tau}. \tag{C5}$$

In the opposite case, this formula can be used to calculate the shift vectors when we know that a set of symmetry operations $\{\boldsymbol{R}|\boldsymbol{\tau}\}$ and $\{\boldsymbol{R}|\boldsymbol{\tau}_1\}$ in two coordinate systems differs only by a shift of origin. The shift vector is not unique, and we only take one of them.

Appendix D: Correspondence table of the two different notations of the point group irreducible representations (irreps)

The notation of the space group irreps by Song, *et al.*[15] is different from that on the BCS.[38] It follows the notation of the point group irreps by Altmann, *et al.*[39] We list herein the correspondence tables of these two notations. In these tables, the first row corresponds to the irreps by Altmann, *et al.*, and the second row corresponds to those on the BCS. Notice that for the space group irreps, one must add the k-point label to the point group irreps when the k-point has the same point group symmetry.

Table D1. Correspondence table of two different notations of the point group irreps.

3, Ci							
A	^{1}E	^{2}E	^{1}E$_{1/2}$	^{2}E$_{1/2}$	A$_{2/3}$		
1	3	2	6	5	4		
$\bar{3}, S_6$							
A$_g$	^{1}E$_g$	^{2}E$_g$	A$_u$	^{1}E$_u$	^{2}E$_u$		
1+	3+	2+	1−	3−	2−		
^{1}E$_{1/2,g}$	^{2}E$_{1/2,g}$	A$_{3/2,g}$	^{1}E$_{1/2,u}$	^{2}E$_{1/2,u}$	A$_{3/2,u}$		
6	5	4	9	8	7		
$4/m, C_{4h}$							
A$_g$	B$_g$	^{1}E$_g$	^{2}E$_g$	A$_u$	B$_u$	^{1}E$_u$	^{2}E$_u$
1+	2+	4+	3+	1−	2−	4−	3−
^{1}E$_{1/2,g}$	^{2}E$_{1/2,g}$	^{1}E$_{3/2,g}$	^{2}E$_{3/2,g}$	^{1}E$_{1/2,u}$	^{2}E$_{1/2,u}$	^{1}E$_{3/2,u}$	^{2}E$_{3/2,u}$
6	8	5	7	10	12	9	11

Table D1. (*Continued*).

$2/m$, C_{2h}									
A_g	B_g	A_u	B_u	$^1E_{1/2,g}$	$^2E_{1/2,g}$	$^1E_{1/2,u}$	$^2E_{1/2,u}$		
1+	2+	1−	2−	4	3	6	5		
$4/mmm$, D_{4h}									
$A_{1,g}$	$A_{2,g}$	$B_{1,g}$	$B_{2,g}$	E_g	$A_{1,u}$	$A_{2,u}$			
1+	3+	2+	4+	5+	1−	3−			
$B_{1,u}$	$B_{2,u}$	E_u	$E_{1/2,g}$	$E_{3/2,g}$	$E_{1/2,u}$	$E_{3/2,u}$			
2−	4−	5−	7	6	9	8			
mmm, D_{2h}									
A_g	$B_{1,g}$	$B_{2,g}$	$B_{3,g}$	A_u	$B_{1,u}$	$B_{2,u}$	$B_{3,u}$	$E_{1/2,g}$	$E_{1/2,u}$
1+	2+	4+	3+	1−	2−	4−	3−	5	6
$\bar{1}$, C_i									
A_g	A_u	$A_{1/2,g}$	$A_{1/2,u}$						
1+	1−	3	2						
$\bar{4}$, S_4									
A	B	^1E	^2E	$^1E_{1/2}$	$^2E_{1/2}$	$^1E_{3/2}$	$^2E_{3/2}$		
1	2	4	3	6	8	5	7		
$\bar{4}2m$, D_{2d}									
A_1	A_2	B_1	B_2	E	$E_{1/2}$	$E_{3/2}$			
1	4	2	3	5	7	6			
$m\bar{3}m$, O_h									
$A_{1,g}$	$A_{2,g}$	E_g	$T_{1,g}$	$T_{2,g}$	$A_{1,u}$	$A_{2,u}$	E_u		
1+	2+	3+	4+	5+	1−	2−	3−		
$T_{1,u}$	$T_{2,u}$	$E_{1/2,g}$	$E_{5/2,g}$	$F_{3/2,g}$	$E_{1/2,u}$	$E_{5/2,u}$	$F_{3/2,u}$		
4−	5−	6	7	10	8	9	11		
$\bar{3}m$, D_{3d}									
$A_{1,g}$	$A_{2,g}$	E_g	$A_{1,u}$	$A_{2,u}$	E_u	$E_{1/2,g}$	$^1E_{3/2,g}$	$^2E_{3/2,g}$	$E_{1/2,u}$	$^1E_{3/2,u}$	$^2E_{3/2,u}$
1+	2+	3+	1−	2−	3−	8	5	4	9	7	6
$\bar{4}3m$, T_d									
A_1	A_2	E	T_1	T_2	$E_{1/2}$	$E_{5/2}$	$F_{3/2}$		
1	2	3	5	4	6	7	8		
$\bar{6}$, C_{3h}									
A'	$^1E'$	$^2E'$	A''	$^1E''$	$^2E''$	$^1E_{1/2}$	$^2E_{1/2}$	$^1E_{3/2}$	$^2E_{3/2}$	$^1E_{5/2}$	$^2E_{5/2}$
1	5	3	2	6	4	12	9	8	7	10	11
$\bar{6}2m$, D_{3h}									
A'_1	A'_2	E'	A''_1	A''_2	E''	$E_{1/2}$	$E_{3/2}$	$E_{5/2}$			
1	2	5	3	4	6	9	7	8			
$6/m$, C_{6h}									
A_g	B_g	$^1E_{1,g}$	$^2E_{1,g}$	$^1E_{2,g}$	$^2E_{2,g}$	A_u	B_u		
1+	2+	6+	4+	5+	3+	1−	2−		
$^1E_{1,u}$	$^2E_{1,u}$	$^1E_{2,u}$	$^2E_{2,u}$	$^1E_{1/2,g}$	$^2E_{1/2,g}$	$^1E_{3/2,g}$	$^2E_{3/2,g}$		
6−	4−	5−	3−	12	9	8	7		
$^1E_{5/2,g}$	$^2E_{5/2,g}$	$^1E_{1/2,u}$	$^2E_{1/2,u}$	$^1E_{3/2,u}$	$^2E_{3/2,u}$	$^1E_{5/2,u}$	$^2E_{5/2,u}$		
10	11	18	15	14	13	16	17		
$6/mmm$, D_{6h}									
$A_{1,g}$	$A_{2,g}$	$B_{1,g}$	$B_{2,g}$	$E_{1,g}$	$E_{2,g}$	$A_{1,u}$	$A_{2,u}$	$B_{1,u}$	
1+	2+	3+	4+	6+	5+	1−	2−	3−	
$B_{2,u}$	$E_{1,u}$	$E_{2,u}$	$E_{1/2,g}$	$E_{3/2,g}$	$E_{5/2,g}$	$E_{1/2,u}$	$E_{3/2,u}$	$E_{5/2,u}$	
4−	6−	5−	9	7	8	12	10	11	

References

[1] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045
[2] Qi X L and Zhang S C 2011 *Rev. Mod. Phys.* **83** 1057
[3] Chiu C K, Teo J C Y, Schnyder A P and Ryu S 2016 *Rev. Mod. Phys.* **88** 035005
[4] Armitage N P, Mele E J and Vishwanath A 2018 *Rev. Mod. Phys.* **90** 015001
[5] Bernevig B A, Hughes T L and Zhang S C 2006 *Science* **314** 1757
[6] Fu L and Kane C L 2008 *Phys. Rev. Lett.* **100** 096407
[7] Kitaev A Y 2001 *Phys. Usp.* **44** 131
[8] Zhang H, Liu C X, Qi X L, Dai X, Fang Z and Zhang S C 2009 *Nat. Phys.* **5** 438
[9] Hsieh T H, Lin H, Liu J, Duan W, Bansi A and Fu L 2012 *Nat. Commun.* **3** 982
[10] Weng H, Fang C, Fang Z, Bernevig B A and Dai X 2015 *Phys. Rev. X* **5** 011029
[11] Huang S M, Xu S Y, Belopolski I, Lee C C and Hasan M Z 2015 *Nat. Commun.* **6** 7373
[12] Wang Z J, Alexandradinata A, Cava R J and Bernevig B A 2016 *Nature* **532** 189
[13] Bradlyn B, Elcoro L, Cano J, Vergniory M G, Wang Z J, Felser C, Aroyo M I and B Andrei Bernevig 2017 *Nature* **547** 298
[14] Po H C, Vishwanath A and Watanabe H 2017 *Nat. Commun.* **8** 50
[15] Song Z D, Zhang T T, Fang Z and Fang C 2018 *Nat. Commun.* **9** 3530
[16] Khalaf E, Po H C, Vishwanath A and Watanabe H 2018 *Phys. Rev. X* **8** 031070
[17] Kruthoff J, De Boer J, Van Wezel J, Kane C L and Slager R J 2017 *Phys. Rev. X* **7** 041069
[18] Song Z D, Zhang T T and Fang C 2018 *Phys. Rev. X* **8** 031069
[19] Zhang T T, Jiang Y, Song Z D, Huang H, He Y Q and Fang Z 2019 *Nature* **566** 475
[20] Vergniory M G, Elcoro L, Felser C, Regnault N, Bernevig B A and Wang Z J 2019 *Nature* **566** 480
[21] Tang F, Po H C, Vishwanath A and Wan X G 2019 *Nature* **566** 486
[22] Kresse G and Hafner J 1993 *Phys. Rev. B* **47** 558
[23] Kresse G and Hafner J 1994 *Phys. Rev. B* **49** 14251
[24] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
[25] Kresse G and Furthmüller J 1996 *Phys. Rev. B* **54** 11169
[26] Gonze X Almbladh C O, Cucca A, Caliste D, Freysoldt C, Marques M A L, Olevanoch V, Pouillonai Y and Verstraetej M J 2008 *Comput. Mater. Sci.* **43** 0
[27] Torrent M, Jollet F, Bottin F, Zérah G and Gonze X 2008 *Comput. Mater. Sci.* **42** 337
[28] Gonze X, Jollet F, Araujo F A, Adams D and Zwanziger J W 2016 *Comput. Phys. Commun.* **205** 106
[29] Gonze X, Amadon B, Anglade P M et al 2009 *Comput. Phys. Commun.* **180** 2582
[30] Togo A and Tanaka I 2015 *Scr. Mater.* **108** 1
[31] Aroyo M I, Orobengoa D, de la Flor G, Tasci E S, Perez-Mato J M and Wondratschek H 2014 *Acta Cryst. A* **70** 126
[32] Elcoro L, Bradlyn B and Wang Z J 2017 *J. Appl. Crystallogr.* **50** 1457
[33] Ong S P, *et al.* 2013 *Comput. Mat. Sci.* **68** 314
[34] Giantomassi M, *et al.* https://github.com/abinit/abipy [2019]
[35] Hellenbrandt M 2004 *Crystallogr. Rev.* **10** 17
[36] Blaha P, Schwarz K, Madsen G K H, *et al.* 2001 WIEN2k: An augmented plane wave + local orbitals program for calculating crystal properties
[37] Tran F and Blaha P 2009 *Phys. Rev. Lett.* **102** 226401
[38] Bradley C J and Cracknell A P 2009 *The Mathematical Theory of Symmetry in Solids* (Oxford: Oxford University Press)
[39] Altmann S L and Herzig P 1994 *Point-group theory tables* (Oxford: Clarendon Press)