

PREDICTIVE MODELING OF SUPERHARD AND TOPOLOGICAL
MATERIALS BY DENSITY FUNCTIONAL THEORY AND MACHINE
LEARNING

by

WEI-CHIH CHEN

CHENG-CHIEN CHEN, COMMITTEE CHAIR
KANNATASSEN APPAVOO
ILIAS E. PERAKIS
YOGESH K. VOHRA
MARY ELLEN ZVANUT

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2021

Copyright by
Wei-Chih Chen
2021

PREDICTIVE MODELING OF SUPERHARD AND TOPOLOGICAL
MATERIALS BY DENSITY FUNCTIONAL THEORY AND MACHINE
LEARNING

WEI-CHIH CHEN

PHYSICS

ABSTRACT

In this dissertation, we focus on the computational prediction for new superhard materials and topological phase transitions in quantum materials. To search for new superhard materials, we employed machine learning techniques, evolutionary crystal structure prediction, and density functional theory to explore new superhard materials in ternary B-C-N and B-N-O compounds. For the B-C-N systems, our machine learning prediction shows that a 1:1 B-N ratio can lead to various superhard compositions. Our subsequent evolutionary algorithm searches indicated that BC_{10}N has an extremely high hardness ~ 87 GPa based on density functional theory. For B-N-O systems, we adopted an self-improvement computational procedure, and found that ternary compounds with a chemical formula $\text{B}_{x+2}\text{N}_x\text{O}_3$ (e.g. $\text{B}_5\text{N}_3\text{O}_3$, $\text{B}_6\text{N}_4\text{O}_3$, etc) are thermodynamically favorable, and several superhard B-N-O compounds can be found when $x \geq 3$. Our studies for new ternary superhard materials show that machine learning is a powerful tool to accelerate material discovery.

In our studies for topological materials, we found that ZrTe_5 is in the vicinity of topological phase transition. By means of the uniaxial tensile strain, ZrTe_5 undergoes a Z_2 transition from a strong topological insulator phase to a weak topological insulator phase. In the case of LaN , we found that LaN has a low-temperature, ferroelectric phase with $P1$ symmetry, and it would undergo a structural phase transition to rock-salt phase ($Fm\bar{3}m$ symmetry) when the temperature is increased. Besides, a hydrostatic pressure ~ 18 GPa can lead to a structural phase transition to tetragonal

structure ($P4/nmm$ symmetry). Our topological studies show that high-temperature rock-salt LaN can change from a topologically trivial phase to a topologically non-trivial phase at $P \sim 14$ GPa, and its topological phase transition can also be induced by the temperature at a pressure between 14 and 18 GPa. These results show that strain engineering and hydrostatic pressure can effectively drive topological phase transition in quantum materials.

DEDICATION

*To God, Tao, and Nature,
The absolute, omnipresent, and eternal Existence*

To my parents, for their endless love and support

ACKNOWLEDGEMENT

Throughout my PhD period, I have received many helps from many people, and this dissertation would not be possible without their efforts, time, and insights. First, wholeheartedly I appreciate my advisor, Dr. Cheng-Chien Chen, who always provides guidance and support for my cultivation of scientific literacy. Especially, I thank Dr. Chen for the detailed instructions on large-scale many-body calculations, machine learning applications, and scientific writings.

This period as an graduate student would be much more difficult without the supports of fellowships. I want to show my biggest appreciation for the support by Dr. Cheng-Chien Chen, Dr. Ilias E. Perakis, and Dr. Mary Ellen Zvanut for their recommendations for UAB Blazer Graduate Research Fellowship. Besides, I thank Dr. Cheng-Chien Chen again and Dr. Yogesh K. Vohra for their supports for CPU2AL Graduate Research Assistantships.

There are many research collaborators that I have worked with during these years. I thank Dr. Yogesh K. Vohra, Dr. Shane Aaron Catledge, Dr. Paul Baker, Kallol Chakrabarty, and Kaleb Burrage in UAB for their experiments on high hardness materials. I also thank Dr. Jiun-Haw Chu and Joshua Mutch in University of Washington for $ZrTe_5$ synthesis and measurements. I also appreciate Dr. Yao Wang in Clemson University for the collaboration and discussion on extended Hubbard model. Besides, I appreciate Dr. Joseph Maciejko in University of Alberta for the discussion of topological materials. In the end, I would like to appreciate my classmate and friend Chia-Min Lin for his company throughout my PhD period.

TABLE OF CONTENTS

	<i>Page</i>
ABSTRACT	iii
DEDICATION	v
ACKNOWLEDGEMENT	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
1 INTRODUCTION	1
Summary of Superhard Materials	1
Topological Materials	11
Overview of Remaining Chapters	20
2 METHODS	21
Density Functional Theory	21
Crystal Structure Prediction and Evolutionary Algorithm	35
3 MACHINE LEARNING AND EVOLUTIONARY DISCOVERY OF TERNARY SUPERHARD MATERIALS	41
B-C-N Superhard Materials	41
B-N-O Superhard Materials	59
4 STRUCTURAL AND TOPOLOGICAL PHASE TRANSITIONS IN TOPOLOGICAL QUANTUM MATERIALS	73

Topological Phase Transition in ZrTe ₅	73
Structrual and Topological Phase Transitions in LaN	81
5 CONCLUSION.....	97
LIST OF PUBLICATIONS	99
REFERENCES	101

LIST OF TABLES

<i>Table</i>		<i>Page</i>
1	Vickers Hardnesses of selected transition metal borides.....	5
2	Mechanical properties of newly-found superhard B-C-N compounds.....	55
3	Mechanical properties of newly-found superhard B-N-O compounds	68

LIST OF FIGURES

<i>Figure</i>	<i>Page</i>
1 Crystal structures of materials with high hardness.....	3
2 Vickers hardnesses of selected hard materials.....	4
3 Schematic plots of two topologically different band structures	16
4 Flowchart of density functional theory calculations.	34
5 Schematic plot of free energy landscape.	36
6 Examples of variation operators in evolutionary algorithm	37
7 Flowchart of evolutionary algorithm.	40
8 Computational flowchart for the B-C-N compounds.....	43
9 Data distribution of bulk and shear moduli for B-C-N systems.	48
10 Evaluation of random forests models in B-C-N study.....	50
11 B-C-N triangular plots for bulk modulus, shear modulus, and hardness...	52
12 Crystal structures, band strcutures, and phonon dispersion of BC ₁₀ N, B ₄ C ₅ N ₃ , and B ₂ C ₃ N.	53
13 The self-improvement recursive procedure used in our B-N-O study.	60
14 Data distribution and predicted cohesive energy in composition space of B-N-O systems	64
15 Machine learning prediction of cohesive energy, density, and hardness as well as corresponding model evaluations.	66
16 Crystal structures, electronic density of states, and phonon density of states of superhard B-N-O compounds.	70

17	Crystal structure of $B_5N_3O_3$ highlighting the vacancies and electron density distribution of the valence band maximum states	71
18	Crystal structure and Brillouin zone of $ZrTe_5$ as well as the universal phase diagram of topological insulators.	74
19	Band structures calculated by DFT and tight-binding model.	76
20	The evolutions of Wannier charge centers for six time-reversal invariant planes of $ZrTe_5$	78
21	Topological phase transition in $ZrTe_5$ driven by uniaxial strain.	79
22	Surface states around Γ point on the (001) surface Brillouin zone in $ZrTe_5$ for STI and WTI phases.....	80
23	Rock-salt crystal structure, Brillouin zone, and phonon dispersion of lanthanum monopnictides.	83
24	Newly found LaN in distorted rock-salt structure and its phonon dispersion. 87	
25	Total energy and relative enthalpy of LaN in different structures.	88
26	Phonon dispersions of LaN in wurtzite and zinc-blende structures	88
27	Double-well potential energy versus polarization in $P1$ -LaN and renormalized phonon dispersion of rock-salt LaN at 500 K.....	90
28	Band structures based on PBE-GGA functional and HSE hybrid functional of LaN in P_1 and $Fm\bar{3}m$ structures	91
29	Topological phase transition driven by hydrostatic pressure in rock-salt LaN	92
30	Schematic phase diagram of LaN.	94
31	Topological phase transition driven by hydrostatic pressure in rock-salt LaP, LaAs, and LaSb	95

CHAPTER 1

INTRODUCTION

Searching for materials with desired properties and controlling functional quantum materials are always central objectives of materials science and condensed matter physics. Traditional materials science research relies on experiments including sample synthesis, characterization, and optimization, which are usually time-consuming and economically expensive. With the advances of quantum-mechanical theories for the many-body problem and efficient algorithms, first-principles calculations have reached reasonable accuracy and have become important tools allowing us to design and study materials before doing experiments. Among methods used to solve the many body problem, density functional theory (DFT) can render several accurate ground state properties, and therefore has been widely used in theoretical studies. Besides, the constantly increasing computer power and the development of efficient DFT codes have facilitated high-throughput screening. The huge amount of data then make machine learning possible to accelerate the material discovery.

In this work, we first focus on searching for new ternary superhard materials including B-C-N and B-N-O compounds. Our studies include DFT calculations, machine learning methods, and crystal structure prediction by the evolutionary algorithm. Another focus of this dissertation is topological materials with quantum spin Hall effect, which are known to host exotic surface states. In particular, we have studied the topological phase transitions in zirconium pentatelluride and lanthanum monopnictides by means of strain engineering and hydrostatic pressure, respectively. In the following, we introduce superhard materials and topological materials.

Summary of Superhard Materials

Hardness is the property that quantifies a material's resistance to plastic deformation. Today, there are three main hardness measurements: scratch, indentation, and rebound. In the studies of superhard materials, the Vickers indentation is the most-used measurement to determine the hardness. The indenter in Vickers hardness test is a piece of diamond in a shape of a pyramid, which has an angle of 136° between two adjacent faces. During the measurement, a load is applied to the sample, and an indentation mark would be left on the sample. The value of Vickers hardness (H_V), with an unit of pressure, can be calculated by F/A , where F is the applied force and A is the area of indentation. In practice, one can determine the value of H_V by

$$H_V = 1854.4 \frac{F}{d^2}. \quad (1)$$

If F is in Newton, and d , the average length of two diagonals, is in μm , then H_V has an unit in GPa. It is important to note that for a piece of material, the value of H_V depends on the applied load, and H_V decreases with an increasing load in an asymptotic way. Generally, superhardness in experiments can be attributed to intrinsic effect and extrinsic effect [1]. The intrinsic effect relies on the compositional elements, strengths of chemical bonds, and underlying crystal structure (eg. sp^3 -bonded diamond and cubic boron nitride). On the other hand, extrinsic effect depends on the appearance like the surface morphology, microstructures, and grain size (e.g. aggregated diamond [2]). Here, we only focus on the intrinsic effect in high hardness materials. In the following, we introduce superhard materials as well as materials with smaller hardness but showing potential for applications in the industry.

Superhard materials are materials with a H_V greater than 40 GPa [3]. They are of great interest with applications like drilling, grinding, cutting, and protective coatings. Their extraordinary properties make them essential in manufacturing, automotive,

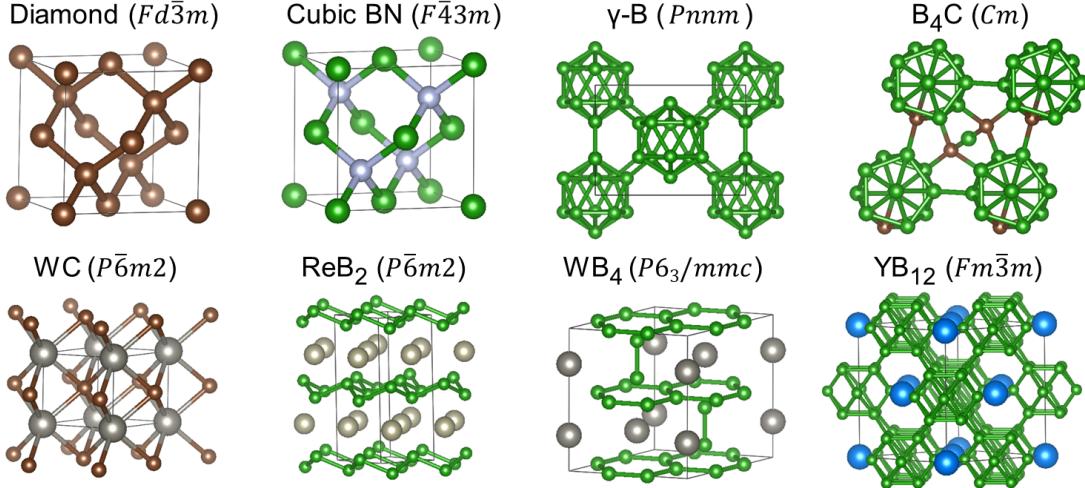


Figure 1: Crystal structures of high-hardness materials. The corresponding space groups are shown in the parentheses.

and petrochemical industries [4]. Diamond is the hardest material ($H_V \sim 90 - 100$ GPa [5]) in the world, and it also has the largest bulk modulus and shear modulus among the known materials. The superior mechanical properties of diamond are the results of its sp^3 -bonded structure and high valence electron density. However, to fulfill different industrial needs, other properties like thermal stability, chemical stability, and fracture toughness are also important factors that affect the performance of a material [6]. Diamond suffers from low oxidation temperature ~ 700 °C, and being reactive with ferrous materials [7], which limits its applications on steel and cast iron in heavy industry.

Cubic boron nitride (c-BN), with the same atomic arrangement as diamond (see Fig. 1), is the second hardest material ($H_V \sim 50 - 65$ GPa) with the advantages of high thermal stability (~ 1300 °C) [8, 9] and being inert in contact with ferrous materials, and thereby has a complementary relationship with diamond [6]. However, its synthesis, similar to diamond, usually requires high pressure and high temperature (HPHT), which takes a great amount of energy and makes it difficult for large-scale production [4]. Other compounds not as hard but widely used in the industry include cemented carbides [10, 11], boron-rich materials [12, 13, 14], SiC [15], and Al₂O₃ [16].

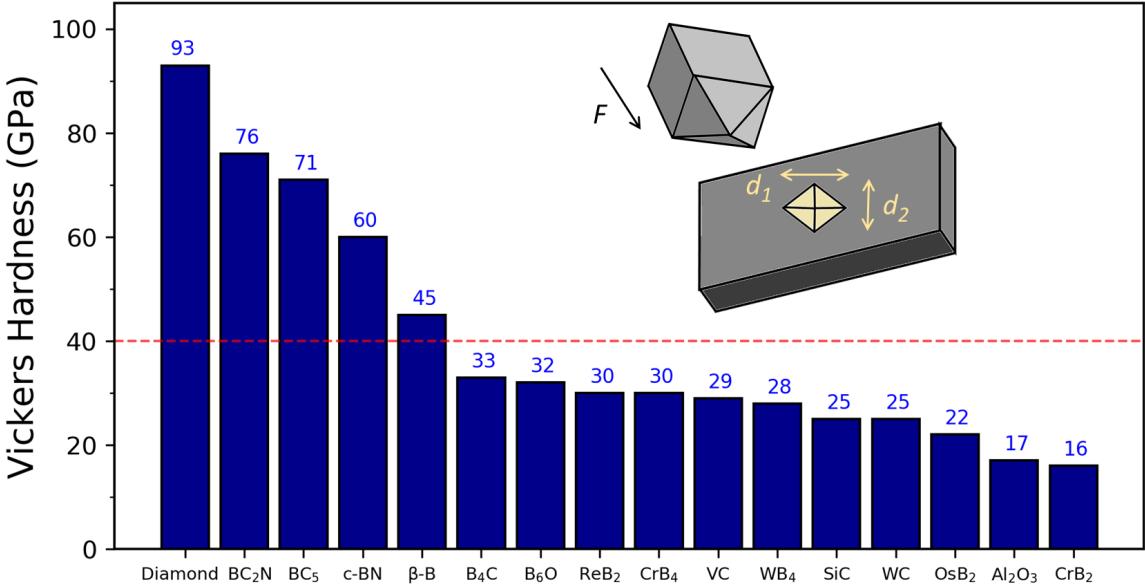


Figure 2: Vickers hardnesses of selected hard materials under a load $F \geq 4.9$ N.

These materials can be easily synthesized on a large-scale through processes like sintering, hot-pressing, and ball-milling. The crystal structures of selected high hardness materials are shown in Fig. 1, and the hardness comparison can be found in Fig. 2.

Boron and boron-rich materials with B₁₂ icosahedra as building blocks also have exceedingly high hardness. For elemental boron, there are three major allotropes worth noting: α -rhombohedral boron (α -B₁₂), β -rhombohedral boron (β -B₁₀₆), and γ -orthorhombic (γ -B₂₈), and they also have superhardness 42, 45, and 50 GPa, respectively [17]. Boron-rich materials like B₄C and B₆O have α -B₁₂-like crystal structures with carbon or oxygen atoms sit at sites between icosahedra. They were reported to have smaller hardnesses around 35 GPa [12, 13, 14]. Other boron-very-rich materials crystallized in tetragonal form like B₅₀C₂ and B₅₀N₂ have high hardnesses like B₄C [18]. Besides, a recent study by Chakrabarty *et al.* showed that the hardness of amorphous B₅₀C₂ can be as hard as c-BN [19].

Transition metal (TM) borides also have superior mechanical properties, and their preparation does not require high pressure environment. Therefore, they are ideal candidates for large-scale production. TM borides have many different crystal structures

ranging from monoborides (e.g. WB, CrB), diborides (e.g. ZrB₂, ReB₂), tetraborides (e.g. WB₄, CrB₄), dodecaborides (e.g. ZrB₁₂, YB₁₂) to higher borides (e.g. YB₅₀, YB₆₆). Monoborides like WB and CrB have relatively small hardness \sim 30 - 35 GPa under a load of 0.49 N, and their hardnesses drops to \sim 20 GPa under a load of 4.9 N [20, 21]. In TM diborides, TiB₂, ZrB₂, and HfB₂ also show relatively low hardness around 20 GPa at high load [22, 23]. It is worth noting that ReB₂ has superhardness around 45 GPa under a 0.49 N load while its hardness drops to \sim 30 GPa under a 4.9 N load. For tetraborides, WB₄ and CrB₄ are superhard with their hardnesses around 44 GPa under 0.49 N, while their high load results are also around 30 GPa. At last, dodecaborides ZrB₁₂ and YB₁₂ have slightly inferior hardnesses compared with ReB₂, WB₄ and CrB₄. Table 1 shows that the experimental results of hardness tests on selected borides are summarized in.

Table 1: Vickers hardnesses of selected transition metal borides

Borides	H_V (GPa) @ 0.49 N load	H_V (GPa) @ high load	Reference
CrB	30	20 @ 4.9 N	[21]
WB	35	18 @ 4.9 N	[20]
WB ₂	30	22 @ 4.9 N	[24]
TiB ₂	-	20 @ 5.0 N	[22]
ZrB ₂	-	17 @ 9.8 N	[23]
HfB ₂	-	20 @ 9.8 N	[23]
CrB ₂	-	16 @ 4.9 N	[25]
OsB ₂	27	18 @ 4.9 N	[26]
ReB ₂	48, 40	30, 29 @ 4.9 N	[27], [28]
WB ₄	43	28 @ 4.9 N	[29]
CrB ₄	44	30 @ 4.9 N	[25]
ZrB ₁₂	42, 40	28, 27 @ 4.9 N	[30, 31]
YB ₁₂	38	26 @ 4.9 N	[30]

With the continuous increase in the global demand of superhard materials, searching for new materials with outstanding hardness, thermal stability, and chemical stability has become an important research topic. In this field, there are two major categories. The first one focuses on hardness enhancement in TM borides by introducing one or more TM elements. This method is known as solid solution hardening, where the atomic size mismatch causes lattice distortions. Such lattice distortions effectively inhibit the dislocation, which is necessary for plastic deformation. Besides, this kind of materials are expected to have higher thermal stability due to the increase in entropy. The second one concerns finding new superhard materials composed of light elements: B, C, N, and O. These materials have short bond lengths and strong covalent bonds, which make them the primary candidates for new superhard materials. Besides, they have the advantages of being abundant in quantity and low price compared with TMs. Next, we introduce the recent developments in these two fields.

In TM bordies, forming solid solutions by introducing one or more TM elements or mixing different borides have been recognized as an effective way to enhance the hardness. In the literature, such experiments were mostly carried out by the research group led by Richard B. Kaner. For example, several experiments have been carried out based on WB_4 , which has a pristine hardness of 43 GPa under 0.49 N. Here, we briefly summarize their findings. Mohammadi *et al.* found that with the addition of 1 at.% rhenium in WB_4 , its hardness increases to 50 GPa at 0.49 N [29]. Besides, when 3 at.% molybdenum is added to WB_4 , its hardness also increases to 50 GPa at 0.49 N [32]. Besides, Akopov *et al.* also carried out a series of experiments by considering three different elements: Ti, Zr, and Hf. Their hardness indentation results show that WB_4 alloys with 8 Ti at.%, 8 at. Zr%, and 6 at. Hf% reach highest hardness values of 51, 56, and 52 GPa at 0.49 N, respectively [33], while they attributed the observed hardness enhancements to extrinsic effect due to the changes in grain morphology. Akopov *et al.* also reported hardness enhancements in WB_4 by introducing different

TMs and rare-earth elements [34]. In their experiments, seven elements (Y, Sc, Gd, Dy, Tb, Ho, and Er) were considered, and their optimum hardness values can reach as high as 46 - 50 GPa at 0.49 N.

In another two studies, Akopov *et al.* also investigated the hardnesses of the mixtures of ZrB_{12} , YB_{12} , ScB_{12} , and UB_{12} [30]. They found that the highest hardness values can be achieved at 1:1 ratio: $Zr_{0.5}Y_{0.5}B_{12}$, $Zr_{0.5}Sc_{0.5}B_{12}$, and $Y_{0.5}Sc_{0.5}B_{12}$, $Zr_{0.5}U_{0.5}B_{12}$ have their hardnesses around 43 - 48 GPa at 0.49 N, which are higher than those of pure diborides (28 - 42 GPa). For systems with lower boron content, tungsten monoboride, WB, has also been studied in the similar way, Yeung *et al.* found $W_{0.5}Ta_{0.5}B$ becomes superhard with a hardness of 43 GPa at 0.49 N, which is a 23 % enhancement compared with WB's 35 GPa [20]. Pangilinan *et al.* studied the solid solution of WB_2 by adding Ta and Nb elements [24]. Their results show that when 6 at.% Nb and 8 at.% Ta are introduced in WB_2 , the values of hardness become greater than 40 GPa at 0.49 N. The hardness enhancements exceed 30 % compared with the hardness of WB_2 (30 GPa at 0.49 N). In the literature, people have also considered mixing icosahedron-based boride with TM diboride. Mnatsakanyan *et al.* synthesized a lightweight superhard B_4C -27wt.% ReB_2 ceramic composite, which has a superhardness of 50 GPa under a load as high as 49 N, which is comparable with cubic boron nitride. [35].

High-entropy materials are solid solutions composed of five or more elements with (nearly) even proportions. Recently, they have attracted increasing attention because these materials are promising to be thermodynamically more stable at high temperatures, and thus have potential for high-temperature applications. In the literature, researchers have found that high-entropy borides and carbides have higher values of hardness than those of simple diborides. Gild *et al.* fabricated a series of single-phase high-entropy TM diborides ($Hf_{0.2}Zr_{0.2}Ta_{0.2}M_{0.2}Ti_{0.2}B_2$; M = Nb, Mo, Cr) [36]. Their results show that high-entropy TM diborides have higher hardness (22 GPa) and im-

proved oxidation resistance than individual diborides. Another similar study carried out by Zhang *et al.* reported that $\text{Hf}_{0.2}\text{Zr}_{0.2}\text{Ta}_{0.2}\text{Cr}_{0.2}\text{Ti}_{0.2}\text{B}_2$ has a hardness of 28 GPa at 1.96 N [37].

Different from borides, Sarker *et al.* synthesized single-phase high-entropy TM carbides with rock-salt structure (like β -WC). Their samples consist of five out of eight TMs (Hf, Nb, Mo, Ta, Ti, V, W, and Zr) and carbon [38]. Their results show that high-entropy carbides can enhance hardness up to 50 % higher than the rule of mixtures. Besides, they also proposed the “entropy forming ability” as an indicator for synthesizability from first-principles calculations. In the following, we introduce the second route to find new superhard materials: ternary compounds that composed of light elements: B, C, N, and O.

Similar to the studies of TM borides, mixing two compounds that composed of B, C, and N in each have been reported in the literature. In the case of elemental carbon and boron, Solozhenko *et al.* synthesized diamond-like BC_5 under HPHT, and found that the synthesized BC_5 has a hardness of 71 GPa [39]. They also reported that BC_5 has relatively high fracture toughness, and exceptional thermal stability (1900 K). Recently, Baker *et al.* used microwave plasma chemical vapor deposition (MPCVD) to synthesize boron-incorporated diamond at low temperature and low pressure, and they found the hardness can reach as high as 62 GPa in their cubic phase sample with a boron content of 7.7 at% [40].

In the literature, ternary B-C-N compounds with superhardness have also been investigated by researchers. Cubic BC_2N , with 1:1 composition ratio of carbon and boron nitride, was reported to have superhardness of 76 or 62 GPa [41, 42, 43]. Not only harder than c-BN, experimental results also show that BC_2N remains stable up to 1800 K, which demonstrates its superior thermal stability than diamond. Other possibilities like BCN and BC_9N in cubic phase have also been synthesized under extreme conditions [44].

Ternary B-C-O materials have also been synthesized in the experiments. Garvie *et al.* used HPHT techniques to synthesized B-C-O compounds, and their results show solid solutions composed of B_4C and B_6O . In particular, $B_6C_{1.1}O_{0.33}$ and $B_6C_{1.28}O_{0.31}$ with α - B_{12} structure were grown [45]. Similarly, Bolotina *et al.* also reported synthesized $BC_{0.155}O_{0.155}$ under HPHT, which also has a structure like α - B_{12} [46]. Researchers have also used computaitonal crystal structure prediction to search for superhard B-C-O compounds. B_2CO were studied independently by two research groups while different structures were discovered [47, 48]. Both of them reported that B_2CO is possible to be superhard based on their density functional theory (DFT) calculations. Another theoretical study used particle swarm algorithm to search for new possibilities in B_2C_xO systems [49]. Their results show that higher carbon content increases the number of sp^3 bonds between carbon atoms and thereby enhances the mechanical properties. They reported that B_2C_3O has a hardness of 51 - 68 GPa. Wang *et al.* used variable-composition evolutionary searches to study the B-C-O systems [50]. They reported several metastable structures including B_4CO_4 , $B_6C_2O_5$, and B_2CO_2 , and their simulated hardnesses are around 35-40 GPa suggesting that they might be superhard materials. Among them, B_4CO_4 is energetically favorable, and it becomes thermodynamically stable at 23 GPa. Hence, these B-C-O compounds have potential to for future industrial applications.

As for ternary C-N-O compounds, very limited results could be found in the literature. Steele *et al.* used evolutionary algorithm to study C-N-O compounds, and they found that C_2N_2O can be stabilized at 10 GPa [51] with a high bulk modulus of 232.5 GPa. Besides, another high-pressure phase, $C_3N_2O_3$, also has a high bulk modulus of 210.5 GPa, and it also becomes thermodynamically stable at 17.3 GPa,

Ternary B-N-O compounds are also not fully-explored in the literature. Li *et al.* used particle swarm algorithm for hardness-driven search, where hardness is the fitness function to be maximized [52]. They reported that orthorhombic B_3NO is

superhard with a hardness greater than 45 GPa. Experimentally, Bhat *et al.* first synthesized single-crystalline B-N-O compounds by mixing hexagonal boron nitride and boron trioxide under HPHT [53]. They reported the synthesized $B_6N_4O_3$ has a zinc-blende-like structure, and their DFT simulations also indicate that it has a high bulk modulus ~ 300 GPa, which is higher than any other known oxynitride.

As discussed above, several ternary compounds composed of B, C, N, and O were studied by crystal structure prediction techniques with density functional theory. However, to explore the enormous phase space in multi-component systems needs tremendous computational time to finish. Besides, traditional methods are also limited by the number of atoms in the unit-cell. Thanks to the developments of machine learning methods, new descriptors for materials, and online materials databases, now we can take advantage of machine learning and existing data to build machine learning models, which in turn can help us predict unknown properties of a material and accelerate the discovery of new materials. In this dissertation, we use two different machine learning study procedures to search for superhard B-C-N and B-N-O materials, which will be discussed in Chapter 3, “Machine Learning and Evolutionary Discovery of Ternary Superhard Materials”.

Topological Materials

Topology in mathematics concerns the properties that are invariants under smooth deformation. In the context of condensed matter physics, topology of a material becomes a new kind of phase classification, which is different from traditional phase classification by Landau’s theory. Therefore, in topological materials, the discrete “topological invariants”, instead of the continuous “order parameter”, is used to label a topological phase.

An example of topological materials is the integer quantum Hall (IQH) effect, which was first discovered by Klaus von Klitzing *et al.* in 1980 [54]. In IQH system, the behaviors of longitudinal and transverse electrical conductivities against external magnetic field cannot be understood by symmetry-breaking. Instead, the topology of electronic structure is the underlying physics behind such quantum phase transitions. In the following, we first introduce the experimental settings and outcomes of IQH effect and how topology comes into play.

The IQH effect was initially observed in silicon-based metal-oxide-semiconductor field-effect transistor (MOSFET), which can be regarded as a two-dimensional electron gas system. When the system is subjected to low temperature and high magnetic field, the electronic structure becomes the discrete Landau levels. As the Fermi level is in the gap between two Landau levels, the longitudinal conductivity σ_{xx} disappears and the transverse conductivity (or Hall conductivity) is quantized,

$$\sigma_{xy} = -\frac{e^2}{2\pi\hbar}C, \quad (2)$$

where C is an integer representing filled Landau levels. This result indicates that the system becomes insulating in the bulk while the edge can conduct electricity. Klaus von Klitzing was latter awarded the Nobel prize in 1985. The precision level of the quantization of σ_{xy} has been found to be one part of a billion [55], and it does

not depend on intricate details like geometry and purity of the sample. Therefore, the quantized Hall conductivity in the IQH effect is a macroscopic phenomenon that manifests its topological property. The mathematical form of the integer C are first obtained by Thouless, Kohmoto, Nightingale, and den Nijs (TKNN) in 1982, where they considered a 2D non-interacting electronic system with periodic potential and an uniform magnetic field [56]. Using the fundamental Kubo formula for electrical conductivity, TKNN found the Hall conductivity has a form

$$\sigma_{xy} = \frac{-e^2}{(2\pi)^2 \hbar} \sum_n^{occ} \int_{BZ} d^2k \left(i \left\langle \frac{\partial u_{n\mathbf{k}}}{\partial k_1} \middle| \frac{\partial u_{n\mathbf{k}}}{\partial k_2} \right\rangle - \left\langle \frac{\partial u_{n\mathbf{k}}}{\partial k_2} \middle| \frac{\partial u_{n\mathbf{k}}}{\partial k_1} \right\rangle \right), \quad (3)$$

where $u_{n\mathbf{k}}$ stands for the periodic part in Bloch wavefunction. Later, Michael Berry found the geometric phase in quantum mechanics [57]. (But this paper was published late in 1984 because one referee lost the manuscript [58]) In 1983, Barry Simon found the integral in TKNN expression is actually Berry curvature [59]. Therefore, TKNN expression can be written as:

$$\sigma_{xy} = \frac{-e^2}{(2\pi)^2 \hbar} \sum_n^{occ} \int_{BZ} d^2k \Omega_{n,z}(\mathbf{k}), \quad (4)$$

where $\Omega_{n,z}$ is the Berry curvature. Chern-Gauss-Bonnet theorem tells us that the integral of the Berry curvature over any closed 2D manifold is quantized to be 2π times an integer. Therefore, we know each band has its own C_n

$$\int_{BZ} \Omega_{n,z}(\mathbf{k}) d^2k = 2\pi C_n. \quad (5)$$

In the end, the Hall conductivity is

$$\sigma_{xy} = \frac{-e^2}{(2\pi)\hbar} \sum_n^{occ} C_n = \frac{-e^2}{2\pi\hbar} C, \quad (6)$$

where C is called the first Chern number or TKNN number. Such classification is known as a Z classification since C runs over the group of all integers. The first Chern number is the topological invariant corresponding to the topological phase in IQH systems. Since topological invariants only take discrete integer numbers, which manifests the robustness of topological phases against small perturbations. Besides, the bulk-edge correspondence tells us that at the interfaces between two topologically different materials, there exist gapless edge states across the bulk band gap. Such edge states are known as chiral, where back scattering is forbidden, and such an exotic property makes them promising for the realization of dissipationless electronics.

With the discovery of the IQH effect, people would like to know if there is similar quantum versions for anomalous Hall effect. Based on the result of Eq. (4), the key ingredient lies in non-trivial Berry curvature. In 1988, Haldane studied a model with a honeycomb lattice and found the quantized Hall conductivity can be realized without a net magnetic field [60]. This model is a tight-binding model with two different site, and the parameters include the on-site energy difference, real nearest-neighbor hopping and complex next-nearest-neighbor hopping. By tuning the on-site energy difference and complex hopping amplitude, one can have a non-zero Chern number. Therefore, Haldane model demonstrate a theoretical realization of Quantum anomalous Hall (QAH) effect. The experimental observation of QAH effect was not found until 2013 in a magnetically-doped Z_2 topological insulator $\text{Cr}_{0.15}(\text{Bi}_{0.1}\text{Sb}_{0.9})_{1.85}\text{Te}_3$ [61]. The topological insulator is a non-trivial phase in quantum spin Hall (QSH) effect, and in $\text{Cr}_{0.15}(\text{Bi}_{0.1}\text{Sb}_{0.9})_{1.85}\text{Te}_3$, the role of chromium is to break time-reversal (TR) symmetry, which results in a non-zero Chern number. To understand topological insulator, we next introduce the Z_2 classification in the quantum spin Hall (QSH) effect.

In a 2D system with strong spin-orbit coupling, electron currents with different spin directions that driven by applied electric filed would drift towards different di-

rections. This phenomenon is known as the spin Hall effect. In such systems, the TR-symmetry is preserved, and hence along the perpendicular direction of electric field, there is no electric current but spin current. The quantum version of spin Hall effect was first reported theoretically by two independent studies, Kane and Mele in 2005 [62], as well as Zhang and Bernevi in 2006 [63]. In Kane and Mele's paper, they considered two copies of Haldane model with different spin and only nearest-neighbor hopping. They also added a spin-orbit coupling term that plays the role of the complex next nearest-neighbor hopping in Haldane model. This model is known as Kane-Mele model, and this model is invariant under TR symmetry. It turns out, the TR symmetry is able to render a new topological phase called QSH phase, and its topological invariant is Z_2 invariant (0 or 1). The non-trivial Z_2 insulator (with $Z_2 = 1$) in 2D system has the so-called TR-symmetry protected helical edge state, which resembles two copies of chiral states flowing in the opposite directions. Moreover, the electrons in edge states have spin locked at a right-angle to their momentum, and they are also free from back scattering.

In Kane-Mele model, graphene is proposed to be a candidate to host QSH effect. However, in reality the spin-orbit coupling in carbon atom is extremely small and such spin-orbit coupling (SOC) gap cannot be observed experimentally. In 2006, Zhang, Hughes, and Bernevig proposed a theoretical model (BHZ model), and predicted a 2D QSH system could be realized in HgTe/CdTe quantum well [64]. In their paper, the thickness of HgTe layer plays an important role that controls the band inversion and hence topological phase transition. One year later, the QSH phase in HgTe/CdTe quantum was observed experimentally by Molenkamp' group [65].

Here we briefly discuss the mathematical formalism to identify the Z_2 invariant in QSH systems. The TR symmetry is the premise for the QSH effect, and it is

represented by an antiunitary operator:

$$\Theta = \exp(i\pi S_y/\hbar)K, \quad (7)$$

where S_y is the spin operator and K is the complex conjugation. For fermionic system, we have $\Theta^2 = -1$. Based on Kramber's theorem, in such systems, all eigenstates must be at least doubly degenerate. A k-space Hamiltonian with the TR-symmetry satisfies

$$\Theta \mathcal{H}(\mathbf{k}) \Theta^{-1} = \mathcal{H}(-\mathbf{k}), \quad (8)$$

In such systems, the Chern number is always 0, but we can use another topological invariant for classification, which is called the Z_2 invariant ν , which can be either 0 or 1. Here, we would like to explain it through the bulk-edge correspondence.

To illustrate the edge states, we need to make a cut on a 2D system in order to obtain a 1D ribbon. The schematic band structures of 1D ribbons are shown in Fig. 3. The edge states at time-reversal invariant momenta (TRIM), 0 and π , are doubly degenerate. This constraint leads to two different ways to connect the edge states (blue lines in Fig. 3). In the case of trivial insulators, the connection is pairwise, and the number of edge states crossing the Fermi level in half of the Brillouin zone is even. Although a conducting edge is allowed to exist (like in Fig. 3(a) left), it might become insulating through perturbation as shown on the right of Fig. 3 (a). In contrast, the edge states in QSH insulators would always cross the bulk band gap, and we say the (helical) edge states are protected by TR-symmetry [66].

There are several approaches to calculate the Z_2 invariant. The most intuitive way is to exploit bulk-boundary correspondence by calculating the edge states dispersion. This method can be directly implemented by means of standard density functional theory (DFT) calculations. Here is an easy example to understand Z_2 invariants by the bulk property. If a system with conserved S_z , up and down spins have their

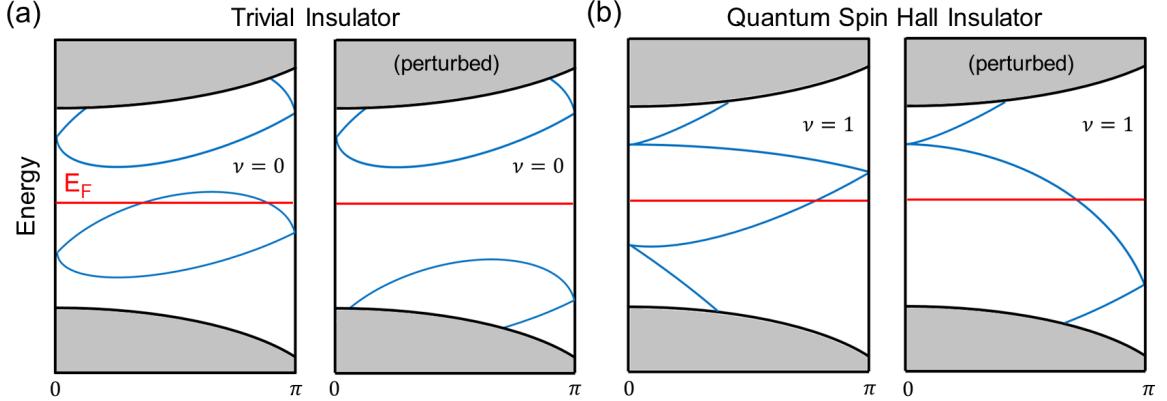


Figure 3: Schematic band structure of 1D ribbon cutting from a 2D (a) trivial insulator ($\nu = 0$) and (b) quantum spin Hall insulator ($\nu = 1$), where the blue lines indicate the edge states.

individual Chern numbers n_\uparrow and n_\downarrow . Due to TR-symmetry, $n_\uparrow + n_\downarrow = 0$. The spin Hall conductivity is defined by $n_\sigma = (n_\uparrow - n_\downarrow)/2$, and the Z_2 invariant can be calculated simply by $Z_2 = n_\sigma \bmod 2$. However, in real systems, S_z is not conserved due to the presence of spin-orbit coupling, so this is not a generic way to calculate Z_2 invariants.

One approach proposed by Fu and Kane [67] is to first define an unitary matrix $w_{ij}(\mathbf{k}) = \langle u_i(\mathbf{k}) | \Theta | u_j(-\mathbf{k}) \rangle$, which has the property, $w^T(\mathbf{k}) = -w(-\mathbf{k})$, due to anti-unitary TR-operator Θ . At four TRIM, \mathbf{k} and $-\mathbf{k}$ are the same, and hence $w(\Lambda_a)$ is anti-symmetric. The Z_2 invariant can be calculated through the following formula when $|u_i(\mathbf{k})\rangle$ is chosen to be continuous throughout the BZ:

$$(-1)^\nu = \prod_{a=1}^4 \frac{\text{Pf}(w(\Lambda_a))}{\sqrt{\text{Det}(w(\Lambda_a))}}. \quad (9)$$

If the crystal structure has inversion symmetry (or centrosymmetry), the Eq. (12) can be simplified by examining the parities for each occupied Kramer pairs at TRIM,

$$(-1)^\nu = \prod_{a=1}^4 \prod_n^{N_{occ}/2} \xi_n(\Lambda_a), \quad (10)$$

where $\xi(\Lambda_a)$ is the parity of wavefunction at Λ_a [68]. Fu and Kane also reported another method [67] to calculate Z_2 invariant by inspecting the Berry connection and curvature in the half Brillouin zone (BZ):

$$\nu = \frac{1}{2\pi} \left[\oint_{\partial B} \mathbf{A} \cdot d\mathbf{l} - \int_B \Omega d^2k \right] \quad \text{mod. 2} \quad (11)$$

with the constraint that the gauge used to obtain the Berry connection \mathbf{A} must respect TR symmetry. For non-trivial case, this is known as the obstruction to the Stokes theorem. In other words, if there exists a smooth gauge in the half BZ, then the Stokes theorem holds, and the system is topologically trivial. This method is also applied by Fukui and Hatsugai in their discretized method to check the number of nontrivial plaquettes [69].

Now, we discuss the situation in three dimension. In quantum anomalous Hall (QAH) effect, the 3D version is associated with stacked 2D QAH layers, and we can use three Chern numbers (C_1, C_2, C_3) to characterize such systems, where C_1, C_2 , and C_3 are corresponding to $(k_2, k_3), (k_1, k_3)$, and (k_1, k_2) planes, respectively. The Hall conductivity in 3D QAH insulators is no longer quantized and is proportional to the density of stacked 2D layers. For example, for a $(0, 0, C_3)$ QAH insulator with a lattice $\mathbf{a}_1, \mathbf{a}_2$, and \mathbf{a}_3 , the Hall conductivity is $\sigma_{xy} = C_3(e^2/2\pi\hbar)(1/|\mathbf{a}_3|)$ [70].

For QSH effect in 3D, the situation is different from QAH effect, and it was studied independently Fu, Kane and Mele [71], as well as Morre and Balents [72], as well as Roy [73]. Moore and Balents gave the non-trivial ones a name, “topological insulator” (TI), and they can be further divided into strong TI insulator and weak TI. In particular, the strong TI has special surface states that distinguished from 3D QAH systems. The QSH effect in 2D, a single Z_2 invariant is used to identify the band topology of a single plane BZ associated with four TRIM. In 3D, we have eight TRIM, and thus there are many choices to get a plane associated with four TRIM.

Each plane is an individual 2D QSH system related to a Z_2 value, and it seems we can choose six planes, and thus six Z_2 invariants to describe a 3D. It turns out, these Z_2 values are not independent with each other. As a matter of fact, it takes only four Z_2 values to fully characterize a 3D QSH system. The Z_2 topological indices are mostly written by $(\nu_0; \nu_1 \nu_2 \nu_3)$, where ν_0 is the strong index, and the other three are the weak indices. Similar to Eq. (12) and (10), the topological indices can be calculated through

$$\begin{aligned} (-1)^{\nu_0} &= \prod_{a=1}^8 \delta_a, \\ (-1)^{\nu_{i=1,2,3}} &= \prod_{a=1}^4 \delta_a. \end{aligned} \tag{12}$$

where $\delta_a = \text{Pf}(w(\Lambda_a)) / \sqrt{\text{Det}(w(\Lambda_a))}$. As described above, the strong index involves all eight TRIM while weak indices only associate with four TRIM on plane $k_{i=1,2,3} = \pi$. In a 3D TR-invariant system, an insulator with $\nu_0 = 1$ is called a strong TI. For those with $\nu_0 = 0$, one of the $\nu_{i=1,2,3} = 1$ is called a weak TI. The one with all four topological indices being 0 is a trivial insulator.

The Z_2 indices characterize the number of intersections between surface states and Fermi level being odd or even. If the number of intersection is odd, the surface states are protected by the TR-symmetry. In a strong TI, the surface Fermi arc encloses an odd number of Dirac points, and this ensures a topologically protected surface Dirac cone on any surface of the system. Besides, the surface states on Dirac cone has a distinct property known as spin-momentum locking – the spin degree of freedom is locked to its momentum. We call it “strong” because these surface properties are robust against small perturbation like small disorder as long as TR-symmetry is present.

On the other hand, the weak TI phase can be regarded as stacked 2D QSH layers along $\mathbf{G} = \nu_1 \mathbf{b}_1 + \nu_2 \mathbf{b}_2 + \nu_3 \mathbf{b}_3$ “module 2” reciprocal lattice vector. On the side surfaces of a stacked QSH layers, the 2D topological surface states are formed due to the

collection of 1D edge states by each QSH layer. However, the top and bottom surfaces are simply trivial surfaces. The term “weak” here is related to the translational symmetry along \mathbf{G} , and we say that surface states in weak TI are not robust against perturbations that break translational symmetry (like dislocation). However, this sensitivity renders us some interesting phenomena, such as anisotropic edge modes, quantized charge polarization, and bound states at dislocation defects.

The topological edge/surface states in QSH effect have great potential to facilitate the advance in science and technology. The followings are some of the possibilities. In magnetic topological insulator or QAH insulator, the back-scattering-immune property is promising for dissipationless topological electronics. The spin-momentum-locked properties in topological insulators can be used for spintronic devices. The junction of a topological insulator and an s-wave superconductor is promising for the realization of Majorana particles, and this can help the realization of fault-tolerant quantum computing. As a result, it is important to know how to manipulate the topological surface states by controlling the transitions between topological phases.

Overview of Remaining Chapters

In the following, we describe the methods employed in this dissertation including density functional theory, evolutionary algorithm and machine learning. The results contain our studies related to computational search for new ternary superhard materials and topological phase transitions in quantum materials. In our B-C-N study, we exploited existing online database to build machine learning models for predicting mechanical properties for the B-C-N system. Guided by machine learning prediction, we then used evolutionary algorithm to search for superhard B-C-N compounds. For the B-N-O study, we adopted a recursive, self-improving method based on the data generated by evolutionary algorithm to find ternary B-N-O compounds with high cohesive energy and high hardness.

For the topological phase transition research, we have studied the transition in ZrTe_5 between strong TI phase to weak TI phase driven by uniaxial strain. Besides, we have also studied the transition between strong TI phase and trivial phase in lanthanum monopnictides. In particular, we found LaN has a low-temperature, ferroelectric structure, which enriches both its structural and topological phase transitions.

CHAPTER 2

METHODS

Density Functional Theory

One of the primary goals in solid state physics and quantum chemistry is to determine the electronic structure because it arguably determines all of the physical properties of a material. In principle, the electronic structure of a system can be obtained by solving the many-body Shrödinger equation. However, only small systems that containing very small number of electrons like helium atom, hydrogen molecule, and lithium atom can be accurately solved using the variational method. The main difficulty in solving electronic structure of many-electron systems lies in the lack of proper treatment of electron-electron interactions. Therefore, solving the electronic structure in an accurate and efficient way is still challenging today.

The Hartree-Fock (HF) theory is the first wavefunction-based approach that captures the fundamental ingredients to solve the many-body Shrödinger equation using a mean-field approximation for electron-electron interaction and also using the Slater determinant to be the approximate many-body wavefunction [74, 75, 76]. However, the electron-electron interaction in the HF method is not precise enough for it to obtain accurate many-body wavefunction. Despite of the inaccuracy, the HF theory laid a good foundation for its successors known as post-HF approaches including Møller–Plesset perturbation theory, Coupled cluster, configuration interaction, etc. The post-HF methods show higher accuracy, and are widely used in modern quantum chemistry calculations. However, such calculations require huge computational

time.

Different from methods focusing on solving many-body wavefunction, another category – density-based method, aims at solving electron density distribution to obtain a system's physical properties. The pioneer of density-based method is the Thomas-Fermi (TF) model , which was independently proposed by Thomas and Fermi [77, 78]. The ideas of using electron density (instead of wavefunction) as the central variable and local density approximation (LDA) to compute the total energy. The LDA basically says that the properties of an inhomogeneous electron gas is locally identical to those of homogeneous one. According to the TF model, the total kinetic energy of the electrons can be expressed in terms of electron density distribution $\rho(\mathbf{r})$ without any wavefunction information. In the TF model, the total energy functional is

$$E_{TF}[\rho(\mathbf{r})] = T_{TF}[\rho(\mathbf{r})] + \int \rho(\mathbf{r})v_{ext}(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (13)$$

where ρ is electron density, v_{ext} is external potential

$$v_{ext} = - \sum_i \int \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} d\mathbf{r}, \quad (14)$$

and T_{TF} is kinetic energy functional, which can be expressed as

$$T_{TF}[\rho(\mathbf{r})] = t_0 \int \rho^{5/3}(\mathbf{r})d\mathbf{r}, \quad (15)$$

where t_0 is a coefficient, $\frac{3}{10}(3\pi^2)^{2/3}$ in atomic unit. The last term in Eq. (13) is the classical electrostatic energy between electrons, which is also known as Hartree energy. In quantum mechanics, one needs to compute the expected value of kinetic energy through the many-body wavefunction, which has $3N$ variables and difficult to solve. Here, the TF model demonstrates that density-based approaches can dramatically reduce the number of variables from $3N$ to 3. However, the TF model is just regarded

as a toy model due to imprecise expression for kinetic energy and omission of quantum mechanical electron-electron interactions. In 1928, Dirac added an exchange-energy functional term to refine the TF model [79], but it did not eliminate the primary errors from crude kinetic energy and electron correlation energy. Hence, density-based methods were not considered to be practical for real materials. Solving a quantum-mechanical many-body problem using only three spacial variables was regarded to be impossible for more than 35 years. Until 1964, the theoretically exact density functional theory was proposed by Hohenberg and Kohn, and today it has become the most popular method to study many-electron systems like solids and molecules.

Density functional theory (DFT) is a quantum-mechanical theory that can be derived from the many-body Schrödinger equation. In the following, we introduce the formalism of DFT from many-body Hamiltonian, Born-Oppenheimer approximation, Hohenberg-Kohn theorems, self-consistent Kohn-Sham equation, and exchange-correlation energy functional used in this dissertation. Besides, we also introduce practical implementation of pseudopotential techniques that largely reduce the computation time.

In a system composed of electrons and nuclei, the behavior of such a system is governed by its total many-body Hamiltonian $\hat{\mathcal{H}}$. It includes nuclear part $\hat{\mathcal{H}}_n$ and electronic part $\hat{\mathcal{H}}_e$. Here, we put the nucleus-electron interaction in $\hat{\mathcal{H}}_e$. Their mathematical forms in atomic units ($\hbar=1$, $m_e=1$, $e=1$) are shown as follows.

$$\begin{aligned}\hat{\mathcal{H}} &= \hat{\mathcal{H}}_n + \hat{\mathcal{H}}_e \\ \hat{\mathcal{H}}_n &= -\frac{1}{2} \sum_N \frac{1}{M_N} \nabla^2 + \sum_{N < M} \frac{Z_N Z_M}{|\mathbf{R}_N - \mathbf{R}_M|} \\ \hat{\mathcal{H}}_e &= -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{Ni} \frac{Z_N}{|\mathbf{R}_N - \mathbf{r}_i|}\end{aligned}\quad (16)$$

One can obtain all of the information of the system by solving the time-independent

Schrödinger equation

$$\hat{\mathcal{H}}\Theta(\{\mathbf{R}\}, \{\mathbf{r}\}) = E_{tot}\Theta(\{\mathbf{R}\}, \{\mathbf{r}\}). \quad (17)$$

However, this is the primary difficulty in electronic structure problem and unsolvable for many-electron systems. The next step to approach the solution is Born-Oppenheimer (BO) approximation, where it assumes the wavefunctions of nuclei and electrons in a system can be separated because the motions of nuclei are much faster than that of electrons based on their mass difference ($M_n \gg m_e$). The mathematical form of BO approximation is expressed as

$$\Theta(\{\mathbf{R}\}, \{\mathbf{r}\}) = \Phi(\{\mathbf{R}\})\Psi(\{\mathbf{r}\}), \quad (18)$$

, where $\Phi(\{\mathbf{R}\})$ is the nuclear wavefunction, and $\psi(\{\mathbf{r}\})$ is the electronic wavefunction. Next, one can solve the electronic Schrödinger by treating the nucleus positions as parameters, and the nuclei-electron interaction becomes the external potential felt by electrons:

$$\hat{\mathcal{H}}_e\Psi(\{\mathbf{r}\}) = E_e\Psi(\{\mathbf{r}\}). \quad (19)$$

In fact, this is the starting point of most of the electronic structure theories. Once the E_e is solved, one can then calculate the total energy E_{tot} of a system by treating nuclei as point charges, *i.e.* in a classical way. In other words, at 0 K, one can neglect the kinetic energy of nuclei, and the total energy E_{tot} is the sum of electronic energy and the electrostatic energy between nuclei:

$$E_{tot} = E_e + E_{NN}, \\ E_{NN} = \sum_{N < M} \frac{Z_N Z_M}{|\mathbf{R}_N - \mathbf{R}_M|}. \quad (20)$$

We next discuss the foundation of the DFT, Hohenberg-Kohn theorems and Kohn-

Sham equation. In the seminal paper by Hohenberg and Kohn [80], they prove that in an non-degenerate electronic system subjected to an external potential $v_{ext}(\mathbf{r})$, such $v_{ext}(\mathbf{r})$ has an one-to-one relationship with electron density. Or more precisely:

Theorem 1: The external potential, and hence the Hamiltonian is uniquely determined by a single and specific electron density.

Based on Theorem 1, in principle, *all* of the properties, or expectation values of *all* observables, of an electronic system can be determined as long as the ground state electron density is given. The second Hohenberg-Kohn theorem follows the idea of the TF model that the total energy of electronic system can be expressed as a density functional:

$$\begin{aligned} E[\rho(\mathbf{r})] &= T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + V_{eN}[\rho(\mathbf{r})] \\ &= F[\rho(\mathbf{r})] + V_{eN}[\rho(\mathbf{r})], \end{aligned} \tag{21}$$

where $F[\rho(\mathbf{r})]$ does not depend on the external potential $V_{ext}(\mathbf{r})$, and is known as universal functional. Hohenber and Kohn shows that minimum of energy functional can be obtained through variational principle, and the second theorem is abtained:

Theorem 2: The electron density that minimizes the total energy functional is the unique ground state density.

Although the powerful Hohenberg-Kohn theorems lay a solid foundation for DFT, they do not provide a practical framework like the mathematical forms of $T[\rho(\mathbf{r})]$ and $V_{ee}[\rho(\mathbf{r})]$ to obtain the ground sate density. One year later, Kohn and Sham proposed the Kohn-Sham equation, which provides a practical scheme to solve ground state density as well as ground state total energy [81]. The key ingredient makes the success of Kohn-Sham equation is that there exist a local single-particle potential, which is able to reproduce the ground state density of the real interacting system. To derive the Kohn-Sham equation, we start from Eq. (21), and our task is to minimize

the total energy $E[\rho(\mathbf{r})]$ for a system with a specific external potential v_{ext} and fixed number of electrons N .

$$0 = \delta \left\{ E[\rho] - \mu \int \rho(\mathbf{r}) d\mathbf{r} \right\} = \int d\mathbf{r} \left[\frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} + v_{ext}(\mathbf{r}) - \mu \right] \delta \rho(\mathbf{r}), \quad (22)$$

where μ is the Lagrange multiplier. The obtained Euler equation is

$$\frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} + v_{ext}(\mathbf{r}) = \mu. \quad (23)$$

The next step is the essence of Kohn-Sham equation: creating a non-interacting system with a different external potential $v_{KS}(\mathbf{r})$, and such a new potential contains the original $v_{ext}(\mathbf{r})$ and all of the electron-electron interactions. Therefore, we can divide the interacting kinetic energy into the non-interacting part and the interacting part,

$$T[\rho] = T_s[\rho] + T_c[\rho], \quad (24)$$

where $T_s[\rho]$ is the non-interacting kinetic energy (denoting s for single-particle), and $T_c[\rho]$ is the correlation part in kinetic energy (denoting c for correlation). The V_{ee} can also be represented by

$$V_{ee}[\rho] = V_H[\rho] + E_x[\rho] + V_c[\rho], \quad (25)$$

where $V_H[\rho]$ is the classic Hartree energy, $E_x[\rho]$ is the quantum-mechanical exchange energy, and $V_c[\rho]$ is the correlation part of electron-electron interaction. The correlation energy is then $E_c[\rho] = T_c[\rho] + V_c[\rho]$. In DFT, the exchange energy and correlation energy functionals are further combined as exchange-correlation energy functional

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho] \quad (26)$$

Therefore, the total energy functional can be represented as

$$\begin{aligned} E[\rho] &= T[\rho] + V_{ee}[\rho] + V_{eN}[\rho] \\ &= T_s[\rho] + V_H[\rho] + E_{xc}[\rho] + V_{eN}[\rho]. \end{aligned} \tag{27}$$

The new universal functional becomes

$$F[\rho] = T_s[\rho] + V_H[\rho] + E_{xc}[\rho]. \tag{28}$$

From Eq. (22) and (23), the new Euler equation hence becomes

$$\begin{aligned} \mu &= \frac{T_s[\rho]}{\delta\rho(\mathbf{r})} + \frac{V_H[\rho]}{\delta\rho(\mathbf{r})} + \frac{E_{xc}[\rho]}{\delta\rho(\mathbf{r})} + v_{ext}(\mathbf{r}) \\ &= \frac{T_s[\rho]}{\delta\rho(\mathbf{r})} + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v_{ext}(\mathbf{r}) \\ &= \frac{T_s[\rho]}{\delta\rho(\mathbf{r})} + v_{KS}(\mathbf{r}), \end{aligned} \tag{29}$$

where $v_H(\mathbf{r})$ is Hartree potential, $v_{xc}(\mathbf{r})$ is exchange-correlation potential, and we also introduce Kohn-Sham potential $v_{KS}(\mathbf{r}) = v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v_{ext}(\mathbf{r})$. Comparing Eq. (29) and (23), we find that by simply rearrange the energy functional terms, we can obtain the ground state density of an interacting system by solving the alternative non-interacting system subjected to Kohn-Sham potential $v_{KS}(\mathbf{r})$.

Now, to deal with the non-interacting system, we can obtain the ground state density by solving the N single-electron Kohn-Sham equations.

$$\left[-\frac{1}{2}\nabla^2 + v_{KS}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \tag{30}$$

The electron density corresponds to a set of Kohn-Sham orbitals ψ_i is

$$\rho_0(\mathbf{r}) = \sum_i^{occ} |\psi_i(\mathbf{r})|^2. \tag{31}$$

The obtained density is not the solution yet. The reason is that the Kohn-Sham potential $v_{KS}(\mathbf{r})$ depends on the density $\rho(\mathbf{r})$, and $v_{KS}(\mathbf{r})$ also depends on $\psi_i(\mathbf{r})$. Hence, we need to solve Eq. (30) and (31) in a self-consistent way like in the HF method. Once solved, the ground state energy can be calculated by

$$E_0 = \sum_i^N \epsilon_i - \frac{1}{2} \int \int d\mathbf{r} d\mathbf{r}' \frac{\rho_0(\mathbf{r})\rho_0(\mathbf{r}')'}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[\rho_0] - \int d\mathbf{r} \rho_0(\mathbf{r}) v_{xc}(\mathbf{r}) \quad (32)$$

The expression of for total energy in Eq. (32) is in principal exact, but in practice, the $E_{xc}[\rho]$ and $v_{xc}(\mathbf{r})$ are unknown, and needed to be approximated.

In our previous discussion for the TF model, the concept of LDA has been considered for the kinetic energy functional. Here, in the context of exchange-correlation energy functional, we first introduce the famous LDA, which has brings tremendous success for DFT in the studies of solids and molecules in the real world. The LDA assumes that the exchange-correlation energy functional is the sum of the pieces of exchange-correlation energies within the infinitesimal volumes, where the exchange-correlation energy is based on the result of homogeneous electron gas (HEG):

$$E_{xc}^{LDA}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{xc}^{LDA}(\rho(\mathbf{r})) d\mathbf{r}, \quad (33)$$

and hence the exchange-correlation potential is

$$v_{xc}^{LDA}(\mathbf{r}) = \frac{\delta E^{LDA}}{\delta \rho(\mathbf{r})} = \varepsilon_{xc}^{LDA}(\rho(\mathbf{r})) + \rho(\mathbf{r}) \frac{\partial \varepsilon_{xc}^{LDA}(\rho)}{\partial \rho(\mathbf{r})}, \quad (34)$$

where ε_{xc}^{LDA} can be divided into exchange part and correlation part,

$$\varepsilon_{xc}^{LDA} = \varepsilon_x^{HEG} + \varepsilon_c^{HEG}. \quad (35)$$

The exchange energy functional in the LDA is

$$E_x^{LDA}[\rho] = \int \rho(\mathbf{r}) \varepsilon_x^{HEG}(\rho(\mathbf{r})) d\mathbf{r}, \quad (36)$$

The exchange part was carried out by Dirac with an exact form [79],

$$\varepsilon_x^{HEG} = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{\frac{1}{3}} \rho^{\frac{1}{3}}. \quad (37)$$

The correlation part does not have an exact form, and its mathematical form is mostly parametrised by the results of using quantum Monte Carlo method to solve HEG systems. There are many examples including Perdew-Zunger [82], Perdew-Wang [83], Vosko-Wilk-Nusair [84], etc. As an example for many different forms, the Perdew-Zunger's LDA formula [82] is expressed as

$$\varepsilon_c^{PZ}(r_s) = \begin{cases} A \ln r_s + B + C r_s \ln r_s + D r_s & r_s \leq 1 \\ \gamma / (1 + \beta_1 \sqrt{r_s} + \beta_2 r_s) & r_s > 1, \end{cases} \quad (38)$$

where r_s is the Wigner-Seitz radius ($\frac{4}{3}\pi r_s^3 = \frac{1}{\rho}$). A , B , C , and D are parameters based on the results of random phase approximation, while γ , β_1 , β_2 are constants based on the precise Quantum Monte Carlo simulation data by Ceperley and Alder [85].

For spin-polarized systems, the LDA of exchange-correlation energy functional is called local spin density approximation (LSDA), which depends on two spin densities:

$$E_{xc}^{LSDA}[\rho_\uparrow, \rho_\downarrow] = \int \rho(\mathbf{r}) \varepsilon_{xc}^{HEG}(\rho_\uparrow, \rho_\downarrow) d\mathbf{r}, \quad (39)$$

where $\rho = \rho_\uparrow + \rho_\downarrow$. The exchange part is simply

$$E_x^{LSDA}[\rho_\uparrow, \rho_\downarrow] = \frac{1}{2}(E_x^{LDA}[2\rho_\uparrow] + E_x^{LDA}[2\rho_\downarrow]). \quad (40)$$

The spin correlation energy density is constructed by interpolating the extreme values. It has a form of $\varepsilon_c = \varepsilon_c(\rho, \zeta)$, where $\zeta = (\rho_\uparrow - \rho_\downarrow)/\rho$ is the relative spin-polarization.

Since in the real systems, the density $\rho(\mathbf{r})$ is spacially varying, adding higher order terms like $\nabla\rho(\mathbf{r})$ and $\nabla^2\rho(\mathbf{r})$ should better describe the exchange-correlation energy functional. The first implementation of this idea is gradient expansion approximation (GEA), which systematically increases higher order gradient terms. However, it was found that the GEA with first order correction hardly improves the results comparing with LDA [86]. Since adding the first order correction fails to improve the functional, including higher order terms cannot be justified. People latter found that the functional could be improved by more general functions with only $\rho(\mathbf{r})$ and $\nabla\rho(\mathbf{r})$ as variables. Such functional is known as generalized gradient approximation (GGA), and they have a form as

$$E_{xc}^{GGA}[\rho_\uparrow, \rho_\downarrow] = \int d\mathbf{r} f(\rho_\uparrow, \rho_\downarrow, \nabla\rho_\uparrow, \nabla\rho_\downarrow) \quad (41)$$

Again, like in LDA, the exchange and correlation functionals are separated, $E_{xc}^{GGA} = E_x^{GGA} + E_c^{GGA}$. The GGA exchange functional has a form like LSDA

$$\begin{aligned} E_x^{GGA}[\rho_\uparrow, \rho_\downarrow] &= \frac{1}{2}(E_x^{GGA}[2\rho_\uparrow] + E_x^{GGA}[2\rho_\downarrow]) \\ E_x^{GGA}[\rho] &= \int \rho(\mathbf{r}) \varepsilon_x^{HEG}(\rho) F_x(s) d\mathbf{r} \end{aligned} \quad (42)$$

where $F_x(s)$ is a function that modifies the LDA exchange functional with s being a function of ρ and $|\nabla\rho|$ [86]. The correlation part of GGA does not have an unified form. One of the expressions introduced by Perdew, Burke, and Ernzerhof (PBE) has

the following form [86]:

$$E_c^{GGA}[\rho_\uparrow, \rho_\downarrow] = \int \rho(\mathbf{r}) [\varepsilon_c^{HEG}(\rho, \zeta) + H^{PBE}(\rho, \zeta, t)], \quad (43)$$

where t is a function of ρ , ζ , and $|\nabla \rho|$.

The performance of DFT results relies on the quality of the approximation for the exchange-correlation energy functional. In general, DFT results based on GGA functional show high accuracy in properties like energies and bond lengths in molecule, as well as lattice constants in solids. However, the LDA and GGA are known to underestimate the energy gap obtained in band structure calculations, which is the so-called “bandgap problem”. It is known that bandgap estimations based on hybrid functional DFT and GW approximation in many-body perturbation are much more accurate. Since we have used hybrid functional to better estimate the bandgap, we give a brief introduction in the following.

In the hybrid functionals, a portion of non-local exact exchange energy from HF theory is incorporated with the local exchange part in LDA or GGA. The HF exact exchange functional can be calculated by

$$E_x^{HF} = -\frac{1}{2} \sum_{\mathbf{k}, \mathbf{k}'} \sum_{i,j}^{occ} \int \psi_{i\mathbf{k}}^*(\mathbf{r}) \psi_{j\mathbf{k}'}^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_{i\mathbf{k}}(\mathbf{r}') \psi_{j\mathbf{k}'}(\mathbf{r}) d\mathbf{r} d\mathbf{r}', \quad (44)$$

where $\psi(\mathbf{r})$ stands for the Khon-Sham orbital. The hybrid functional then has a general form:

$$E_{xc}^{Hybrid} = E_{xc}^L + a(E_{xc}^{HF} - E_x^L), \quad (45)$$

where E_{xc}^L is the local exchange-correlation energy functional (like LDA and GGA). A hybrid functional based on PBE local functional is known as PBE0 [87], where the value of a is 0.25. Another well-known hybrid functional proposed by Heyd, Scuseria, and Ernzerhof is HSE hybrid functional, where an error function is used to separate

the long-range and the short-range parts [88].

$$\frac{1}{r} = \frac{\text{erfc}(\omega r)}{r} + \frac{\text{erf}(\omega r)}{r}. \quad (46)$$

In the first term, $\text{erfc}(\omega r) = 1 - \text{erf}(\omega r)$. The first and the second terms are the short-range and long-range parts, respectively. ω is the is an adjustable parameter. The complete form of HSE hybrid functional is shown as

$$E_{xc}^{HSE} = aE_x^{HF,SR}(\omega) + (1-a)E_x^{PBE,SR}(\omega) + E_x^{PBE,LR}(\omega) + E_c^{PBE}. \quad (47)$$

In HSE06 hybrid functional, used in this dissertation, the values of a and ω are 0.25 and 0.2, respectively [89]. Basically, GGA functionals (like PBE) can give us reliable results for most properties including total energy, lattice parameters, and strengths of chemical bonds (covalent, ionic, metallic and hydrogen bridge). To study the energy gap between valence and conduction bands, the hybrid functionals can be taken into consideration while it is computationally expensive.

After the discussion of how we can use DFT to calculate the total energy, we next discuss how forces and stresses are computed based on Hellman-Feynman theorem. For a Hamiltonian which can be parametrized by a variable λ , we have

$$\hat{H}(\lambda)\Psi(\lambda) = E(\lambda).\Psi(\lambda) \quad (48)$$

Based on Hellman-Feynman theorem, the derivative of E with respect to λ can be calculated by

$$\frac{\partial E(\lambda)}{\partial \lambda} = \left\langle \Psi(\lambda) \left| \frac{\partial \hat{H}}{\partial \lambda} \right| \Psi(\lambda) \right\rangle. \quad (49)$$

Therefore, the force on the nucleus I is

$$\mathbf{F}_I(\mathbf{R}) = -\nabla_{\mathbf{R}}E(\mathbf{R}) = -\langle \Psi(\mathbf{R}) | \nabla_{\mathbf{R}}H | \Psi(\mathbf{R}) \rangle. \quad (50)$$

Stresses can also be calculated using a similar but more complicated expression based on Nielsen *et al.* [90].

With the information of forces and stresses, we can do structure relaxation by minimizing both forces and stresses. Here, the procedure of structure relaxation in a DFT code is shown in Fig. 4. At first, we only need to provide the initial crystal structure, and then the code enter the electronic loop to obtain the ground state electron density by solving the Kohn-Sham equation self-consistently, which is also known as the self-consistent field calculation (shown in the yellow box in Fig. 4). The update of electron density is done by mixing the new and the old densities, and the convergence criteria include the energy difference (usually less than 10^{-5} eV/atom), and also the differences in eigenvalues. Once the self-consistent field is done, the code will calculate the forces and stresses. By means of algorithms like quasi-Newton, and conjugate-gradient methods, the crystal structure can be gradually optimized. The convergence criterion for structure relaxation is usually the force on each atom less than 10^{-2} or 10^{-3} eV/Å. The structure relaxation is also important in the procedure of crystal structure prediction, which will be introduced in the next section.

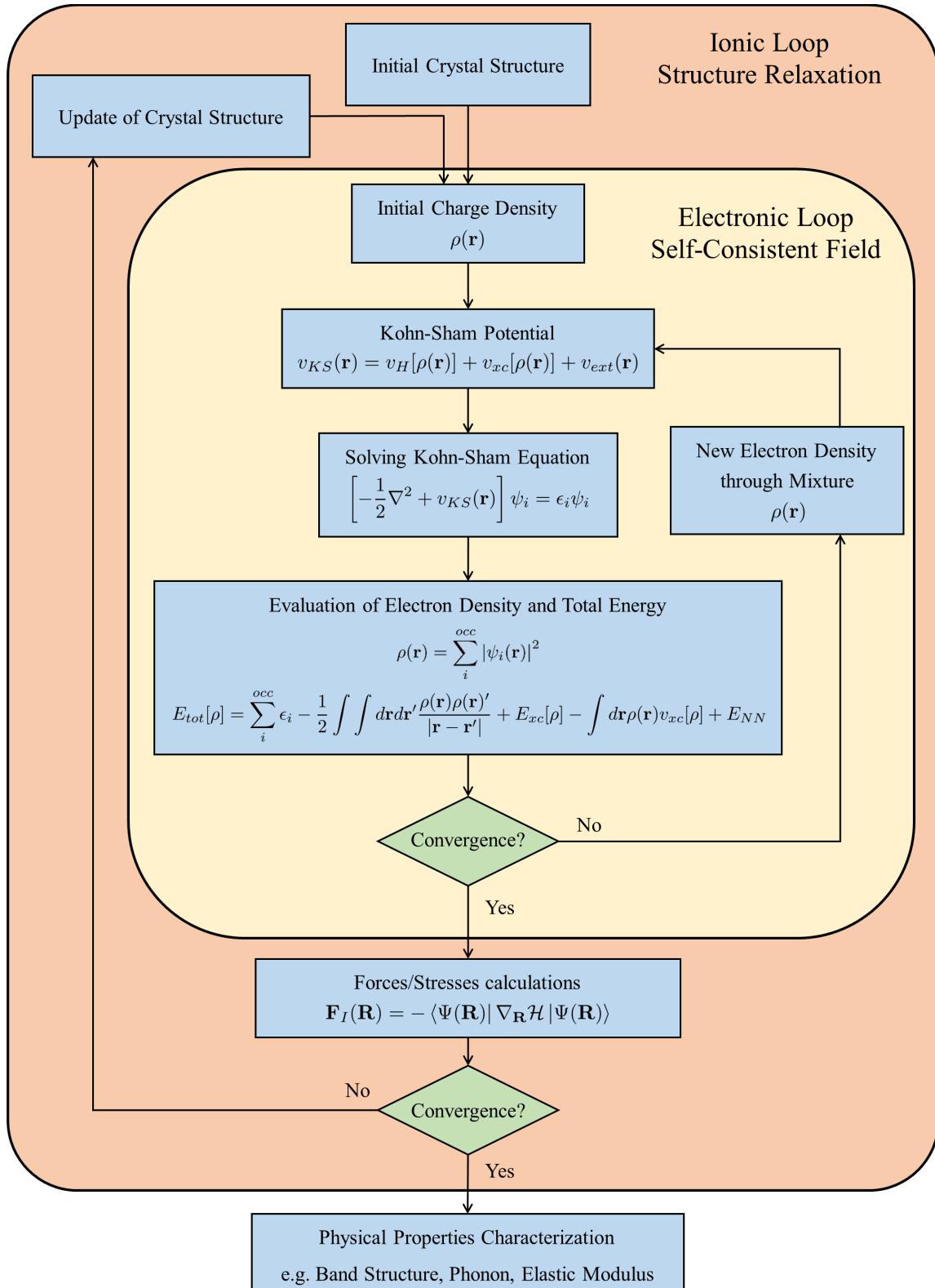


Figure 4: Flowchart of density functional theory calculations including the ionic loop (outer loop) for crystal structure relaxation and electronic loop (inner loop) for self-consistent field.

Crystal Structure Prediction and Evolutionary Algorithm

The crystal structure, the arrangement of compositional atoms, in a piece of material is the most important information that determines arguably all of the intrinsic physical properties of a material. With the information of crystal structure, one can compute several properties related to electronics, thermodynamics, vibration, optics, and magnetism through first-principles calculations like DFT. This allows researchers to search for materials with desired properties before conducting experiments like material syntheses and measurements. In certain situations, theoretical results might be the only way that one can resort to. For example, investigating materials at condition that cannot be approached by current experimental tools like extremely high pressure research. However, to obtain the crystal structure of a material fully by theoretical and computational tools is still a challenging task today, which is known as crystal structure prediction (CSP) [91].

The goal of CSP is to search for stable or metastable structures by exploring the gigantic configuration space of high-dimensional potential-energy surface (PES) with the knowledge of chemical compositions only. In the generic case, CSP is an optimization problem to locate the global or local minima in the Gibbs free energy landscape. The Gibbs free energy (G) is defined as

$$G = E_{tot} + PV - TS, \quad (51)$$

where E_{tot} is the total energy, P is the pressure, V is the volume, T is the temperature, and S is the entropy. Calculations for entropy involves heavy phonon calculations in solid systems, so in practice the enthalpy $H = E_{tot} + PV$ is considered as the fitness function to be minimized. With the progress of first-principles calculations, we can accurately estimate the enthalpy by DFT-based packages like VASP and QUANTUM ESPRESSO. However, this optimization problem is challenging due to the facts of high

dimensional configuration space and the huge number of local minimums when the number of atoms is large. We can estimate the difficulty of CSP problem by following the arguments in Ref. [92]. The total number of configuration is $C = \binom{V/\delta^3}{N} \prod_i \binom{N}{n_i}$ if not considering correlations between atoms. In this expression, V is the volume, δ is the length-scale of the grid, $N = \sum_i n_i$ is the total number of atoms, and n_i is the number of atom of i type.

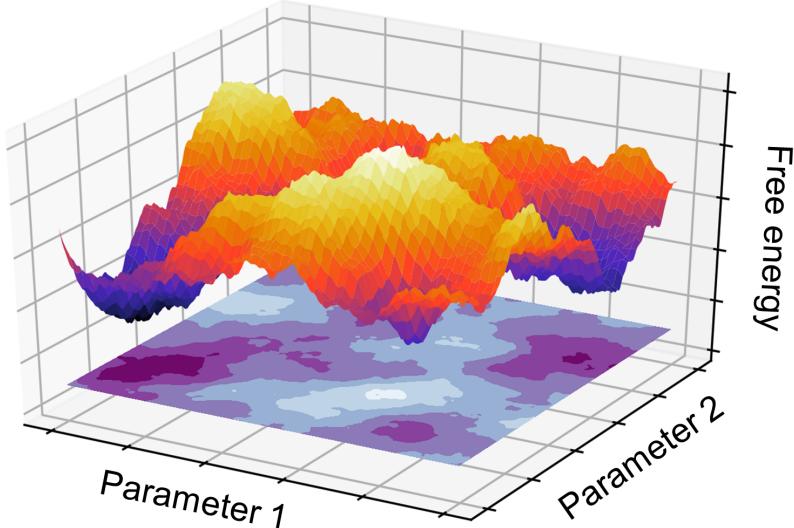


Figure 5: Schematic plot of free energy landscape.

Besides, the dimensionality of the configuration space is $3N + 6 - 3$ ($3N$ for atomic position, 6 for lattice geometry, 3 for the consideration of translational symmetry). To roughly estimate the number C , we consider a cubic box with a volume of 100 \AA^3 that contains 10 atoms in two kinds, 5 for each. If we choose δ to be 1 \AA , then what we are dealing with is an optimization process with 33 variables, and the total number of configuration is $\sim 1.1 \times 10^{18}$. In reality, we do not need to deal with such a huge phase space. Instead, this problem becomes searching for local minimums when we can take advantage of local optimization, which is the (local) structural relaxation built in modern DFT codes (see Fig. 4). The local optimization is usually based on methods like steepest-descent, conjugate-gradient, Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, etc. However, the number of

local minimum also grows exponentially with increasing number of atoms.

To search for the local minimum efficiently, we need reliable global optimization algorithms to explore the configuration space. It is important to note that no method can guarantee to find the global minimum within a certain amount of time. Here we list some of popular global optimization algorithms: quasi-random, simulated annealing, meta-dynamics, minimum-hopping, particle swarm algorithm and evolutionary algorithm. In the following, we discuss the procedure of the method used in this dissertation – evolutionary algorithm.

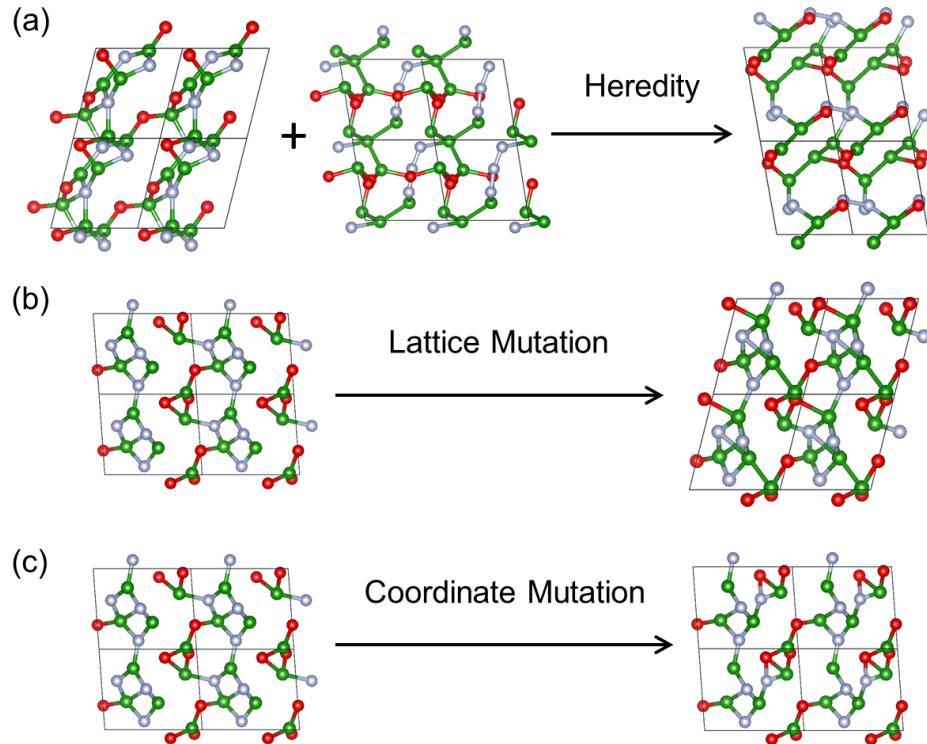


Figure 6: Examples of variation operators used to create new structures in the evolutionary algorithm for crystal structure prediction.

Evolutionary algorithm is a population-based optimization method aiming at finding the best solution in configuration space in a heuristic and also stochastic way. The process of evolutionary algorithm is inspired by the biological evolution, where the new generation can be created by operators like heredity or mutation. The heredity operator is also known as crossover operator or recombination, where a new candi-

date is a combination of two parts from its parents (one for each). In the context of crystal structure prediction, two sub-structures from parent structures are combined together to form a new structure. As for the mutation operator, it produces a new candidate from an old candidate in the population. The mutation operator can be lattice mutation or coordinate mutation that modifies an old structure by changing parts of structure variables. The heredity, lattice mutation, and coordinate mutation operators in crystal structure prediction can be found in Fig. 6.

Today, evolutionary algorithm packages for crystal structure prediction are usually combined with density functional theory to accurately evaluate the fitness function. In most of the cases, the fitness function is enthalpy, and the minimized enthalpy is corresponding to the thermodynamic favorable structure at a given pressure. Here, we describe the main steps in evolutionary algorithm:

step 1: Generating random structures

step 2: Local optimization and fitness evaluation for the first generation

step 3: Construction of initial population

step 4: Generating new structures through evolutionary operators.

step 5: Local optimization and fitness evaluation for new structures

step 6: Iterating *step 4* and *step 5* until fulfilling the convergence criteria

The flowchart of evolutionary algorithm can be found in Fig. 7 .The fitness function can also be a physical property like the bulk modulus or dielectric constant, and the desired property can be optimized. However, such structures might not be thermodynamically stable in reality.

In this dissertation, we use USPEX (Universal Structure Predictor: Evolutionary Xtallography) package to carry out evolutionary searches in B-C-O and B-N-O systems. USPEX, designed by Oganov *et al.* [93, 94, 95], is an highly-efficient in crystal structure prediction, and its success is a result of carefully designed varia-

tion operators as well as efficient constraints that avoid unphysical and redundant structures in the search space. Besides, it also contains alternative methods including random sampling, evolutionary metadynamics, and corrected particle swarm optimization algorithms. These alternative methods with evolutionary algorithm allow one to explore the search space in a more complete way. In particular, evolutionary metadynamics is very powerful not only for global optimization and searching for low-energy metastable structures but also predicting phase transition paths. In fact, this feature allows one to evaluate the energy barrier between two meta-stable structure, and therefore is highly-related to another important problem – the synthesizability of a metastable structure.

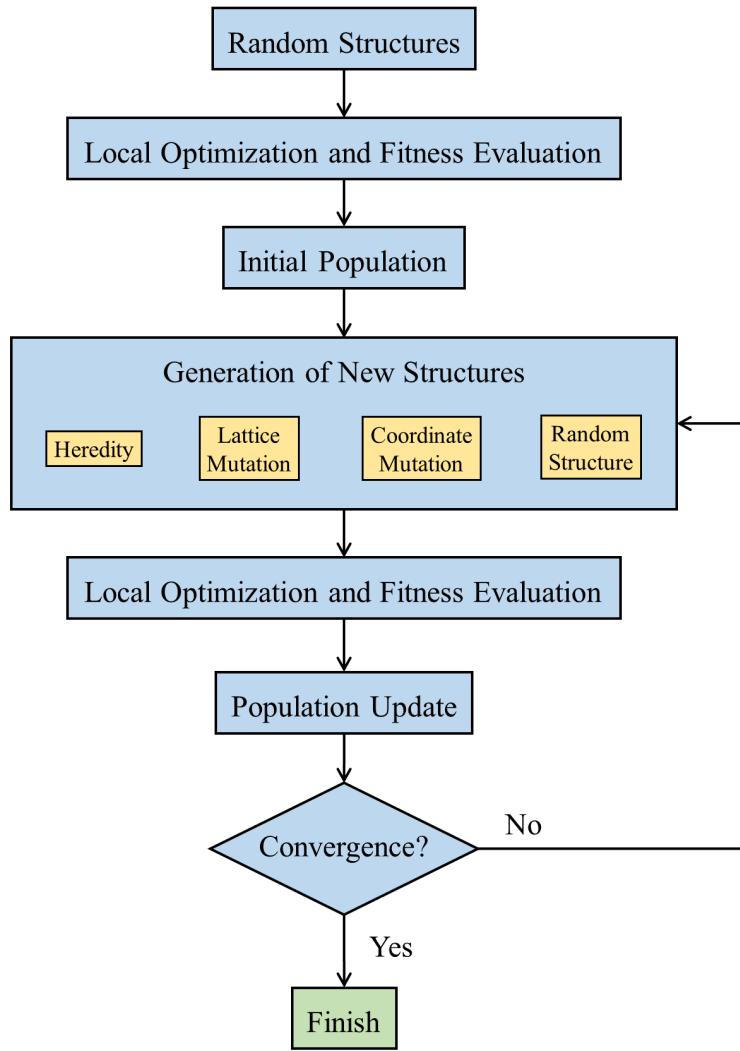


Figure 7: Flowchart of evolutionary algorithm.

CHAPTER 3

MACHINE LEARNING AND EVOLUTIONARY DISCOVERY OF TERNARY SUPERHARD MATERIALS

B-C-N Superhard Materials

First-principles simulations based on density functional theory have played important roles in predicting new superhard compounds. However, *ab initio* methods are still computationally expensive and size-limited. On the other hand, data-driven approaches have proven to be powerful and efficient in exploring new materials [96, 97, 98, 99, 100, 101] – thanks to recent advance in computing hardwares, development in machine learning algorithms, and availability of online materials database. For example, Meredig *et al.* [102] have constructed a machine learning model to *screen over 1.6 million ternary compositions* and predicted 4,500 novel, potentially stable ternary materials. Therefore, data-driven machine learning approaches are promising for large-scale materials design and discovery.

In principle, a machine learning framework can be implemented with different material features or descriptors for a wide range of target properties. Two popular properties to predict are bulk and shear moduli [103, 104, 105, 106, 107, 108], which are also correlated with the material hardness. For example, de Jong *et al.* [104] developed a technique based on gradient boosting and used features like the volume per atom and cohesive energy. Mansouri *et al.* [107] used support vector machines and combined elemental and structural properties as descriptors, where the cohesive energy was also identified as a crucial feature. These machine learning studies typically can achieve high prediction accuracy with only a few thousands of training data

points. However, using cohesive energy, volume, melting point, crystal symmetry and so on as features may be less ideal, as obtaining these information for new compounds would require additional measurements or calculations.

To explore the vast configuration space of ternary compounds, we develop random forests models to predict material mechanical properties, by using only features that can be derived directly from the chemical formula. The resulting machine learning models thereby can achieve large-scale prediction of new superhard and ultra incompressible materials for extreme environment applications. We also employ evolutionary structure prediction and density functional theory calculations to further validate the machine learning results. In particular, we propose three new superhard compositions – BC_{10}N , $\text{B}_4\text{C}_5\text{N}_3$, and $\text{B}_2\text{C}_3\text{N}$ – and fully characterize their structural, phonon, and electronic properties. These new superhard compounds are all dynamically stable with relatively low formation energy, so they can potentially be synthesized by low-temperature plasma methods, *without the need of* high-temperature high-pressure conditions. It is noted that our newly suggested compound BC_{10}N has a computed hardness value ~ 87 GPa; once synthesized, the compound would become the second hardest material. Our computational flowchart is summarized in Fig. 8 and discussed in detail in the next section.

Data acquisition – There exist several online computational materials databases, such as AFLOW [114], Materials Project [109], NOMAD Encyclopedia [115], and the Open Quantum Materials Database (OQMD) [116]. Here we use the Materials Project [109], which provides open access to various computed properties of known and predicted crystalline compounds. The corresponding PYMATGEN (Python Materials Genomics) library [110] is utilized to extract the target properties of bulk modulus (K) and shear modulus (G). The Materials Project [109] database also contains high-pressure phases and artificial crystal structures, which can exhibit extreme values of bulk and shear moduli. Therefore, we exclude those extreme outliers and

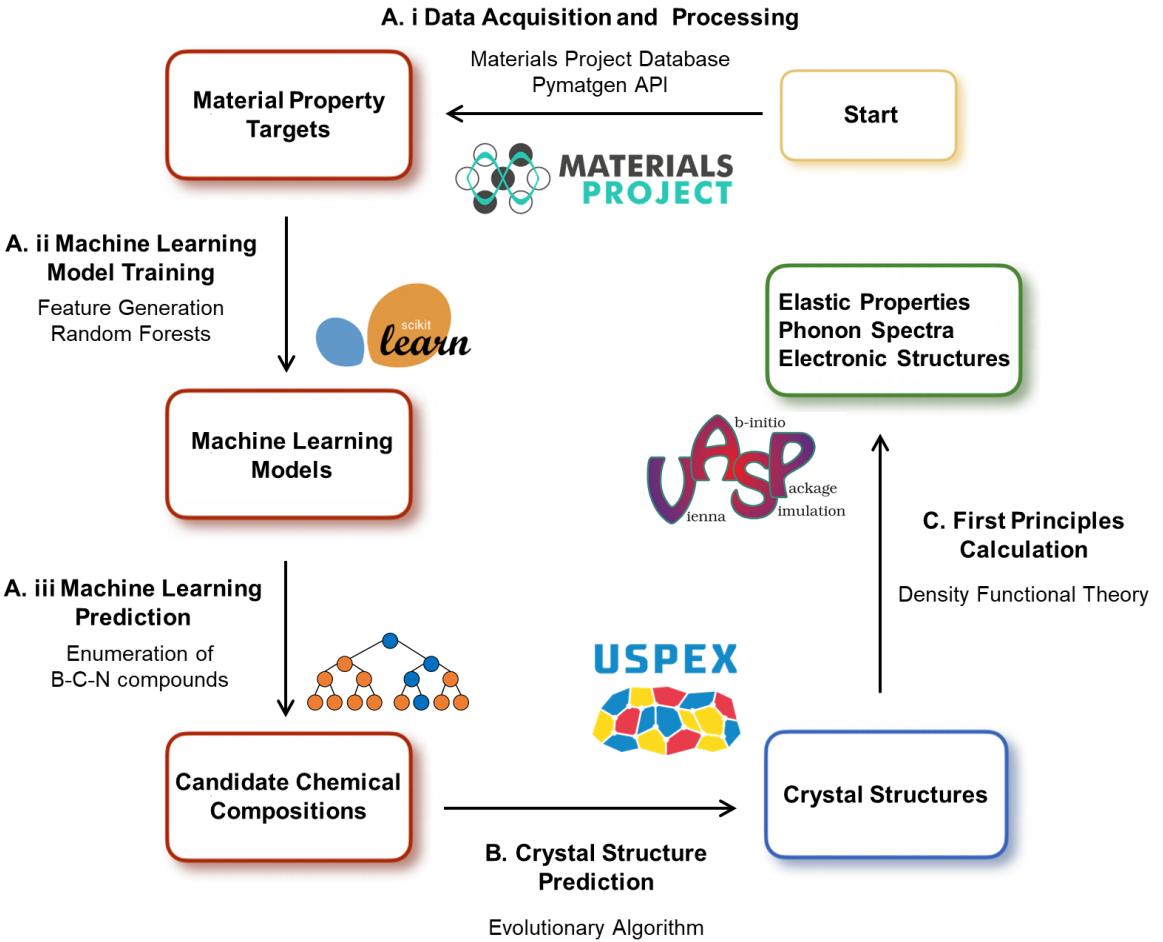


Figure 8: Computational flowchart of data-driven discovery of new superhard materials in our B-C-N study: **A.i** Data acquisition and processing using the Materials Project [109] database and its application programming interface (API) PYMATGEN [110]; **A.ii** Machine learning model training with handcrafted features and regression algorithms implemented in the SCIKIT-LEARN library [111]; **A.iii** Random forests prediction of chemical compositions for candidate superhard materials. **B.** Crystal structure prediction of a given chemical formula using evolutionary algorithms implemented in the USPEX program [93, 94, 95]. **C.** First-principles validation of the machine learning results with density functional theory calculations using the VASP software [112, 113]

focus on 10,421 selected compounds with K and G values both in the ranges of 0–550 GPa.

Feature generation – To build a supervised learning model using only chemical composition as input, we generate features (or descriptors) based on a compound’s chemical formula. By following Ref. [117], we consider features related to stoichiometric attributes, elemental properties, orbital occupations, and ionic levels. Part of the features can be generated with the Python library MATMINER [118]. We *do not* consider structural or electronic features like crystal symmetry, volume, melting point, band gap, etc. While including these additional features could improve the model performance, these information is *a priori* unknown for new compounds. To expedite materials discovery, we thereby do not include features that require additional first-principles calculations.

First, the stoichiometric features are L^p norm $\|x\|_p = (\sum_i |x_i|^p)^{1/p}$, where x_i is element i ’s atomic fraction. These attributes capture the changes in atomic fraction, independent of the actual elements. As an example, the $p = 2$ norm of Fe₂O₃ is $\|x\|_2 = ((\frac{2}{5})^2 + (\frac{3}{5})^2)^{1/2} \simeq 0.721$ [117]. Here we consider 3 stoichiometric features, including the $p = 0$ norm (i.e. the number of chemical components), and the $p = 2, 3$ norms. The $p = 1$ norm is equal to unity regardless of the chemical composition, so it is not considered. In addition, we do not find an apparent model improvement with more higher order norms ($p > 3$), so they are not included.

Second, the elemental features are computed using the minimum, maximum, and range for properties of each element present, as well as the values of fraction-weighted mean $\bar{f} = \sum_i x_i f_i$ and average deviation $\hat{f} = \sum_i |f_i - \bar{f}|$. Here, f_i is the property of element i , and x_i is the atomic fraction. We consider the following 10 properties: atomic number, atomic mass, element column number, row number, atomic radius, electronegativity, and the numbers of valence electrons in s , p , d , and f orbitals, respectively. Therefore, there are 50 elemental-property features (= 5 values

$\times 10$ properties). Using again Fe_2O_3 as an example [117], for the “atomic number” property, $\bar{f} = \frac{2}{5}(26) + \frac{3}{5}(8) = 15.2$ and $\hat{f} = \frac{2}{5}|26 - 15.2| + \frac{3}{5}|8 - 15.2| = 8.64$.

Third, 4 orbital-occupation features are computed using the fraction-weighted average of the number of valence electrons respectively in s , p , d , and f orbitals, divided by the fraction-weighted average of the total number of valence electrons. For example, Fe_2O_3 ’s p -orbital occupation feature is $F_p = \frac{2/5 \times (0) + 3/5 \times (4)}{2/5 \times (8) + 3/5 \times (6)} \simeq 0.353$ [117].

Finally, 3 features are based on ionic levels. The first is a Boolean number denoting whether it is possible to form a neutral ionic compound, by assuming that each element takes exactly one of its common charge states. The other two features are based on the “ionic character” of a chemical bond: $I(\chi_i, \chi_j) = 1 - \exp(-(\chi_i - \chi_j)^2 / 4)$, where χ_i and χ_j are electronegativities for elements i and j , respectively. In Pauling scale, fluorine has the highest electronegativity value of $\chi_F = 3.98$, and francium has the lowest electronegativity value of $\chi_{Fr} = 0.70$. The two features we consider are respectively the maximum ionic character I between any two elements in a compound, and the mean ionic character $\bar{I} = \sum_{i,j} x_i x_j \chi_i \chi_j$.

In total, 60 features are created. To simplify the training task, we do not consider additional feature engineering such as degree-2 polynomials, which otherwise could lead to thousands of new features and cause overfitting. The chemical compositions and their target properties of bulk and shear moduli for the 10,421 compounds considered here are written as a Python dictionary object saved in a JSON file. The features for all compounds are provided as a csv file accordingly. Both the data and our machine learning codes can be found in our github repository [119].

Model training, validation, and application – For regression task, we choose the random forests algorithm [120, 121], which is a tree-based ensemble method. A random forests model builds multiple decision trees, by taking a random sample with replacement from the training set and a random subset of features to split tree nodes. The results averaged over individual trees serve as the final predictions, which help re-

duce variance and improve accuracy. However, without restriction on the tree depth, the model can become very deep and cause overfitting. Therefore, we constrain the pre-pruning parameter of tree depth to regularize the models.

The model training is implemented with the SCIKIT-LEARN library [111]. We use 90% of our samples as the training and validation set, which is then used to determined the tree depth by 10-fold cross-validation. The remaining 10% is the test set used for an unbiased evaluation of the final model. We build two separate models to predict the bulk modulus (K) and shear modulus (G), respectively. We do not train a model for predicting the Vickers hardness (H), as the target hardness value is not as widely available as K and G . On the other hand, there exist several empirical models for evaluating H [122, 123, 124, 125, 126, 127], based on physical properties such as bond length, bond strength, electronegativity, and covalent radius. Here we adopt hardness models that require only bulk and shear moduli as inputs [128, 129], so that our regression results of K and G can be employed directly to predict H .

After training and evaluation, we apply the models to predict mechanical properties of B-C-N compounds and search for new superhard ternary materials. For candidate compositions identified with superhardness (i.e. $H \geq 40$ GPa), we then perform crystal structure prediction and first-principles calculations to further validate the machine learning predictions.

Crystal structure prediction (CSP) concerns finding the stable structure of a compound knowing only its chemical formula [130, 131, 132], and it requires an accurate and efficient estimation of the system’s total energy U (usually based on density functional theory), and an efficient optimization technique. Here we utilize the highly efficient implementation of evolutionary algorithm in USPEX package [93, 94, 95]. In USPEX, several features have been added for the enhancement of searching efficiency. For example, USPEX implements efficient constraint techniques being able to neglect unphysical and redundant regions in the search space. In our calculations, we

examine over at least 1,000 structures for a given composition. The first generation of structures are randomly created. Subsequent generations are created with 20% from random structures and 80% from heredity, softmutation, and transmutation operators.

Our first-principles density functional theory (DFT) calculations are performed with the Vienna Ab initio Simulation Package (VASP) [112, 113]. The Monkhorst-Pack sampling scheme [133] is used with a Γ -centered k -point mesh of $21 \times 21 \times 5$ (resolution = $0.02 \times 2\pi/\text{\AA}$) points in the Brillouin zone. The convergence criteria of self-consistent and structural relaxation calculations are set to 10^{-6} eV/unit-cell and 10^{-3} eV/ \AA , respectively. We adopt a plane wave energy cutoff of 520 eV, which is sufficient to converge the DFT total energy difference $< 10^{-4}$ eV/atom. For each crystal structure, we first fully relax the lattice parameters and atomic positions. After structure relaxation, we then compute the corresponding mechanical, electronic, and phonon properties. All calculations use projector augmented wave (PAW) [134, 135] pseudopotentials and the Perdew-Burke-Ernzerhof generalized gradient approximation (GGA) functional [86].

We employ the strain-stress method [136] in VASP to compute the elastic constants C_{ij} , which in turn can determine the bulk and shear moduli using the Voigt-Reuss-Hill formula [137, 138, 139]. Phonon dispersion spectra are computed using the PHONOPY package [140]. Density functional perturbation theory with $2 \times 2 \times 1$ supercells are adopted to evaluate the second-order force constants.

Figure 9 shows histograms of the bulk and shear moduli computed by DFT (denoted respectively as K_{DFT} and G_{DFT}) for 10,421 samples acquired from the Materials Project database [109]. Here, the DFT modulus values represent the Voigt-Reuss-Hill average moduli [137, 138, 139], and the medians of K_{DFT} and G_{DFT} in Fig. 9(a) are 84 GPa and 40 GPa, respectively. Since the accuracy and applicability of a machine learning model largely depend on the training data, we have set a few criteria to select

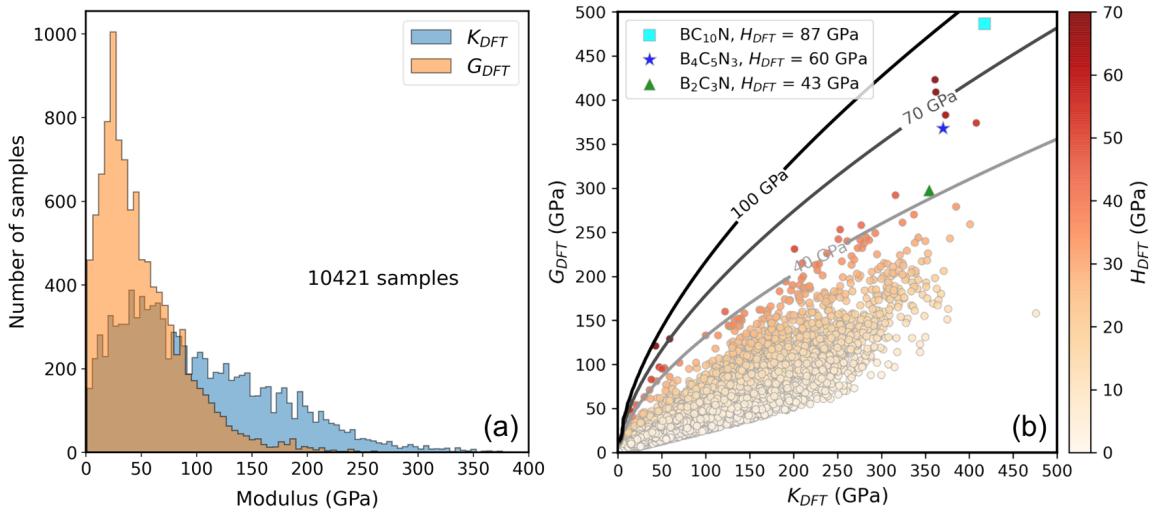


Figure 9: (a) Histogram and (b) scatter plot of bulk (K) and shear (G) moduli for 10,421 samples acquired from the Materials Project [109] database based on density functional theory (DFT) calculations. In (b), the false-color intensity represents the Vickers hardness (H) computed by Tian's empirical model using K and G as inputs. The solid curves represent the hardness contours using Tian's model [129]. These contour lines can help quickly locate compounds with superhardness, with the caveat that the model's applicability might be more limited in the low-bulk/low-shear modulus region. The three newly proposed superhard compounds BC_{10}N , $\text{B}_4\text{C}_5\text{N}_3$, and $\text{B}_2\text{C}_3\text{N}$ are highlighted respectively by the square, star, and triangle symbols.

suitable samples during data acquisition.

First, we have excluded sample materials with a formation energy ≥ 0.2 eV/atom, as they are thermodynamically unfavorable. Second, we have neglected samples whose Voigt and Reuss modulus values differ by more than 50 GPa; this class of samples are typically layered quasi-two-dimensional materials, like graphite or hexagonal boron nitride, which are not the focus of our study. Third, we have utilized the Pugh's ratio k ($\equiv G/K$) [141] to further filter out materials with $k < 0.25$ due to their extremely small hardness, as well as materials with $k > 4.0$, which represents an extreme high-hardness structure (with $H > 200$ GPa) usually computed under high pressure. With these selection criteria, there are in total 10,421 samples considered in our machine learning study.

Figure 9(b) shows the scatter plot for the distribution of bulk and shear moduli. The false-color intensity represents the corresponding material hardness (H), which is calculated by using Tian's empirical model [129]:

$$H_V = 0.92k^{1.137}G^{0.708}. \quad (52)$$

This empirical formula dictates that superhardness requires a large Pugh's ratio k and/or a high shear modulus G . Using Tian's model, we also plot hardness contour lines in Fig. 9(b). Materials in the contour region between $H = 40 - 100$ GPa are mostly compounds like C, BC₂N, c-BN, and M_xB_y (M: Be or transition metal). The three newly proposed superhard ternary compounds – BC₁₀N, B₄C₅N₃, and B₂C₃N – are highlighted respectively by the square, star, and triangle symbols in Fig. 9(b). These materials will be discussed later in the paper.

After data acquisition, we split the whole data into the training-validation set (90%) and the test set (10%). The training-validation set is used for grid search with 10-fold cross validation to search for a proper tree depth of the random forests models.

We find that a maximum depth of 12 layers (with 100 estimators) is reasonable for obtaining a good balance between bias and variance. A deeper tree would not improve the model performance. After deciding on the maximum tree depth, we also utilize the training-validation set to further refine the machine learning models. For an unbiased evaluation of the final model performance, we use the test set and consider the metric of Pearson correlation coefficient r :

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (53)$$

Here, x_i is the machine-learning predicted value for a single entry, and y_i is the corresponding “actual” value computed by DFT. \bar{x} and \bar{y} represent respectively the mean values of the predicted and the actual (DFT) values in the test set, which contains $n \sim 1,000$ sample points. The r value can range between -1 to 1, and $r = 1$ means that the prediction is 100% accurate. Figure 10 shows the r -value

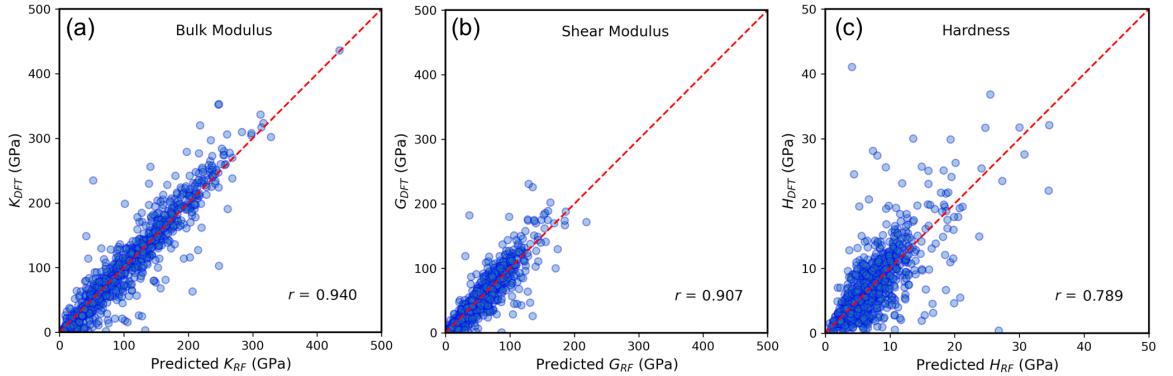


Figure 10: Evaluation of random forests (RF) models using the Pearson correlation coefficient (r) as a metric, for (a) bulk modulus (K), (b) shear modulus (G), and (c) hardness (H). The machine learning models are trained to predict respectively K and G , and both can achieve $r > 0.9$ when applied to the test set based on density functional theory (DFT) calculations. The predicted hardness H_{RF} is obtained by using Tian’s empirical formula with K_{RF} and G_{RF} as inputs, which results in an inferior correlation coefficient as anticipated.

plots using the test set. For bulk and shear moduli [Fig. 10(a) and 10(b)], the data distribution follows closely the $r = 1$ dashed line. The r values of K and G are respectively 0.940 and 0.907 (, and their coefficients of determination r^2 are respectively 0.885 and 0.822). For the hardness H_{RF} in Fig. 10(c), we note that the prediction is *not* obtained directly from a machine learning model. Instead, we use the machine learning predicted K_{RF} and G_{RF} with Tian’s empirical formula in Eq. (1) to compute H_{RF} . Therefore, the data distribution in Fig. 10(c) is more dispersing with a r -value ~ 0.79 , which is slightly inferior as expected. We also note that a higher r -value (or r^2 score) could be achieved by including additional features like volume, crystal symmetry, and cohesive energy, as done in previous studies [103, 104, 105, 106, 107, 108]. However, here we do not consider features that require additional measurements or calculations, but focus only on features that can be derived directly from the chemical formula, in order to achieve efficient large-scale materials discovery.

Our machine learning models also provide information on feature importance to help reveal features that are more correlated with the bulk or shear moduli. Among the 60 features in our study, the atomic radius and d electron occupation are the most important ones. For example, the average atomic radius of a given compound and the bulk modulus exhibit a negative correlation with $r \sim -0.24$. Similarly, the r value between the largest atomic radius and the bulk modulus is $r \sim -0.41$. The results indicate that in general, a smaller crystal unit cell will favor a higher bulk modulus. This is consistent with the facts that most superhard materials consist of light and small elements like Be, B, C, N, and O, and that diamond has the smallest volume per unit cell among all crystalline materials. On the other hand, the d electron occupation rate is positively correlated with the bulk modulus, with a r value ~ 0.58 . This corresponds to the fact that many ultra incompressible materials are transition-metal borides like ReB_2 and Os_2B_3 [142, 143].

With the random forest models, we can predict quickly the mechanical properties of a given chemical formula. Here we apply our models to ternary B-C-N compounds by enumerating a series of $B_xC_yN_z$ compositions, with $x, y, z \in \{1, 2, 3, \dots\}$.

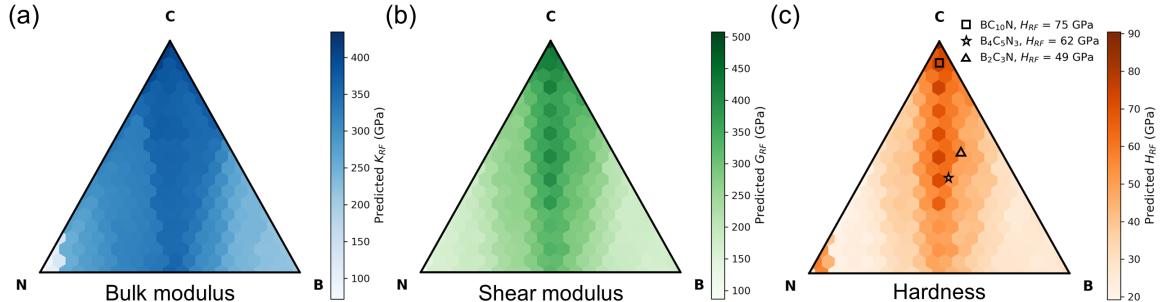


Figure 11: Triangular (or ternary) B-C-N graphs for (a) bulk modulus (K), (b) shear modulus (G), and (c) hardness (H), predicted by random forests (RF) machine learning models. Panel (c) indicates that a 1:1 B:N composition ratio can lead to various superhard compounds such BC_2N ($H_{RF} = 74$ GPa) and BC_4N ($H_{RF} = 65$ GPa). The hardness of three newly proposed superhard compounds, $BC_{10}N$, $B_4C_5N_3$, and B_2C_3N are consistent with the DFT results. The triangular graphs are visualized by using the Python Ternary Plots library [144].

Figure 11 shows the predicted triangular graphs, where the corner points correspond to unitary elemental compounds. For example, the pure boron phase in Fig. 11(c) is predicted to have $H_{RF} \sim 30$ GPa, which could be regarded as the hardness for α -B, β -B, γ -B, or tetragonal B_{52} . For the pure carbon phase, the predicted $H_{RF} \sim 90$ GPa could be related to cubic or hexagonal diamond (lonsdaleite). One caveat is that the relatively high hardness predicted near the pure nitrogen phase may be unrealistic; this is due to small bulk moduli of nitrogen-dominated compounds, which leads to a large Pugh's ratio k and a high hardness when Tian's model in Eq. (1) is used. If we implement more data selection rules by restricting the K and G values to be > 50 GPa, then the artifact near the pure nitrogen phase could be avoided. However, this would cause overestimation in the overall mechanical properties.

Figure 11 also shows that B-C-N compositions with a 1 : 1 B:N ratio can result in several superhard compounds with hardness > 60 GPa. For example, the predicted hardness values of BC_2N and BC_4N by machine learning are 74 and 65 GPa,

respectively. These predictions are consistent with previous experimental findings of superhardness in BC_2N (76 or 62 GPa) [41, 42, 43] and BC_4N (68 GPa) [43], which are synthesized under high-pressure and high-temperature conditions.

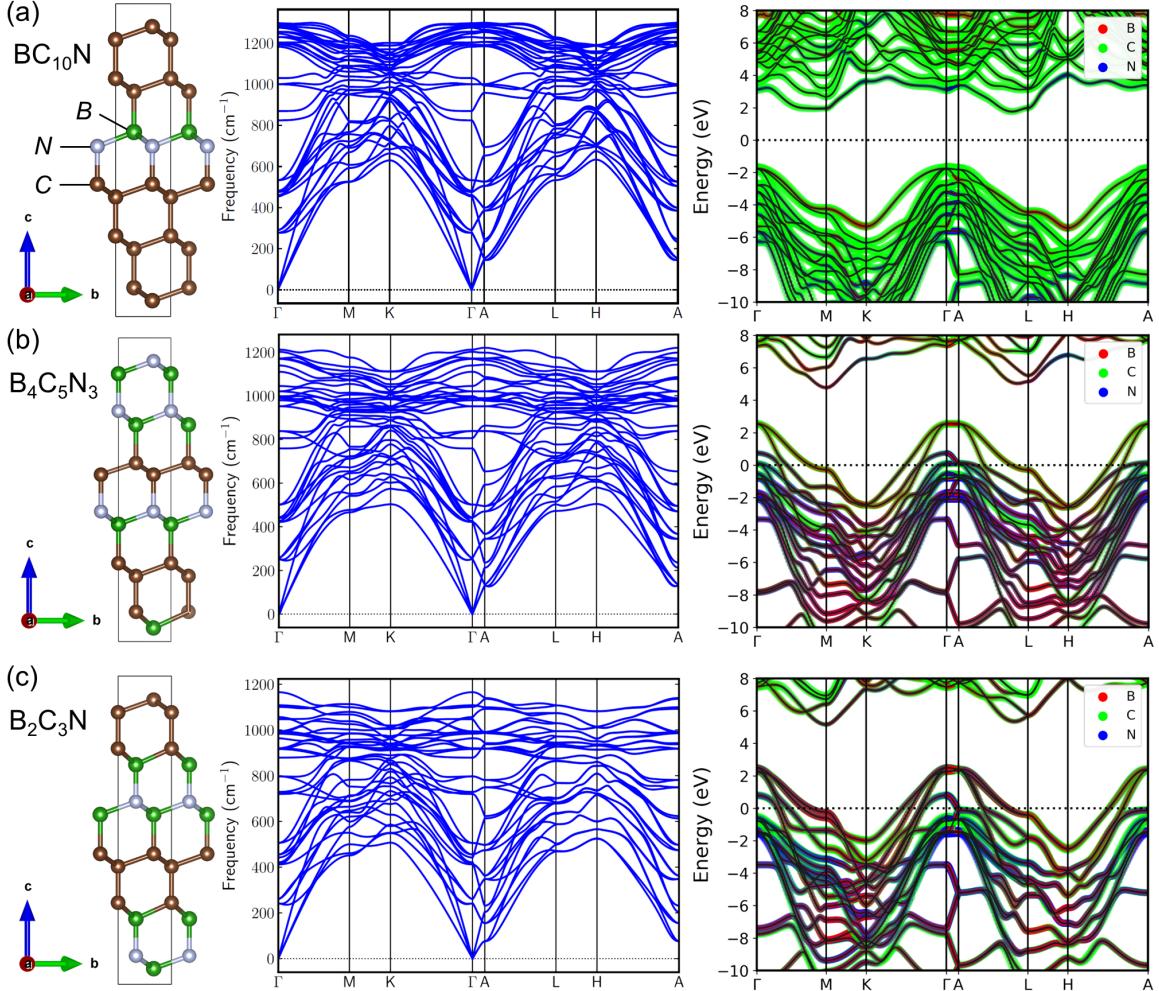


Figure 12: Theoretical crystal structures (left panels), phonon dispersion spectra (middle panels), and electronic band structures (right panels) from evolutionary algorithm and density functional theory calculations for (a) BC_{10}N , (b) $\text{B}_4\text{C}_5\text{N}_3$, and (c) $\text{B}_2\text{C}_3\text{N}$. BC_{10}N is a wide-band-gap insulator, while $\text{B}_4\text{C}_5\text{N}_3$ and $\text{B}_2\text{C}_3\text{N}$ are both metals. All three compounds are dynamically stable (i.e. without negative phonon modes). The crystal structures are visualized by the VESTA software [145].

Motivated by the machine learning results in Fig. 11, we next employ evolutionary prediction with USPEX to search for potential superhard structures around the region with a B:N ratio $\sim 1:1$. The calculations are performed under an applied pressure of 15 GPa to help locate stable structures of smaller volumes and larger hardness.

We first consider 15 trial chemical formulae with even number of valence electrons, including BC_3N , BC_5N , BC_6N , BC_3N_2 , $\text{B}_2\text{C}_3\text{N}$, $\text{B}_4\text{C}_5\text{N}_2$, etc., with a single-formula unit cell. However, most of the structures we found are graphite-like structures with sp^2 bonding, so they are not superhard. On the other hand, we find a diamond-like structure with sp^3 bonding for $\text{B}_2\text{C}_3\text{N}$ [Fig. 12(b)], which exhibits a hardness value > 40 GPa and a relatively low formation energy as discussed later.

$\text{B}_2\text{C}_3\text{N}$ is hexagonal with a superlattice structure along the (111) direction of cubic diamond. By comparing the $1 \times 1 \times 2$ supercell of $\text{B}_2\text{C}_3\text{N}$ (i.e. $\text{B}_4\text{C}_6\text{N}_2$), we find that such structure is similar to the $1 \times 1 \times 2$ BC_4N [146] (i.e. $\text{B}_2\text{C}_8\text{N}_2$) with 2 carbon atoms replaced by 2 boron atoms, or the structure of $1 \times 1 \times 3$ BC_2N [147] (i.e. $\text{B}_3\text{C}_6\text{N}_3$) with 1 nitrogen replaced by 1 boron. In fact, such atomic replacement is also the case of boron-substituted diamond BC_5 [148] ($H \sim 70$ GPa) in a 12-atom unit cell, with 2 carbon atoms replaced by 2 boron atoms. If one further replaces a carbon by boron in BC_5 , the resulting B_2C_4 (BC_2) [149] structure is also superhard ($H \sim 56$ GPa). Similarly, based on the 12-atom unit cell of diamond, the aforementioned superhard structure of BC_2N [147] (BC_4N [146]) also can be generated by replacing 6 (4) carbons with 3 (2) BN pairs.

Using a 12-atom unit cell with a 1:1 B:N ratio, we first create the structure of BC_{10}N [Fig. 12 (a)] by replacing 2 carbon atoms in diamond by 1 pair of BN. Our random forests models predict that BC_{10}N has bulk and shear moduli equal to $K_{RF} = 379$ GPa and $G_{RF} = 422$ GPa, respectively, which corresponds to a hardness $H_{RF} = 75$ GPa by Tian's model. Using a similar rule, we also generate a new superhard composition $\text{B}_4\text{C}_5\text{N}_3$, by replacing 1 carbon with 1 boron in $\text{B}_3\text{C}_6\text{N}_3$. $\text{B}_4\text{C}_5\text{N}_3$ is predicted to have $K_{RF} = 359$ GPa and $G_{RF} = 369$ GPa, with $H_{RF} = 62$ GPa. Other superhard B-C-N compounds also could be generated in a similar way. For example, $\text{B}_3\text{C}_7\text{N}_2$ and $\text{B}_2\text{C}_9\text{N}$ could be obtained respectively from $\text{B}_3\text{C}_6\text{N}_3$ and $\text{B}_2\text{C}_8\text{N}_2$ by atomic substitution. Before we shift the focus to first-principles DFT validation of

Table 2: Mechanical properties calculated by density functional theory (DFT), including density ρ (atom/ \AA^3), bulk modulus K (GPa), shear modulus G (GPa), Young’s modulus E (GPa), Pugh’s ratio k , Poisson’s ratio ν , universal elastic anisotropy A^U , hardness H (GPa), and formation energy ΔE (meV/atom).

Formula	ρ	K	G	E	k	ν	A^U	H_{RF}	H_{DFT}	ΔE
BC_{10}N	0.173	417	487	1052	1.166	0.080	0.058	75	87	79.5
$\text{B}_4\text{C}_5\text{N}_3$	0.166	370	368	829	0.995	0.125	0.221	62	60	141.3
$\text{B}_2\text{C}_3\text{N}$	0.164	354	298	697	0.840	0.172	0.792	49	43	155.9
diamond	0.175	432	518	1110	1.199	0.072	0.044	90	94	0

machine learning results, some comments are in order: (i) We have considered boron substitution in a 8-atom unit cell of BC_2N (i.e. $\text{B}_2\text{C}_4\text{N}_2$) to obtain $\text{B}_3\text{C}_3\text{N}_2$. However, we find that B-C-N compounds in a 8-atom unit cell have higher formation energies > 300 meV/atom, which is consistent with early study on BC_2N by Chen *et al.* [150]. (ii) We did not consider nitrogen substitution, because such structures tend to be thermodynamically more unstable.

We next discuss DFT calculations of three new superhard B-C-N phases predicted by machine learning: BC_{10}N , $\text{B}_4\text{C}_5\text{N}_3$, and $\text{B}_2\text{C}_3\text{N}$. The structure of $\text{B}_2\text{C}_3\text{N}$ [Fig. 12(c)] is discovered by USPEX using a 6-atom unit cell. For BC_{10}N [Fig. 12(a)] and $\text{B}_4\text{C}_5\text{N}_3$ [Fig. 12(b)], they are generated by using the aforementioned rule of atomic substitution in a 12-atom diamond unit cell. We have performed additional USPEX calculations for BC_{10}N and $\text{B}_4\text{C}_5\text{N}_3$, but did not find any lower enthalpy structure. The structures in Fig. 12 can be treated as superlattices along the (111) direction of cubic diamond. All structures are trigonal systems with the hexagonal lattice space group $P3m1$ (No. 156), which has 6 independent elastic constants C_{ij} : C_{11} , C_{12} , C_{13} , C_{14} , C_{33} , C_{44} , ($C_{66} = (C_{11} - C_{12})/2$). There are 4 necessary and sufficient mechanical stability conditions based on Born’s criteria [151]: $C_{11} > C_{12}$, $C_{44} > 0$, $C_{13}^2 < \frac{1}{2}C_{33}(C_{11} + C_{12})$, and $C_{14}^2 < \frac{1}{2}C_{44}(C_{11} - C_{12})$. The three structures in Fig. 12 all fulfill these criteria, so they are mechanically stable.

Figure 12 middle panels show phonon spectra for the corresponding structures in the left panels, and all structures are dynamically stable without negative modes. Their phonon dispersions are fairly similar, due to the similarity in the crystal structures. Among them, BC_{10}N has the highest phonon frequency above 1250 cm^{-1} , while $\text{B}_4\text{C}_5\text{N}_3$ and $\text{B}_2\text{C}_3\text{N}$ have slightly lower phonon frequencies near 1200 cm^{-1} at the top of the phonon bands. We note that cubic diamond has the highest phonon frequency above 1300 cm^{-1} . The results indicate that phonon bands are softened with increasing B/N content. This phonon softening is consistent with the phenomenon observed in boron-incorporated diamond BC_5 [40].

The electronic band structures are shown accordingly in Fig. 12 right panels. BC_{10}N exhibits a wide band gap $\sim 3.5 \text{ eV}$, so it is a superhard insulator. On the other hand, the electron-deficit $\text{B}_4\text{C}_5\text{N}_3$ and $\text{B}_2\text{C}_3\text{N}$ are metals, where their valence band maximums are shifted towards the conduction bands. Due to the similarity of phonon and electronic dispersion relations between $\text{B}_4\text{C}_5\text{N}_3/\text{B}_2\text{C}_3\text{N}$ and BC_5 [40, 148, 152], superconductivity may be observed in the predicted $\text{B}_4\text{C}_5\text{N}_3$ and $\text{B}_2\text{C}_3\text{N}$ compounds as well. However, discussion related to superconducting properties is beyond the scope of this study.

The elastic constants C_{ij} computed by DFT for the structures in Fig. 12 can be utilized to derive other mechanical properties, such as the bulk modulus (K), shear modulus (G), and Young's modulus (E). The Vicker's hardness H also can be calculated by Eq.(1) with K and G as input. Details of the calculation results are given in Table 2. Notably, BC_{10}N has supreme mechanical properties with a hardness of 87 GPa, which is comparable to diamond. Once synthesized, BC_{10}N would be the second known hardest material.

Table I also indicates that when the B/N content increases, the mechanical strengths like bulk and shear moduli as well as hardness will tend to decrease, which is consistent with the trend of phonon softening. The computed Cauchy pressures

($= C_{12} - C_{44}$) [153] of BC_{10}N , $\text{B}_4\text{C}_5\text{N}_3$, and $\text{B}_2\text{C}_3\text{N}$ are all negative (and equal to -405 GPa, -209 GPa, and -121 GPa, respectively), which suggests their brittle properties and strong covalent bondings. In addition, the Pugh's ratio ($k = G/K$) [141] is greater than 0.571 for all three ternary compounds, indicative of their brittle properties as well. By increasing the B/N content, other B-C-N compounds like $\text{B}_4\text{C}_5\text{N}_3$ and $\text{B}_2\text{C}_3\text{N}$ can become more ductile. Table I also demonstrates that the density ρ (atom/ \AA^3) is positively correlated with mechanical strength, while the poisson's ratio ν and universal elastic anisotropy A^U [154] have negative correlations with elastic moduli and hardness.

Finally, we evaluate the thermodynamic stability of the newly proposed $\text{B}_x\text{C}_y\text{N}_z$ compounds, by calculating the formation energy ΔE (also shown in Table 2):

$$\Delta E = \frac{E(\text{B}_x\text{C}_y\text{N}_z) - yE(\text{C}) - zE(\text{BN}) - (x - z)E(\text{B})}{x + y + z}, \quad (54)$$

which is the difference between the total energy E of $\text{B}_x\text{C}_y\text{N}_z$ and the atomically weighted reference total energies E of diamond, cubic BN, and α -B. The formation energies of the three compounds are positive, which is consistent with previous study on BC_2N [150], suggesting that ΔE increases due to B-C and C-N bondings. Among the three proposed compositions, BC_{10}N has the least number of B-C and B-N bonds, and it has the lowest formation energy < 100 meV/atom. Since ΔE of BC_{10}N is smaller than those of BC_2N and BC_4N , it is likely that BC_{10}N can be synthesized without extreme conditions [155], using e.g. low-temperature plasma methods.

To conclude this section, we have built random forest models to predict bulk and shear moduli by using target elastic properties in the Materials Project database [109]. The machine learning models utilize only materials features that can be derived directly from a given chemical formula, so they are suitable for large-scale materials characterization and discovery. We have applied the resulting models to B-C-N com-

pounds to search for new superhard ternary materials. The machine-learning predicted ternary graphs indicate that a 1:1 B:N ratio can lead to various promising superhard materials with hardness > 40 GPa. We also have utilized evolutionary structure prediction together with first-principles density-functional-theory calculations to further validate the machine learning results. We have proposed three new potential superhard ternary compounds – BC_{10}N , $\text{B}_4\text{C}_5\text{N}_3$, and $\text{B}_2\text{C}_3\text{N}$ – and fully characterized their properties. In predicted ternary compounds, BC_{10}N is a wide band-gap semiconductor, while $\text{B}_4\text{C}_5\text{N}_3$ and $\text{B}_2\text{C}_3\text{N}$ show metallic behavior. Among them, BC_{10}N has a hardness value ~ 87 GPa and a relatively low formation energy. Therefore, BC_{10}N may be synthesized without high-pressure and high-temperature conditions, for example using low-temperature plasma methods. Once synthesized, BC_{10}N would become the second known hardest material with a wide range of potential applications in extreme environments.

This section is based on our study published in *npj Comput Mater* **7**, 114 (2021) [156].

B-N-O Superhard Materials

B-N-O compounds have attracted considerable attention because it shows potential to several research areas including chemical adsorption [157], fluorescent dots [158], water splitting [159], wide bandgap insulator with large dielectric strength [160], as well as electrochemical applications [161]. Although the synthesized B-N-O compounds have several forms including amorphous, porous, layered, thin-film, and nano-particle, our knowledge to the mechanical properties of bulk B-N-O are still limited. Recently, Bhat *et al.* employed HPHT methods to synthesize crystalline novel B-N-O compounds with zinc-blende-like structure [53]. In their previous experiments for the synthesis of B-C-N compounds, they found the initial materials tended to decompose into diamond and c-BN. Moreover, they found some crystals contained considerable amount of oxygen as well as boron and nitrogen, which motivated their HPHT synthesis for B-N-O compounds. They used hexagonal BN and B_2O_3 in a ratio of 3 : 1 as the initial materials. Afterwards, the crystalline B-N-O material with composition ratio of 6:4:3 was synthesized under HPHT condition.

Besides, Bhat *et al.* also conducted crystal structure prediction and density functional theory characterization, their computational results showed that the most stable structures contained ordered vacancies in zinc-blende type structure, and such structures have a high bulk modulus of 300 GPa, which indicates that B-N-O compounds might have superhardness [53]. Motivated by the experimental findings by Bhat *et al.*, here we again use evolutionary crystal structure prediction, random forests models, and density functional theory to search for new superhard materials in ternary B-N-O compounds. However, we found that the number of B-N-O data in Materials Project database is small, which reflects that crystalline B-N-O compounds are not fully explored. Moreover, the lack of relevant data might lead to inaccurate machine learning prediction. Therefore, here we adopt another study procedure based on the data generated from our evolutionary searches.

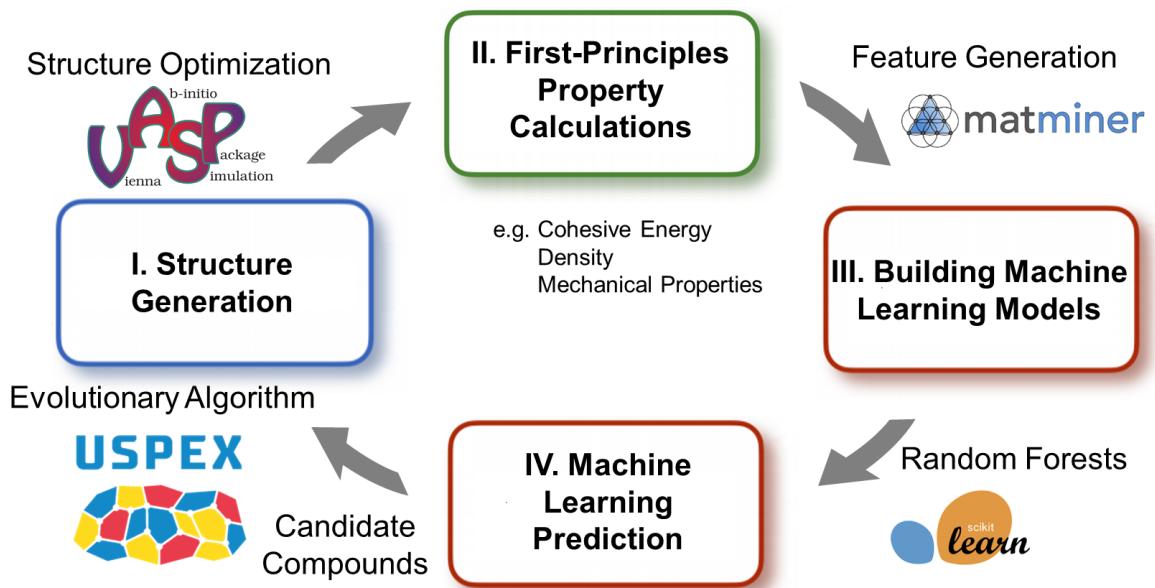


Figure 13: The self-improvement recursive procedure used in our B-N-O study. This procedure helps us gradually improve our ML models and locate atomic compositions of superhard B-N-O compounds. **I.** Structure generation as data generation for machine learning. **II.** First-principles property calculations as targets including cohesive energy, density, and mechanical properties. **III.** Building random forests models based on the features generated by MATMINER library. **III.** Machine learning prediction by mapping out the properties in triangular plots. The promising B-N-O materials are then selected in next round of structure generation.

We first introduce the new study procedure for B-N-O compounds. The new flowchart is shown in Fig. 13. To achieve high-accuracy ML prediction for exploring B-N-O materials, we first generate B-N-O data through evolutionary crystal structure prediction using DFT enthalpy as the fitness. Here we adopt a recursive process shown in Fig. 13. to gradually improve our ML models and to locate atomic compositions of superhard B-N-O compounds. This process involves four steps: (i) Structure generation, (ii) First-principles property calculation, (iii) Building ML models, and (iv) ML prediction. First, we use evolutionary algorithm for CSP to create the first generation B-N-O data. In the evolutionary searches, we conducted two separate evolutionary searches for each chemical formula, and the optimized structures at the end of each search become a data point. Next, we use DFT to calculate physical properties including cohesive energy, density, and elastic tensor. With the elastic tensor, we can calculate the bulk modulus and shear modulus, and also Vickers hardness. Besides, we can also evaluate the elastic stability by examining whether the elastic tensor is positive-definite.

We build ML models for the cohesive energy, density and Vickers hardness. Afterwards, we carry out ML prediction by mapping out the aforementioned properties in triangular plots. The cohesive energy is related to the formation energy if using isolated atoms' total energies as reference points. Therefore, we use cohesive energy to evaluate the thermodynamic stability. Based on the ML prediction of the cohesive energy and hardness, we select promising compounds with high thermodynamic stability and high hardness to be the candidates of next generation data. In this study, we use four data generations to explore superhard B-N-O compounds. In the following, we elaborate the computational details of CSP, DFT calculations, and ML.

Crystal structure prediction as Data Acquisition – We generate our B-N-O data through CSP using DFT enthalpy as the fitness. The purpose of CSP is to find the stable or metastable structures of a compound knowing only the chemical for-

mula [130, 131, 132]. In our B-N-O study, we use evolutionary algorithm implemented in USPEX package [93, 94, 95], and the new structures are generated by heredity (50%), mutation (20%), and random (10%) operators. During the data generation for ternary B-N-O compounds, the number of atoms in an unit-cell is between 9 to 19 atoms. During our search, we only consider Since our study focuses on superhard compounds, an external pressure $P = 15$ GPa is added to avoid graphite-like, layered structures. For each atomic composition, we carry out two separate USPEX searches, and more than 1200 structures are explored for each. Upon completion, the optimized structures of both USPEX searches are further relaxed without external pressure. In the end, we have two data points for a single chemical formula.

Density functional theory (DFT) calculations – Our DFT calculations are performed by VASP software [112, 113], which adopts the pseudopotential method and plane-wave basis sets. We use the projector augmented wave (PAW) [134, 135] method and generalized gradient approximation (GGA) exchange-correlation energy functional based on Perdew-Burke-Ernzerhof (PBE) [162] formalism. The kinetic energy cutoff for wavefunction expansion is 520 eV, and k-points are sampled by a Γ -centered Monkhorst-Pack mesh with a resolution $\sim 0.02 \times 2\pi/\text{\AA}$. The convergence criteria of electronic self-consistency and structural relaxation are set to 10^{-6} eV/unit cell and 10^{-3} eV/ \AA , respectively.

The strain-stress method [136] implemented in VASP is used to compute the elastic constants C_{ij} , which in turn allows us to evaluate the elastic stability and mechanical properties. In our calculations, we carried out properties including eigenvalues of elastic tensor, Voigt-Reuss-Hill (VRH) average [137, 138, 139] for bulk/shear modulus, and Vickers hardness by Chen’s model [128] directly through MechElastic python library [163, 164]. Phonon density of states are computed using the PHONOPY package [140] through density functional perturbation theory (DFPT) scheme implemented in VASP, and all of the convergence tests are examined carefully.

Machine Learning models and prediction – After the preparation of ML data, we build our ML models for cohesive energy, density, and hardness. Since the quality of ML data determines the predictive capability, we first exclude elastically unstable structures in the data set by examining whether the elastic tensor is positive-definite. Next, we create features or descriptors to train our ML model using MATMINER [118] python library. For the prediction of cohesive energy, we adopt the composition features proposed by Meredig *et al.* [102]. In our B-N-O study, a total number of 16 features are considered: fraction of B/N/O, mean of atomic weight, mean column, mean row, range of atomic number, mean of atomic number, range of atomic radius, mean of atomic radius, range of electronegativity, mean of electronegativity, average of s/p electron, fraction of s/p electron. For the hardness model, we add one extra structural property, density, to improve the model performance. As a result, we also build the density model using 16 compositional features to evaluate the density for hardness prediction. The adopted ML algorithm in our study is random forests, which is implemented in SCIKIT-LEARN python library [111].

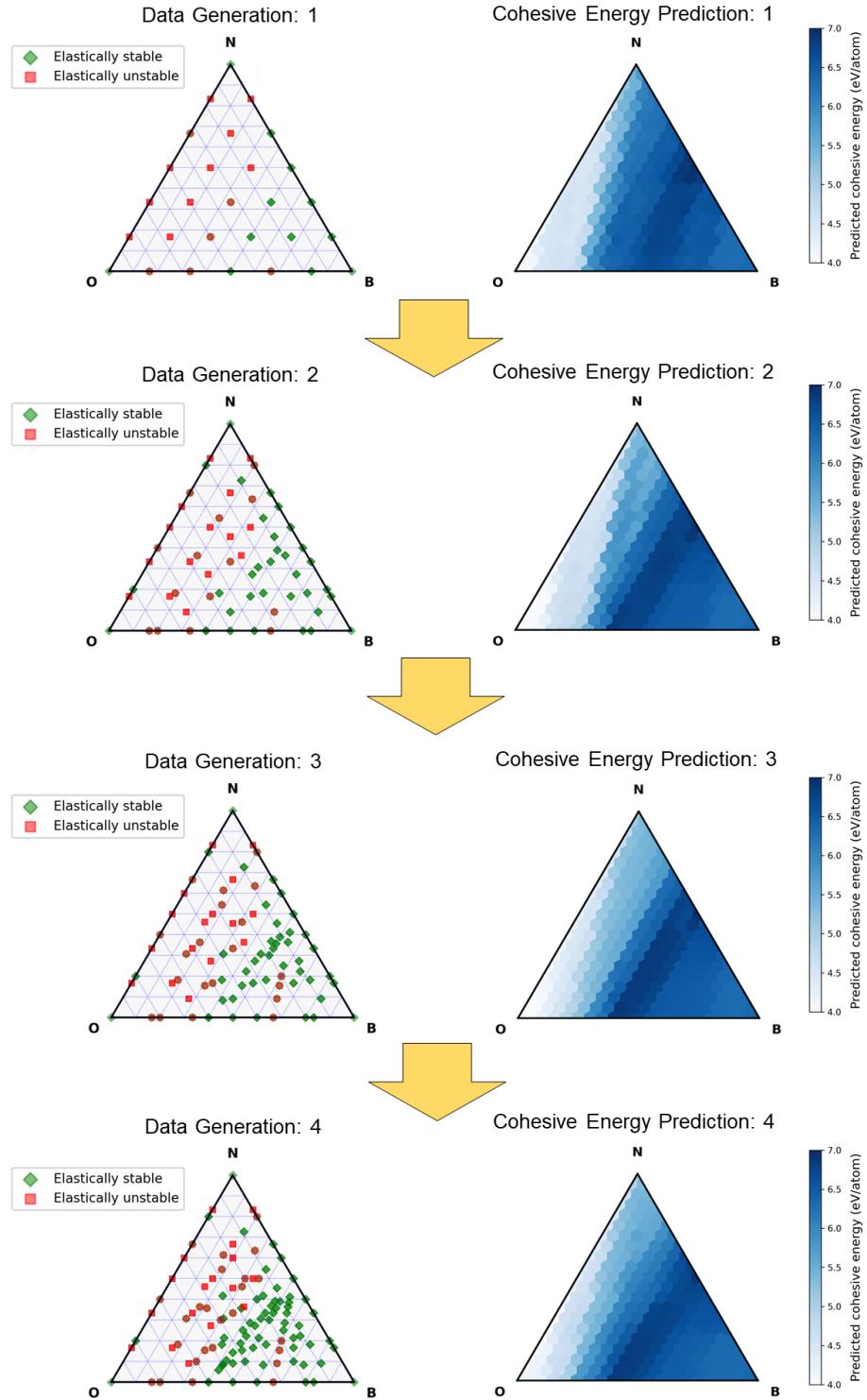


Figure 14: Data distribution (top) and predicted cohesive energy (bottom) in composition space of B-N-O systems. In the top row, the green diamond and red rectangle represent the elastically stable and unstable structures from evolutionary algorithm searches.

In Fig. 14, we show the evolution of our machine data generated by USPEX package and the corresponding random forests models for cohesive energy. In the beginning, we have an uniform sampling with a number of atoms equal to 12 for binary and ternary compounds. (The elemental B, N and O are based on the known structures.) For example, pure B here is a α -B₁₂. In the first generation sampling, we have found the rate between elastically stable and unstable structures is roughly 1:1. Although the grid is course, we can see that if a ternary compound is composed of mostly nitrogen and oxygen atoms, they tend to be elastically unstable. On the other hand, if the boron content is high, then such systems are more likely to be elastically stable. To ensure high accuracy of ML prediction, we only consider elastically stable structures to build our ML models. In these four generations, the numbers of (elastically stable) data involved in building ML models are 27, 78, 104, and 153, respectively.

From cohesive energy prediction of first generation, we notice that when the boron content is large, the cohesive energy is also high, indicating higher thermodynamic stability. Therefore, the behavior of elastic stability and thermodynamic stability are consistent with each other. Besides, based on the such predicting results, starting from data generation 2, we neglect the composition with simultaneous high nitrogen and high oxygen contents. In the subsequent generations, we manually select more compositions based on the ML prediction of previous generation. In particular, we gradually increase the compositions with high predicted cohesive energy and predicted hardness. We find that the machine learning prediction results have higher and higher resolution during this process. Importantly, we found that in cohesive energy prediction, a region connecting BN and B₂O₃ have high cohesive energy. This result is consistent with the recently synthesized B₆N₄O₃ by Bhat *et al.*, where a mixture of hexagonal BN and B₂O₃ is used as starting materials for HPHT treatment.

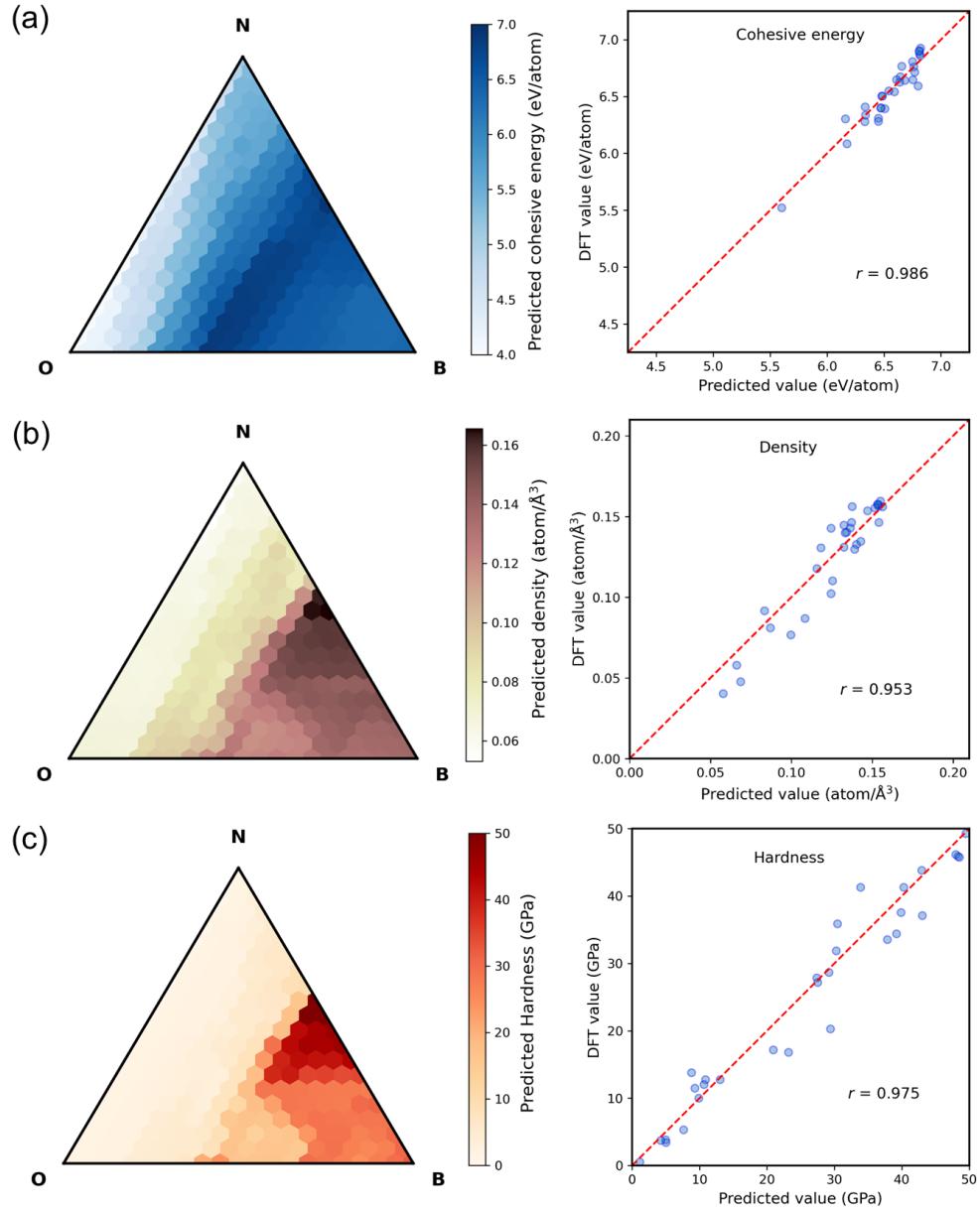


Figure 15: Machine learning prediction of (a) cohesive energy, (b) volumetric density, and (c) hardness using the data of generation 4 (left column) and corresponding model evaluation using test set (20 % of data) (right column).

Fig. 15 shows the predicting results of our random forests models for cohesive energy, density, and hardness built from data generation 4. Here, 80 % of data are used to construct random forests models (for training and validation), and the remaining 20 % of data are used for final evaluation. Our cohesive energy model shows a Pearson correlation coefficient $r = 0.98$, which demonstrates the accuracy of the model.

Based on our previous ML study for B-C-N compounds, we have noticed that the hardness value has a positive correlation with volumetric density. Therefore, here we also consider density as a feature to build our hardness model. The hardness values here are computed using Chen's model [128]

$$H_V = 2(k^2 G)^{0.585} - 3, \quad (55)$$

where $k = G/K$ is pugh's ratio, G is shear modulus, and K is bulk modulus.

In order to map out the hardness values of ternary B-N-O compounds, we first build a density model based on compositional features. The built density model is used to predict density, which in turn becomes the density feature for hardness prediction. As shown in Fig. 15, the predicted density and hardness show a highly positive correlation with each other. The results show that ternary B-N-O compounds can have superhardness around a region near binary BN, which has a predicted hardness like cubic boron nitride. Following our study procedure shown in Fig. 13, now we can find the superhard B-N-O compounds directly by inspecting the data generated from evolutionary algorithm. The predicted ternary B-N-O compounds with superhardness are listed in Table 3. We found that the compositions of discovered superhard B-N-O compounds have two kinds. For the first kind, the number of boron is equal to the sum of nitrogen and oxygen (e.g. $B_5N_3O_2$). The second kinds has one more oxygen, so the number of boron is equal to the sum of nitrogen and oxygen plus 1

Table 3: Mechanical properties of superhard B-N-O compounds with cohesive energy > 6.75 eV discovered in this study. Density ρ (atom/ \AA^3), bulk modulus K (GPa), shear modulus G (GPa), cohesive energy E_{coh} (eV/atom), and formation energy E_{form} (meV/atom).

Crystal	ρ	K	G	H	E_{coh}	E_{form}
$\text{B}_5\text{N}_3\text{O}_2$	0.151	286	257	42	6.761	
$\text{B}_5\text{N}_3\text{O}_3$ (WZ)	0.150	279	255	43	6.935	83
$\text{B}_6\text{N}_4\text{O}_2$	0.156	322	284	44	6.810	
$\text{B}_6\text{N}_4\text{O}_3$ (ZB)	0.155	292	268	45	6.900	120
$\text{B}_7\text{N}_5\text{O}_2$	0.146	272	241	40	6.776	
$\text{B}_7\text{N}_5\text{O}_3$ (ZB)	0.158	306	283	47	6.904	117
$\text{B}_9\text{N}_7\text{O}_2$	0.161	330	296	46	6.802	
$\text{B}_9\text{N}_7\text{O}_3$ (ZB)	0.160	318	301	50	6.926	97
cubic BN	0.168	373	383	64	7.028	0
B_2O_3 (P3 ₁ 21)	0.101	35	33	12	7.008	0

(e.g. $\text{B}_5\text{N}_3\text{O}_3$). First, by examining the density of states, these materials are insulators with wide bandgaps greater than 4.5 eV based on modified Becke Johnson (mBJ) potential [165, 166]. To evaluate their thermodynamic stability, we can compare their cohesive energy,

$$E_{coh} = \frac{xE(B_{atom}) + yE(N_{atom}) + zE(O_{atom}) - E(B_xN_yO_z)}{x + y + z}, \quad (56)$$

which can be regarded as the formation energy with a minus sign using the total energies of isolated atoms as references. Therefore, higher cohesive energy means higher thermodynamic stability of a material. We found that the second kind have a higher cohesive energy than the first kind. Besides, we also find that the cohesive energies of c-BN and B_2O_3 are higher than discovered B-N-O compounds. This result implies that the discovered B-N-O compounds are metastable at ambient pressure

condition. Besides, the chemical formula of the second kind has a form: $B_{x+2}N_xO_3$, and this allows us to calculate the formation energy using the total energies of c-BN and B_2O_3 as references. Hence, the formation energy for the second kind can be computed by

$$E_{form} = \frac{E(B_{x+2}N_xO_3) - xE(c-BN) - E(B_2O_3)}{2x + 5}. \quad (57)$$

We found that $B_5N_3O_3$ has the lowest formation energy in our study, and this property might be related to its superlattice structure shown in Fig. 16(a). Besides, the report from Bhat *et al.* also found their ordered structure models are in better agreement with experiment experiment. As a result, the superlattice structure might exist in B-N-O compounds.

Next, we conduct more theoretical characterization for the more stable B-N-O compounds with their chemical formulae as $B_{x+2}N_xO_3$. In Fig. 16, we show crystal structure, electronic density of states (DOS), and phonon DOS for $B_5N_3O_3$, $B_6N_4O_3$, $B_7N_5O_3$, and $B_9N_7O_3$. From their crystal structure, only $B_5N_3O_3$ has a wurtzite-type structure, and the other three compounds have zinc-blende-type structure. We note that if comparing these chemical formulae with BN, we can find that the one boron atom is missing (due to B_2O_3), and this is associated with vacant boron sites in these materials. The vacant boron sites are shown in Fig. 17. By calculating the phonon DOS (shown in the right column of Fig. 16), our results ensure the dynamical stability in these B-N-O compounds.

We also calculate the electronic DOS of these four structures. We find that the three zinc-blende type structures have similar electronic DOS around Fermi level. The DOS of wurtzite-type $B_5N_3O_3$ is also similar, but it shows a very wide bandgap roughly 2 eV greater than zinc-blende type B-N-O compounds. This result is unusual because in binary BN structures, wurtzite BN has a smaller bandgap than cubic BN,

which is completely different from the ternary B-N-O compounds.

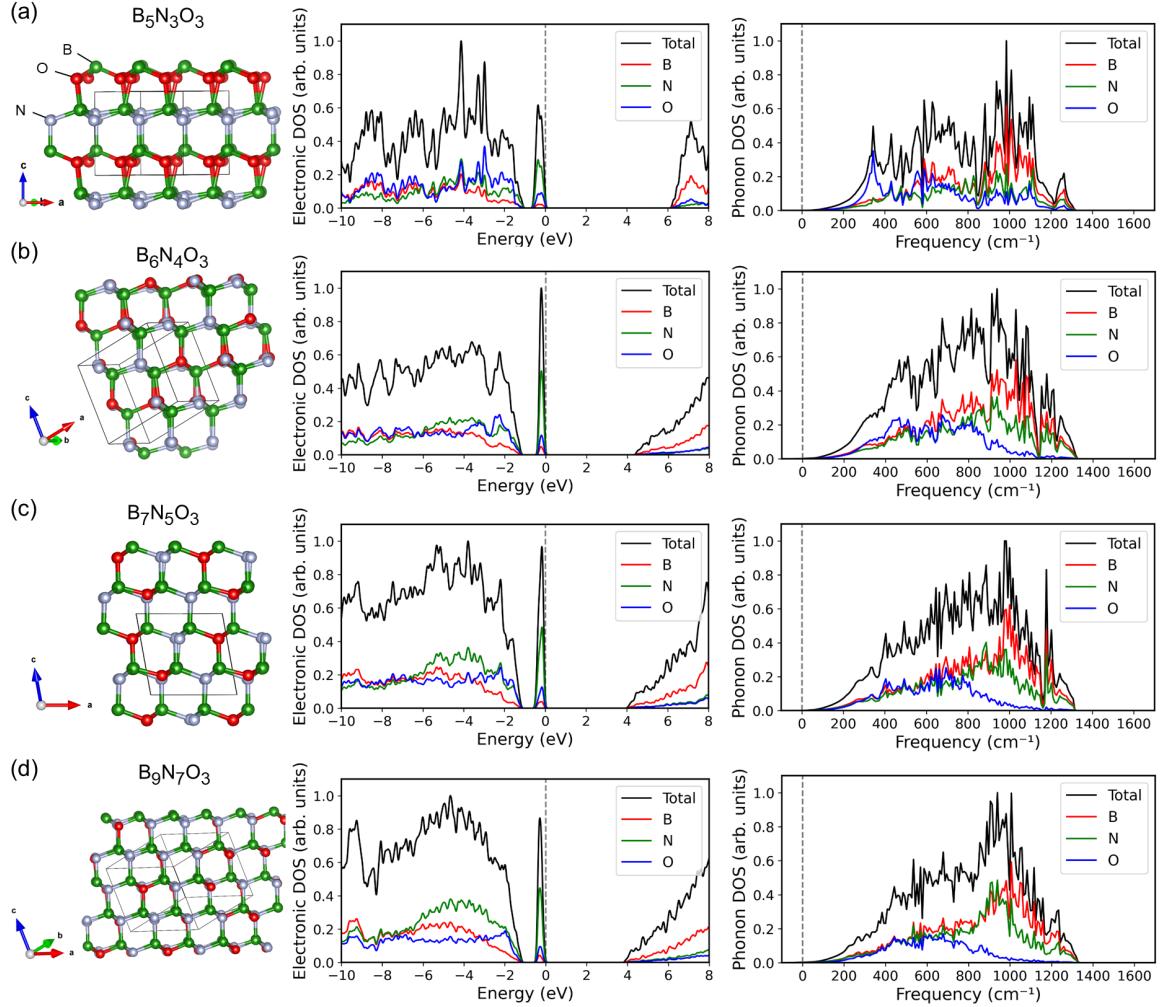


Figure 16: Crystal structure [left column], electronic density of states [middle column], and phonon density of states of superhard (a) $\text{B}_5\text{N}_3\text{O}_3$, (b) $\text{B}_6\text{N}_4\text{O}_3$, (c) $\text{B}_7\text{N}_5\text{O}_3$, and (d) $\text{B}_9\text{N}_7\text{O}_3$ discovered by evolutionary crystal structure prediction.

A remarkable thing to be mentioned here is that all of the density of states show a sharp peak at valence band maximum. This peak is contributed from the localized states of nitrogen atoms being adjacent to the vacant boron sites. The charge distribution of the peak in the DOS of $B_5N_3O_3$ is shown by the contour in Fig. 17. Besides, we find that in zinc-blende-type structures, the bandgap increase with increasing oxygen content. This result is reasonable because B_2O_3 has an extremely large band gap greater than 10 eV [167] comparing with the bandgap of 6 eV for c-BN. Overall, by introducing oxygen atoms in binary BN compounds, the electronic structures can be modified substantially, which might also lead to changes in the thermal and chemical stabilities.

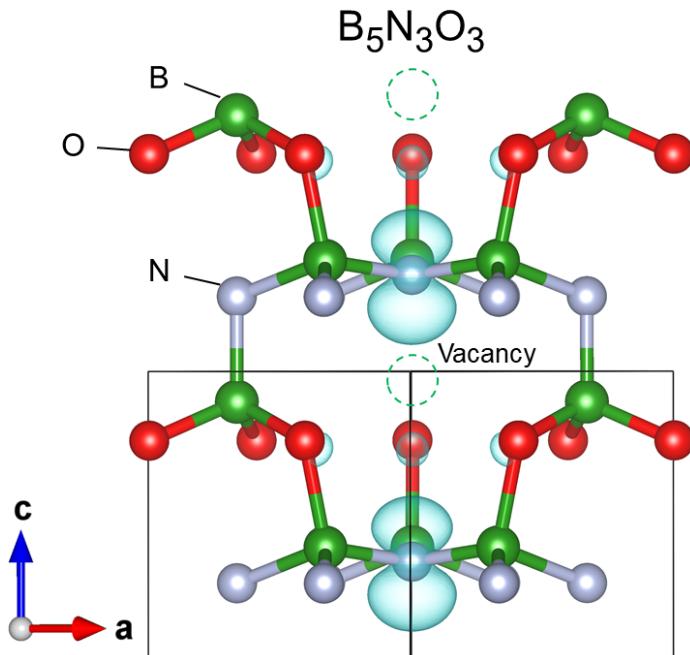


Figure 17: Crystal structure of $B_5N_3O_3$ highlighting the vacancies near three oxygen atoms and one nitrogen atom. The contour represents the electron density distribution of the electronic states at valence band maximum.

In summary, we have used a recursive study procedure to search for superhard B-N-O compounds using machine learning techniques, where we generated the data based on the evolutionary searches and density functional theory. Our results show that thermodynamically stable and superhard B-N-O materials have a chemical formula $B_{x+2}N_xO_3$ ($x \geq 3$), and their hardness values are around 45 GPa. Besides, that are all have very large bandgap > 4 eV based on PBE functional. Since the DFT-PBE results tend to underestimate bandgaps, the actual values of bandgaps should be even larger. Their electronic density of states all have a prominent peak around valence band maximum (VBM), which is corresponding to a localized state at nitrogen atoms around the vacant sites. Due to the change of electronic structure, the newly-found B-N-O materials might have a better performance than cubic BN in the humid condition at high temperature.

CHAPTER 4

STRUCTURAL AND TOPOLOGICAL PHASE TRANSITIONS IN TOPOLOGICAL QUANTUM MATERIALS

Topological Phase Transition in ZrTe₅

Topological materials with symmetry-protected boundary modes has fundamentally opened a entire new filed in condensed matter physics. Three dimensional band insulators with time-reversal (TR) symmetry can be classified into three kinds: normal insulators (NIs), weak topological insulators (WTIs), and strong topological insulators (STIs). Such classification is based on four topological Z_2 invariants or Z_2 indices. Here, our goal is to study topological phase transition by manipulating the band structure, which requires closing and reopening the bandgap. In order to achieve band inversion, it is easier to choose a material with small gap between valence and conduction bands. Zirconium pentatellurides ZrTe₅ has a mall direct band gap at Γ point, and hence is an ideal material to study topological phase transition.

ZrTe₅ is a van der Waals (vdW) layered material crystallized in the *Cmcm* orthorhombic space group as shown in Fig. 18(a). Each layer consists of ZrTe₃ (1 Te_a and 2 Te_d) chains extending along the x direction, and the layers are stacked along the y direction. The material has received substantial interest because its monolayer form is predicted to be a large bandgap quantum spin Hall insulator [170]. It was also suggested that the 3D bulk band structure is very close to the phase boundary between WTI and STI.

Early studies of optical conductivity [171], quantum oscillations [172, 173], photoemission, and negative longitudinal magnetoresistance (NLMR) [174] have supported

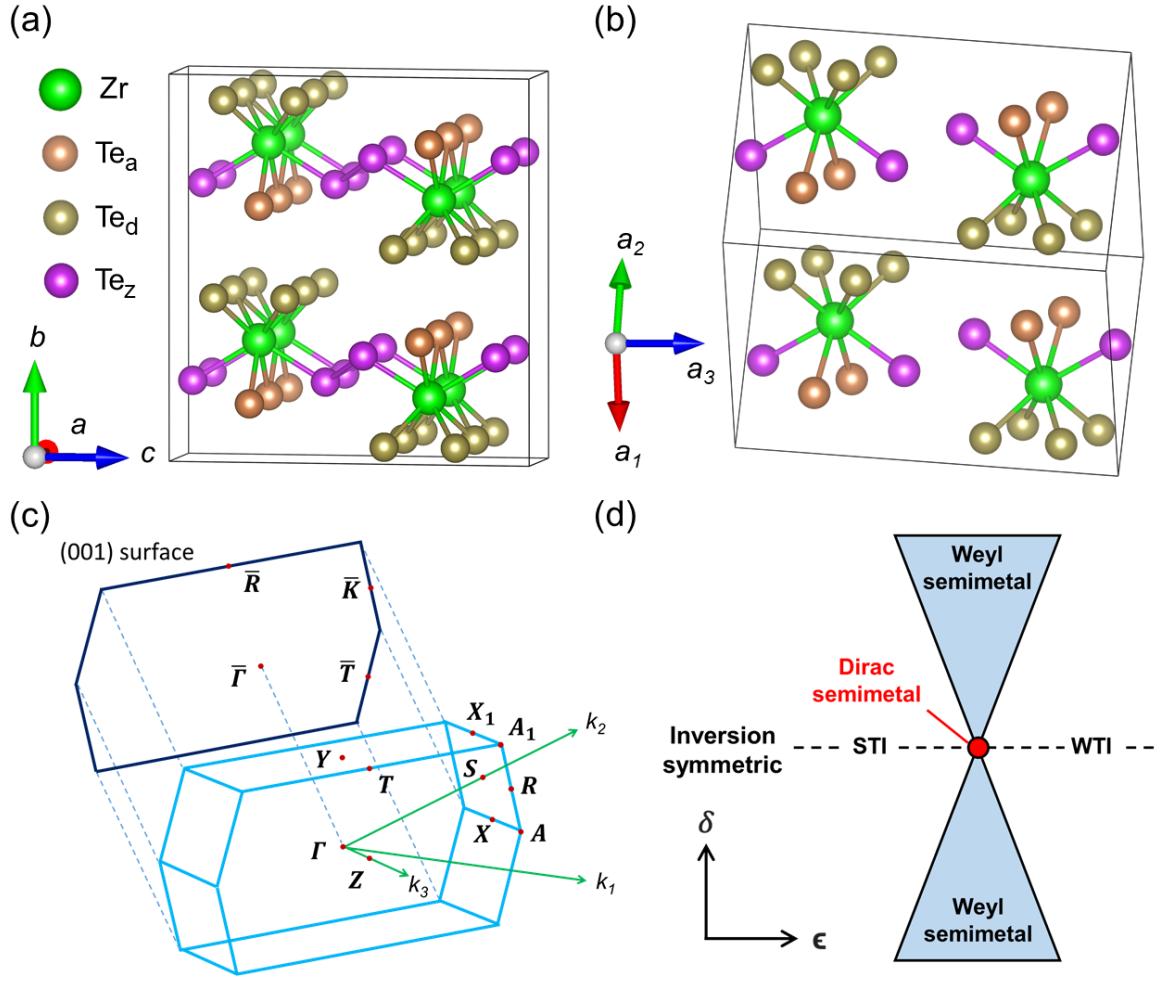


Figure 18: Crystal structure of ZrTe₅ in (a) conventional cell, and (b) primitive cell. (c) Bulk and surface Brillouin zones of ZrTe₅. (d) Universal phase diagram of topological insulators proposed by Murakami for a 3D system [168, 169]. The control parameter δ describes the breaking of inversion symmetry. The control parameter ϵ does not break inversion symmetry.

the predicted Dirac semimetal-like band structure in the bulk, with a single Dirac point at the center of the Brillouin zone. Unlike topological Dirac semimetals such as Na_3Bi or Cd_3As_2 , there is no additional crystalline symmetry to protect the Dirac point in ZrTe_5 , and more recent spectroscopy measurements revealed that its band structure is better described by a massive Dirac dispersion, where mass plays the role of bandgap. The bandgap size measured by different experiments varies, ranging from 10 to 80 meV, and there are conflicting reports on whether the material is a WTI or STI [175, 176, 177, 178], or changes with temperature [179]. These experimental findings suggest that the electronic structure of ZrTe_5 may be very sensitive to external perturbation, such as lattice distortion.

The relationship between these phases is summarized in the general phase diagram proposed by Murakami and Kuga [168, 169], as shown in Fig. 18 (d). These topological phases have been intensively studied in the past decade. In contrast, the transition between these phases is less explored because it requires changing the band parameters of the solid. It has been demonstrated that topological phase transitions can be induced either by chemical doping or by thermal lattice expansion [180, 181, 182, 183]. Nevertheless, the precise *in situ* control of topological phase transition in a three-dimensional (3D) system is still challenging.

Methods –The electronic structure and energy gap of ZrTe_5 under strain were calculated by the Quantum Espresso package [184] based on DFT. The Perdew-Burke-Ernzerhof exchange-correlation functional (PBE-GGA) [162] with spin-orbit coupling and projector augmented wave [134] method were used. The unstrained lattice parameters ($a = 3.97976 \text{ \AA}$, $c = 13.6762 \text{ \AA}$) were determined by experimental data at 10 K [185], and an 11×11 grid in the parameter ranges of $1.0a$ to $1.02a$ and $0.99c$ to $1.01c$ was considered for studying the gap behavior with lattice variation. An $8 \times 8 \times 4$ momentum grid was used in the self-consistent calculation, and the kinetic energy cutoff and convergence criterion were set to 30 and 10^{-7} rydberg, respectively.

The topological Z_2 indices of different strained structures were also computed to identify their topological nature. With input from the electronic structure calculations of Quantum Espresso, the maximally localized Wannier functions were first computed using the Wannier90 package [186], which, in turn, allowed the determination of Z_2 indices by tracking hybrid Wannier charge centers using the WannierTools package [187]. In this study, phase transition between the STI and WTI states was directly controlled by closing the energy gap at the Brillouin zone center.

Additional structure relaxation calculations were performed with the vdW-DFT [188, 189], corrected using the exchange-hole dipole moment model [190]. The vdW-DFT fully relaxed lattice parameters of layer-structured ZrTe_5 are within 1% error compared to the unstrained experimental data. To determine the Poisson ratio, conventional cells of different fixed lattice parameters along the a axis were considered, while the b and c lattice parameters were allowed to evolve freely in the structure relaxation calculations.

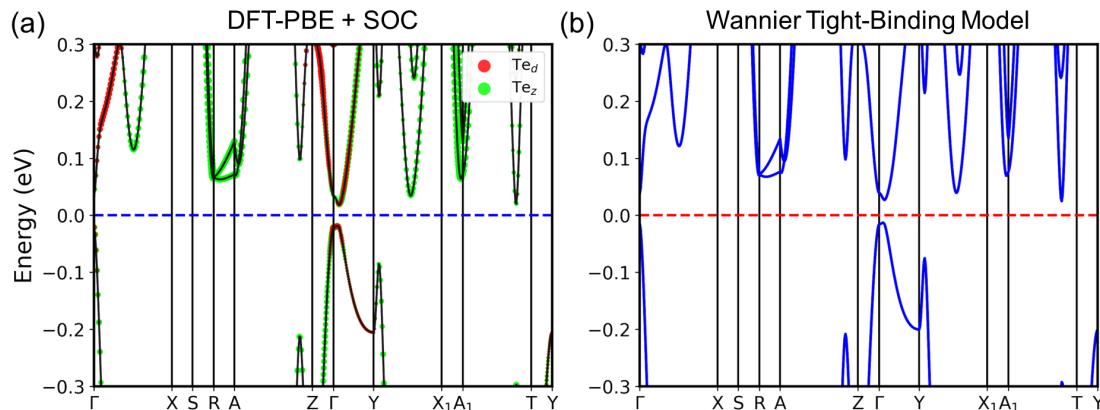


Figure 19: Band structures calculated by (a) DFT and (b) Wannier tight-binding model from Wannier90 package [186]. The high symmetry points can be found in Fig. 18(c). Here, the Wannier tight-binding model perfectly reproduces the DFT band structure around Fermi level.

In order to identify the Z_2 indices of ZrTe_5 , we first calculate the band structure (with spin-orbit coupling) based on experimental crystal structure [185], and we next use Wannier90 package [186] to obtain the corresponding tight-binding model. The

band structures calculated by DFT and the Wannier tight-binding model are shown in Fig. 19. The band structure based on Wannier tight-binding model is identical to DFT band structure around Fermi level. Therefore, we use the Wannier tight-binding model to compute the topological Z_2 indices using WannierTools package [187]. In WannierTools, it is capable of identifying the Z_2 indices by tracking hybrid Wannier charge centers [191]. In this method, first the hybrid Wannier function is defined as

$$|nk_x l_y\rangle = \frac{1}{2\pi} \int_0^{2\pi} dk_y e^{-ik_y l_y} |\psi_{nk}\rangle, \quad (58)$$

where $|\psi_{nk}\rangle$ is the Bloch wave function. The hybrid Wannier charge center (WCC) \bar{y}_n can be defined as a function of k_x by integrating over k_y :

$$\begin{aligned} \bar{y}_n(k_x) &= \langle nk_x 0 | y | nk_x 0 \rangle \\ &= \frac{i}{2\pi} \int_{-\pi}^{\pi} dk_y \left\langle u_{n,k_x,k_y} \left| \frac{\partial}{\partial k_y} \right| u_{n,k_x,k_y} \right\rangle, \end{aligned} \quad (59)$$

where $|u_{n\mathbf{k}}\rangle$ is the periodic part of Bloch function $|\psi_{n\mathbf{k}}\rangle$. One can find more information in Ref. [187].

Figure 20 shows the WCC evolution for six time-reversal invariant planes corresponding to the BZ of primitive cell. The strong index can be obtain by checking either (a), (b), or (c). One plane has a $Z_2 = 1$ and the other has a $Z_2 = 0$, so the strong index is 1. For the weak indices, by checking the right column of Fig. 20, we know $(\nu_1 \nu_2 \nu_3)$ is (110). Therefore, the band structure of ZrTe₅ shows a strong topological insulator phase. Our results are consistent with the calculations done by Fan *et al.* [192]. Here, we note that a small gap at Γ point gives us a chance to restore the band inversion between p orbitals of Te_d and TE_z, and this is corresponding to a topological phase transition from (1;110) to (0;110).

In order to achieve the topological phase transition by manipulating the lattice structure, it is important to know the response of the bandgap at Γ point during the

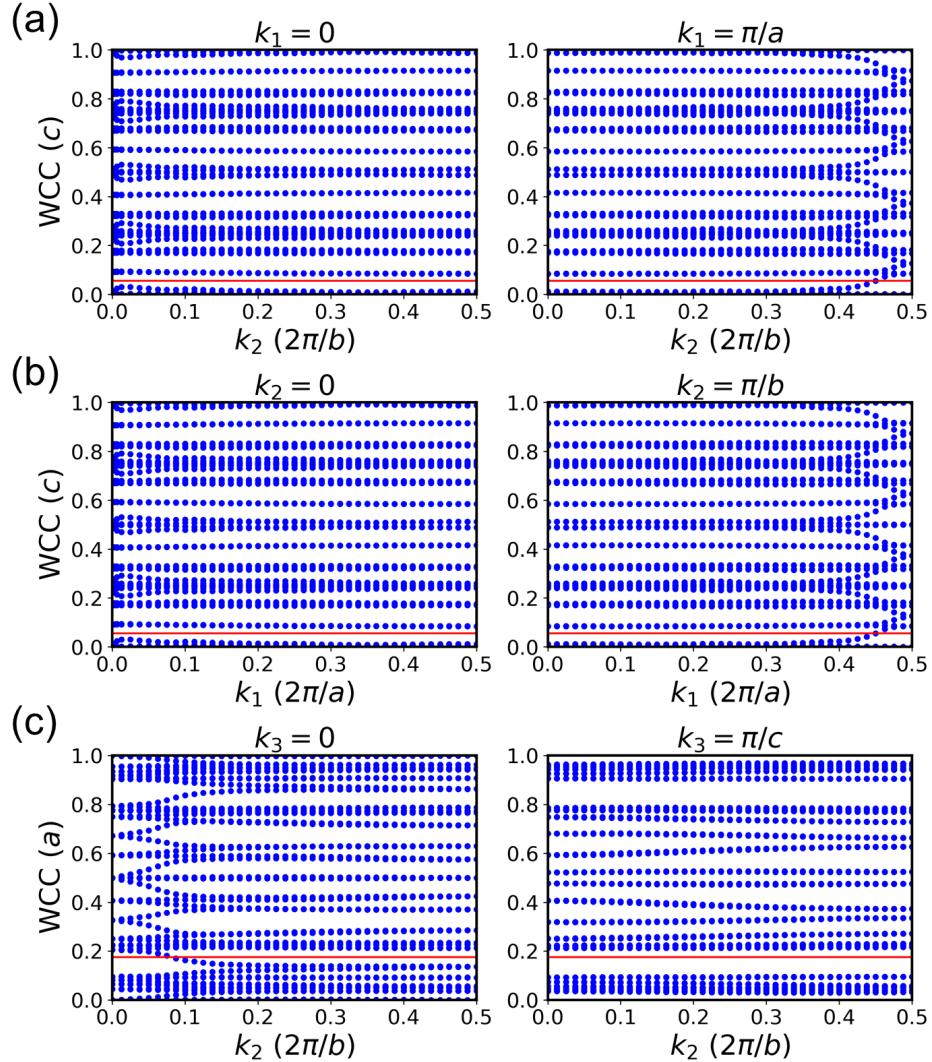


Figure 20: The evolutions of Wannier charge centers for six time-reversal invariant planes of ZrTe_5 . (a) $Z_2 = 0$ for $k_1 = 0$ and $Z_2 = 1$ for $k_1 = \pi/a$. (b) $Z_2 = 0$ for $k_2 = 0$ and $Z_2 = 1$ for $k_2 = \pi/b$. (c) $Z_2 = 1$ for $k_3 = 0$ and $Z_2 = 0$ for $k_3 = \pi/c$. Therefore, ZrTe_5 is a strong topological insulator with Z_2 indices = (1;110).

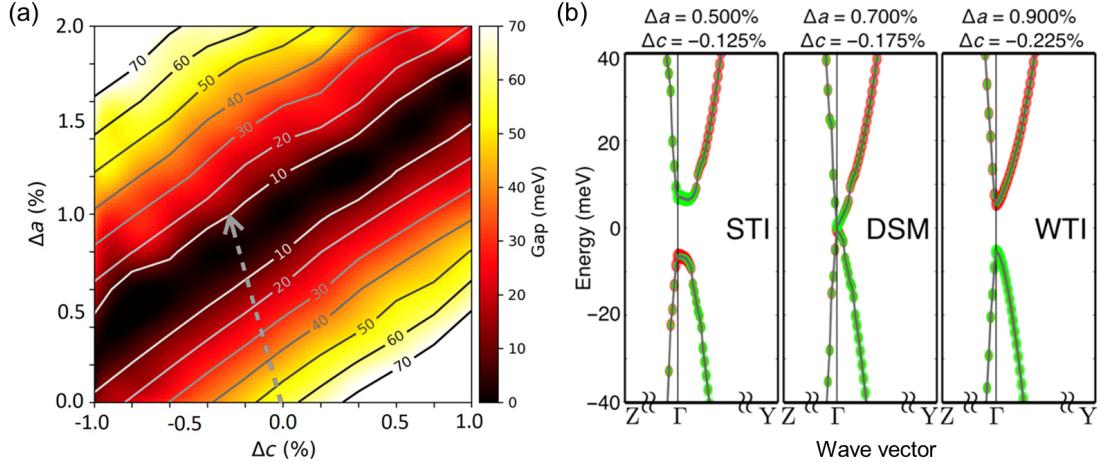


Figure 21: (a) Bandgap E_g at the Γ point as functions of strains in a and c lattice directions. The dashed gray arrow indicates the anisotropic strain induced by a uniaxial stress along the a axis direction (b) Band structures for different strain states taken at points along the Poisson’s ratio path. These points (from left to right) correspond to STI, Dirac semimetal, and WTI, respectively. A band inversion involving Te_d and Te_z p orbitals (shown as red and green colors, respectively) is seen in the STI phase.

change in of lattice parameters. Fig. 21 (a) shows a contour map of the energy gap E_g at Γ as a function of ϵ_{aa} and ϵ_{cc} , i.e., strain (%) along the a and c lattice directions, calculated by density functional theory (DFT). It shows a V-shaped valley, with the minimum of the valley (corresponding to $E_g = 0$) extending along the diagonal direction. Z_2 topological indices were also computed, and the $E_g = 0$ line is the phase boundary between STI and WTI. Fig. 21 (a) reveals a highly anisotropic strain dependence of E_g : The steepest gradient of E_g is along the direction in which ϵ_{aa} and ϵ_{cc} have opposite signs, as the case when an uniaxial stress is applied along the a lattice direction. In contrast, the strain induced by applying hydrostatic pressure corresponds to a trajectory that is almost parallel to the contour lines. The Poisson’s ratio of the anisotropic strain, $\epsilon_{cc}/\epsilon_{aa} = 0.25$, induced by a -axis uniaxial stress, was obtained using a fully relaxed vdW-DFT calculation, and it is indicated as the gray arrow in Fig. 21 (a). It requires less than a percent of ϵ_{aa} to reach the WTI-STI phase boundary. We note that there is uncertainty in the DFT bandgap size for a given set

of lattice constants. For example, several spectroscopy measurements reported E_g as low as 10 meV, which is significantly lower than the calculated 60-meV bandgap at zero strain [176, 177]. Hence, the actual strain required to pass through $E_g = 0$ could be smaller.

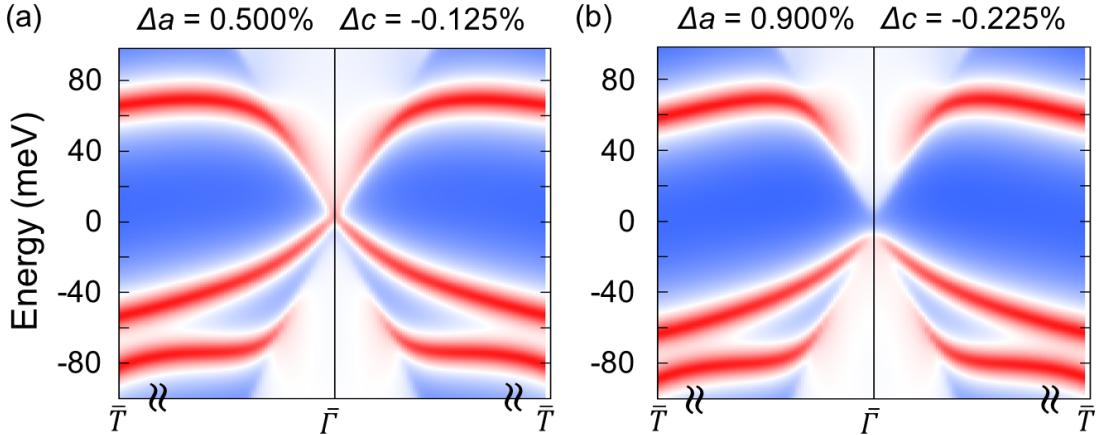


Figure 22: Surface states (represented by red lines) around Γ point on the (001) surface Brillouin zone in ZrTe_5 for (a) STI and (b) WTI phases. (The white regions indicate the bulk bands.)

Based on the band structures of Fig. 21(b), during the topological phase transition, one might observe the electrical transport properties, such as resistivity, will fall and rise again. This peculiar phenomenon was observed in the experiments done by Mutch *et al.* [193], and they also found that the positive longitudinal magnetoconductance also shows an unusual behavior that the value becomes maximal at the critical strain of phase transition. Besides, we further calculate the surface states around Γ point in the (001) surface Brillouin zone [shown in Fig. 18(c)]. Figure 22(a) shows the topologically-protected surface states (represented by red lines) with a Dirac point at Γ point for STI phase. For WTI phase, the surface states do not connect valence band and conduction band anymore. In summary, uni-axial strain is an effective tool to drive topological phase transition in small bandgap materials.

This section is based on our study published in *Science advances*, 5(8), p.eaav9771 [193].

Structural and Topological Phase Transitions in LaN

Semimetallic lanthanum monopnictides such as LaAs, LaSb, and LaBi have attracted much attention because of their unusual extreme magnetoresistance due to electron-hole compensation [194, 195, 196, 197]. These materials typically assume stable rock-salt structures, and they are promising for magnetic sensor and spintronics applications. Different from the above semimetals, LaP is found theoretically to be semiconducting with a narrow band gap (< 0.1 eV), and it displays high thermoelectric performance tunable by strain and pressure [198, 199]. In addition to the intriguing magnetic and electronic behaviors of lanthanum monopnictides, researchers have recently paid attention to their non-trivial topological properties.

In the single-particle picture, topological bulk materials can be roughly categorized into topological semimetals (TSMs) and topological insulators (TIs). TSMs are characterized by topologically protected band crossings [200]; they include nodal semimetals such as Weyl and Dirac semimetals [201, 202], as well as more exotic varieties like nodal-line semimetals [203], type-II Weyl semimetals [204], and multi-fold fermions [205, 206, 207]. For insulators with the time-reversal symmetry (TRS), topology arises from band inversion at TR-invariant momenta caused by strong spin-orbit coupling and can be classified by the Z_2 indices [208, 209]. The topologically non-trivial phases exhibit robust helical surface states with spin-momentum locking, which forbids back-scattering and thus renders new potential applications in low-dissipation devices.

LaX (X = N, P, As, Sb, and Bi) lacks conducting f -electrons and preserves TRS, and their band structures also show an energy gap between valence and conduction bands at each k -point in the Brillouin zone (BZ). Hence, LaX still permits the Z_2 classifications, even though some lanthanum monopnictides are semimetallic. Experimentally, the signature of topologically nontrivial (bulk) insulating phases can be determined by the conducting surface states near high symmetry points in the surface

BZ, and they can be observed directly by angular-resolved photoemission spectroscopy (ARPES). In ambient conditions, LaBi is confirmed to be topologically nontrivial by several experimental groups [210, 211, 212, 195, 213, 214]. In contrast, LaAs and LaSb are found to be topologically trivial in ARPES experiments [215, 195, 214]. Recently, topological phase transitions (TPTs) are especially intriguing, because of the possibility to control different topological phases [216, 217, 192, 193, 218, 219, 220]. In LaAs and LaSb, pressure-induced TPTs have been proposed [221, 222]. Since LaN and LaP are expected to show similar crystal and electronic structures as other lanthanum monopnictides, TPTs also may be achieved by uniaxial strain or hydrostatic pressure.

In this paper, we use first-principles calculations to study the structural, electronic, and topological phase transitions of LaN. By re-examining the dynamic stabilities of all rock-salt LaX , we find that only LaN is dynamically unstable with imaginary phonon modes at zero temperature. Using evolutionary crystal structure prediction techniques, we find a new, dynamically stable structure of LaN with $P1$ symmetry. In our calculations, the $P1$ -LaN also exhibits spontaneous electronic polarization, and it can undergo concomitant structural and ferroelectric transitions driven by temperature or pressure. The rest of the paper is organized as follows. First, we discuss the computational methods based on density functional theory, evolutionary crystal structure prediction, and *ab initio* molecular dynamics. Next, we present the theoretical results of structural, electronic, and topological properties of LaN under different temperature and pressure. We al discuss additionally topological phase transitions in rock-salt LaP, LaAs, and LaSb under hydrostatic pressure. Sec. IV concludes the paper by summarizing the conditions for stabilizing different phases under study. In the end,

Our density functional theory (DFT) calculations are conducted using the Vienna Ab initio Simulation Package (VASP) [112, 113, 223, 224]. In particular, the

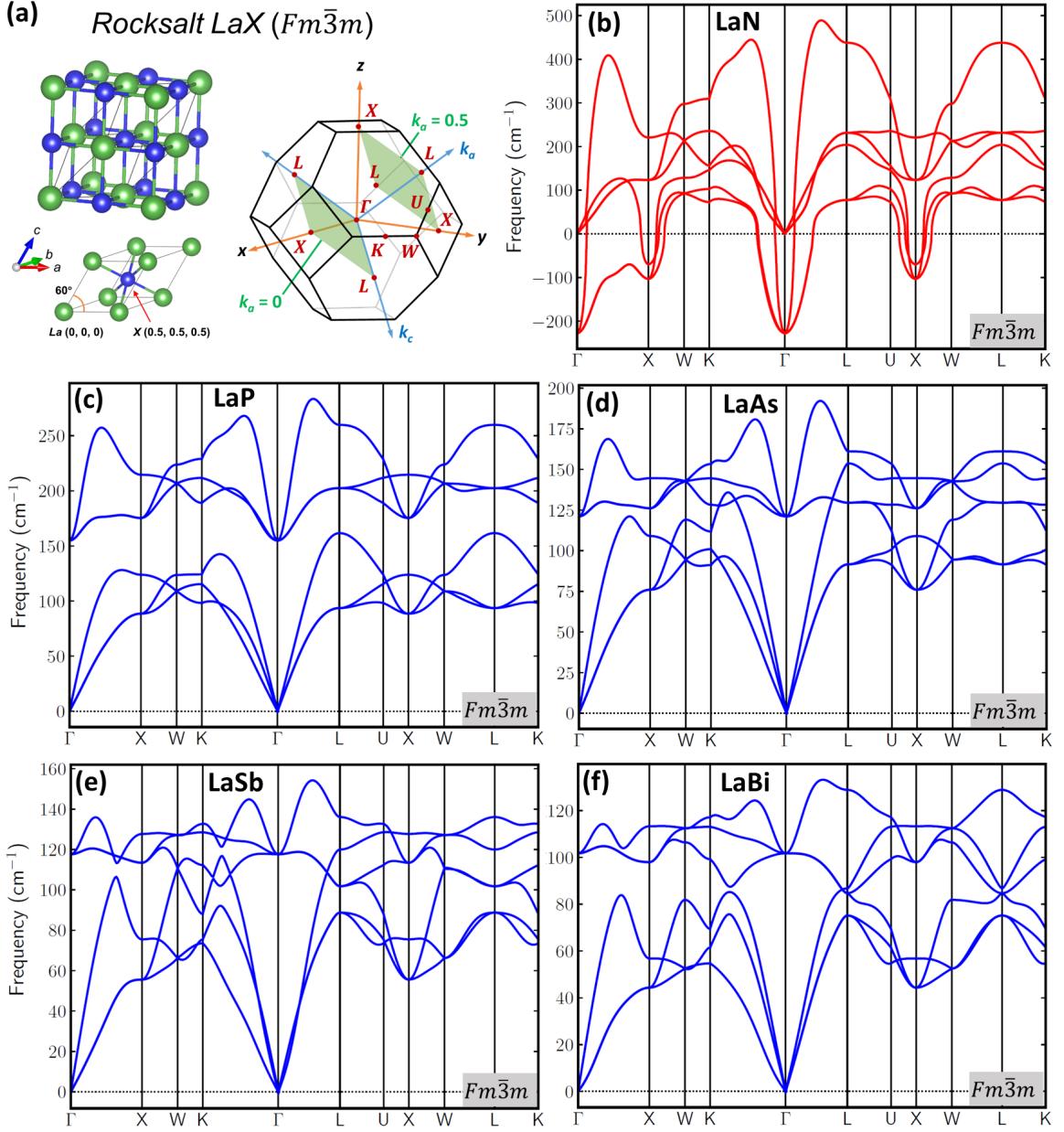


Figure 23: (a) Rock-salt structure of LaX (with $X = \text{N}, \text{P}, \text{As}, \text{Sb}$, or Bi), and the corresponding Brillouin zone with projected (001) and (111) surfaces. (b)-(f) Phonon dispersions (without non-analytical term correction) of LaN , LaP , LaAs , LaSb , and LaBi , respectively. The imaginary phonon modes in (b) show that the rock-salt LaN is dynamically unstable at 0 K .

Perdew-Burke-Ernzerhof generalized gradient approximation (PBE-GGA) [162] functional and the projector augmented wave (PAW) method [134, 135] are adopted. The wavefunctions are expanded in plane-wave basis sets with a kinetic energy cutoff of 500 eV, and a Γ -centered $11 \times 11 \times 11$ Monkhorst-Pack k mesh [109] is used for the Brillouin zone sampling. In structure relaxation calculations, the convergence criteria for the force on each atom and the total energy are set to 1×10^{-3} eV/ \AA and 1×10^{-8} eV, respectively.

To search for the zero-temperature (0 K) crystal structure of LaN, the USPEX package [93, 94, 95] is used for crystal structure prediction (CSP). USPEX implements a highly efficient evolutionary algorithm for finding stable and meta-stable low-enthalpy structures corresponding to local minima in the potential energy surface. In our CSP calculations, 2-, 4-, and 8-atom unit cells are considered. The first generation of structures are randomly created. Subsequent generations are produced with 20% from random structures and 80% from heredity, soft mutation, and transmutation operators.

The phonon dispersions at 0 K are calculated by the frozen-phonon approach with a $5 \times 5 \times 5$ supercell (250 atoms) using VASP and PHONOPY [140]. Here, VASP calculates the force on each atom in the supercell and provides the force information for PHONOPY, which in turn computes interatomic force constants and obtains phonon eigenvalues by diagonalizing the dynamical matrix. The atomic displacement is set as 0.02 \AA , and a $3 \times 3 \times 3$ k -mesh is employed in the self-consistency calculations. Since most of the rare-earth monopnictides are semimetallic, the non-analytical term correction is not included in our calculations.

For finite-temperature phonon spectra, we adopt the temperature-dependent effective potential (TDEP) [225] technique with *ab initio* molecular dynamics (AIMD). The AIMD simulations use a time step of 2 fs and a total simulation length of 20 ps. We have checked the convergence of our results with respect to the total simulation

time. The final 200 structures are then utilized to estimate the force constants via least-square fitting by using the ALAMODE package [226].

For electronic structure calculations, band orderings are essential for capturing the correct topological nature. However, standard DFT calculations based on local functionals like LDA or GGA typically underestimate the band gap, which may lead to an incorrect topological classification. To more adequately compute the band gap, we further apply an advanced non-local hybrid functional HSE06 [227] with spin-orbit coupling (SOC). Since hybrid DFT is computationally more demanding, we reduce the k points to $9 \times 9 \times 9$ and the cutoff energy to 400 eV (and we have checked that the band structure is quantitatively consistent with that obtained by using a $11 \times 11 \times 11$ k mesh and 500 eV energy cutoff). With the eigenvalues and eigenfunctions from hybrid-DFT, we construct maximally localized Wannier functions (MLWFs) and obtain corresponding *ab-initio* tight-binding models using the Wannier90 package [186]. We then use the WannierTools package [187] to identify topological Z_2 indices by tracking the Wannier charge centers [228].

Several rare-earth monopnictides are known to be stabilized in the cubic rock-salt structure (space group No. 225, $Fm\bar{3}m$ symmetry), as shown in Fig. 23(a). Indeed, a previous X-ray diffraction (XRD) experiment has reported the observation of rock-salt LaN [229]. However, a recent experiment using magnetron sputtering found only wurtzite and zinc blende structures for LaN [230]. As a result, we begin by using DFT to examine the structural stability. We first fully relax the rock-salt structures for all lanthanum monopnictides using both LDA and GGA functionals without SOC. The resulting lattice parameters obtained by GGA (LDA) for LaX (with X ranging from N, P, As, Sb, to Bi) are 5.3076 (5.2140), 6.0496 (5.9377), 6.1925 (6.0656), 6.5466 (6.4019), and 6.6596 (6.5065) Å, respectively. Compared with the experimental lattice parameters of 6.1670(7) Å for LaAs [197], 6.5008 Å for LaSb [196], and 6.5799 Å for LaBi [196], GGA slightly overestimates the lattice parameters (with respective errors

of 0.4%, 0.7%, and 1.2%), while LDA underestimates them (with respective errors of -1.6%, -1.5%, and -1.1%). We note that the SOC effect on lattice parameters is less than 0.1%. Since overall the GGA results agree better with experiments, below we consider the fully relaxed structures obtained by GGA as the unstrained structures.

To examine the structural stabilities, we first compute the elastic constants using VASP [231] to ensure that the Born-Huang criteria for cubic crystals are met [232]: $C_{11} - C_{12} > 0$, $C_{11} + 2C_{12} > 0$, and $C_{44} > 0$. Our calculations indicate that all LaX rock-salt structures are mechanically stable. We next study the dynamical stability by calculating the phonon dispersion at 0 K. The phonon spectrum of LaN in Fig. 23(b) contains soft modes with imaginary frequencies around the Γ and X points, which indicates the dynamical instability of LaN's rock-salt structure. Our LaN phonon dispersion calculated by a finite-displacement method is consistent with the study of Gökoglu *et al.* [233] based on density functional perturbation theory. For other lanthanum monopnictides, the phonon spectra in Figs. 23(c)-(f) do not exhibit any soft mode, showing that the corresponding rock-salt structures are dynamically stable. This is also consistent with recent experiments [215, 195, 214]. We further note that in our electronic structure calculation using the HSE06 hybrid functional [227], unstrained rock-salt LaP is a semiconductor with an indirect band gap ~ 82 meV. Therefore, a longitudinal optical-transverse optical (LO-TO) phonon splitting near Γ may be observed experimentally.

The imaginary phonon modes in rock-salt LaN imply a different stable crystal structure at 0 K. Using the evolutionary algorithm implemented in USPEX [93, 94, 95], we conduct crystal structure searches at 0 K and find a lower-symmetry structure with space group No. 1 ($P1$ symmetry) shown in Fig. 24(a) as the ground state. The lattice parameters of the primitive cell are $a = 3.7245 \text{ \AA}$, $b = 3.8192 \text{ \AA}$, and $c = 3.7557 \text{ \AA}$. The angles between them are $\alpha = 58.8990^\circ$, $\beta = 61.3750^\circ$, and $\gamma = 59.7033^\circ$. The Wyckoff positions for La and N are $(0, 0, 0)$ and $(0.42, 0.52, 0.58)$, respectively. This

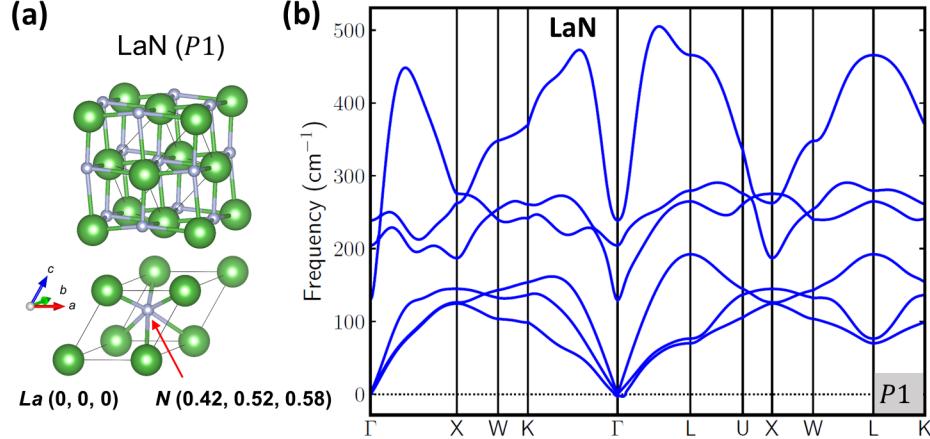


Figure 24: (a) Distorted rock-salt structure obtained by evolutionary algorithm for LaN. This structure lacks an inversion center and belongs to space group No. 1 ($P1$ symmetry). The lattice parameters are $a = 3.7245 \text{ \AA}$, $b = 3.8192 \text{ \AA}$, $c = 3.7557 \text{ \AA}$, $\alpha = 58.8990^\circ$, $\beta = 61.3750^\circ$, and $\gamma = 59.7033^\circ$. The Wyckoff positions are La (0, 0, 0) and Na (0.42, 0.52, 0.58). (b) Phonon dispersion of the $P1$ -LaN structure. The labelings of high-symmetry points are based on the fcc lattice. The absence of imaginary phonon mode indicates that $P1$ -LaN is dynamically stable at 0 K.

structure can be regarded as a distorted rock-salt structure, where the N position is slightly shifted away from the more symmetric (0.5, 0.5, 0.5) position. Figure 24(b) shows the corresponding phonon dispersion, and the absence of imaginary mode indicates that $P1$ -LaN is dynamically stable at 0 K.

To ensure that $P1$ -LaN is indeed the ground state structure, we also compute the total energy versus volume for LaN in five different crystal structures: the low-symmetry structure $P1$ (space group No. 1), rock salt $Fm\bar{3}m$ (No. 225), PbO-like tetragonal structure $P4/nmm$ (No. 129), wurtzite $P6_3mc$ (No. 186), and zinc blende $Fm\bar{4}3m$ (No. 216). As shown in Fig. 25(a), the $P1$ structure has the lowest total energy, which is 29 meV (per formula unit of LaN) lower than that of the rock-salt structure. The lowest possible total energy of the wurtzite structure is also 16 meV (per formula unit) lower than that of the rock-salt structure. In the phonon dispersions for the wurtzite and zinc-blende structures shown in Fig. 26, both of them are dynamically stable at 0 K. Our theoretical results thereby support the recent experimental findings of synthesized wurtzite and zinc-blende LaN [230].

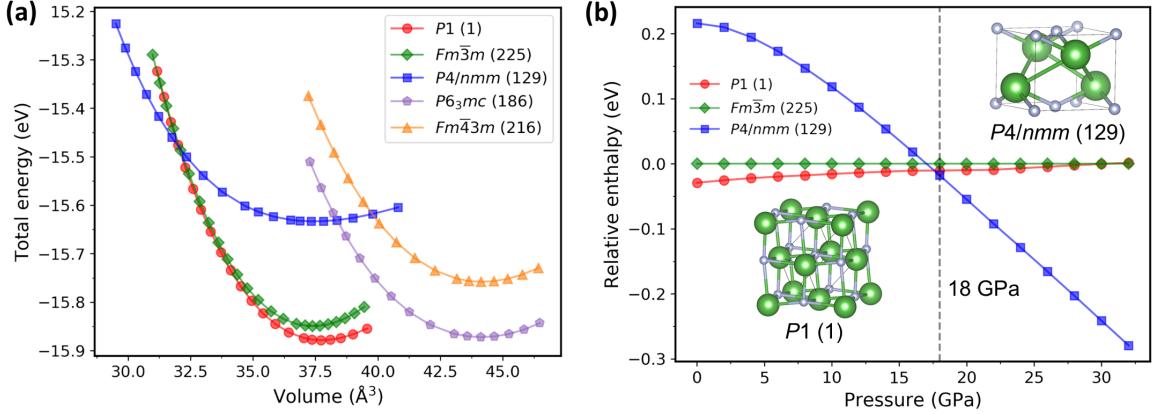


Figure 25: (a) Total-energy versus volume curves of LaN in different crystal symmetries, including the low-symmetry structure $P1$ (space group No. 1), rock salt $Fm\bar{3}m$ (No. 225), PbO-like tetragonal structure $P4/nmm$ (No. 129), wurtzite $P6_3mc$ (No. 186), and zinc blende $Fm\bar{4}3m$ (No. 216). (b) Relative formation enthalpy as a function of pressure using rock-salt LaN as the reference. The result indicates a pressure-induced structure transition from $P1$ to $P4/nmm$ symmetry near 18 GPa.

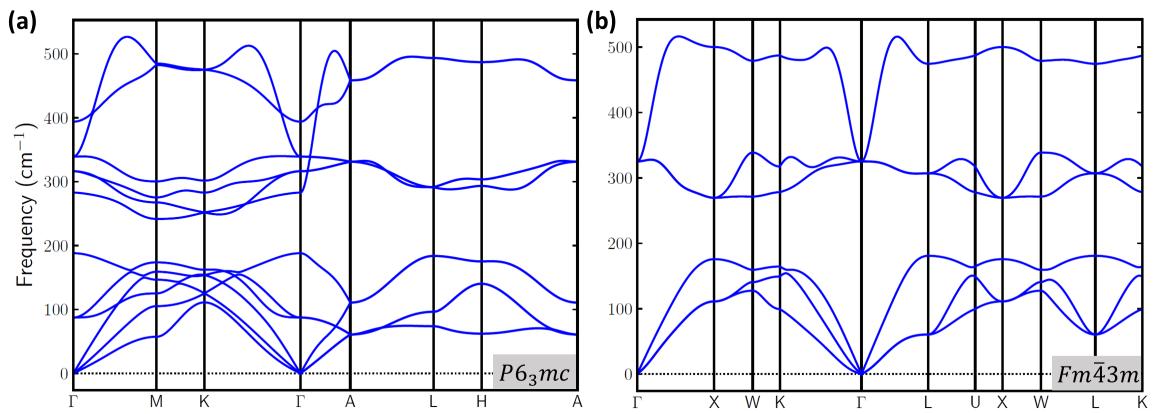


Figure 26: Phonon dispersions of LaN in (a) wurtzite and (b) zinc-blende structures, and both of them are dynamically stable.

Figure 25(a) also shows that the total energy-volume curves of $P1$ and $P4/nmm$ intersect around 32 \AA^3 . Therefore, applying external pressure can favor the PbO-like $P4/nmm$ tetragonal structure. Figure 25(b) shows the calculated relative formation enthalpy as a function of pressure, using the rock-salt $Fm\bar{3}m$ structure as the reference. The result indicates that when the external pressure is below 18 GPa, $P1$ -LaN has the lowest enthalpy. Above 18 GPa, $P4/nmm$ -LaN becomes the lowest enthalpy structure. In other words, our DFT calculations suggest a pressure-induced structural phase transition from $P1$ to $P4/nmm$ symmetry in LaN near 18 GPa at 0 K. Our predicted transition pressure is similar to that from a previous theoretical study by Mukherjee *et al.*, who proposed a transition pressure of 19 GPa for LaN to transit from $Fm\bar{3}m$ to $P4/nmm$ symmetry [234]. The theoretical results are also consistent with the experimental finding by Schneider *et al.*, who observed the $Fm\bar{3}m$ structure in ambient conditions, and it undergoes a phase transition to the PbO-like $P4/nmm$ tetragonal structure at ~ 22.8 GPa [235].

On the other hand, although our DFT calculation predicts a stable $P1$ -LaN at 0 K, this structure has not been reported before. As mentioned above, the rock-salt $Fm\bar{3}m$ -LaN is generally acknowledged in the literature as the correct structure in ambient conditions. These results imply that LaN may exhibit a temperature-induced structural transition from $P1$ to $Fm\bar{3}m$ symmetry. In fact, the situation here for LaN is very similar to that in ferroelectric materials GeSe and GeTe [236, 237]. There, the structural transitions induced by temperature are usually accompanied by ferroelectric transitions. For example, rhombohedral α -GeTe (space group No. 160, $R3m$ symmetry) undergoes a ferroelectric transition and a structural change to rock-salt β -GeTe (No. 225, $Fm\bar{3}m$) near 700 K [237]. An important signature of ferroelectric material is the double-well potential energy surface (PES), which can be obtained by the nudged elastic band method. Motivated by the similarity between GeTe and LaN, we compute the spontaneous electric polarization of $P1$ -LaN based

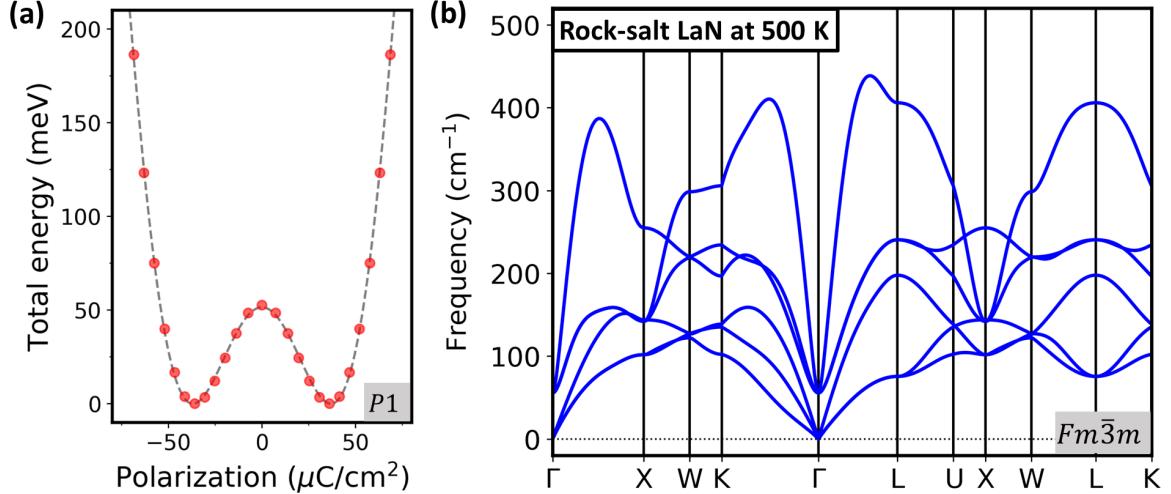


Figure 27: (a) Double-well potential energy versus polarization in $P1$ -LaN using the ground state energy as a reference, where the ground state shows a spontaneous polarization of $36 \mu\text{C}/\text{cm}^2$. The gray dashed line is the fitting function up to 6th order. (b) Renormalized phonon dispersion of rock-salt ($Fm\bar{3}m$) LaN at 500 K. Compared to the phonon dispersion at 0 K in Fig. 23(b), the absence of imaginary mode here shows that rock-salt LaN can be stabilized at high temperature.

on the modern theory of polarization [238]. Figure 27(a) shows the potential energy as a function of polarization, which exhibits a double-well energy landscape. In our calculations, $P1$ -LaN has a switching barrier of 53 meV per formula unit, and the ground state structure displays a net polarization of $36 \mu\text{C}/\text{cm}^2$. Therefore, LaN is likely to undergo a concurrent ferroelectric transition and structural change from $P1$ to $Fm\bar{3}m$ symmetry when temperature is increased.

To confirm our conjecture that rock-salt LaN can be stabilized at high temperature, we calculate its finite-temperature (or renormalized) phonon dispersion. To include anharmonic effects, we use the TDEP method [225] to obtain the effective 2nd-order force constants and fit them via AIMD simulations. Figure 27(b) shows the renormalized phonon dispersion of rock-salt LaN at 500 K. No imaginary mode exists, indicating the dynamical stability of rock-salt LaN at high temperature. Here, we do not attempt to discuss the exact critical temperature, because it is computationally demanding to calculate and also depends on the chosen DFT functional.

Here, we do not attempt to discuss the exact critical transition temperature, because it is computationally demanding to calculate and also depends on the chosen DFT functional. In particular, our phonon dispersions computed with the PBE functional still contain negative frequencies of -60 and -44 cm^{-1} at the Γ point at temperatures 100 K and 300 K , respectively. We have also calculated the phonon dispersion at 300 K using the LDA functional, and the result shows that the negative phonon frequency at 300 K is reduced to only -19 cm^{-1} .

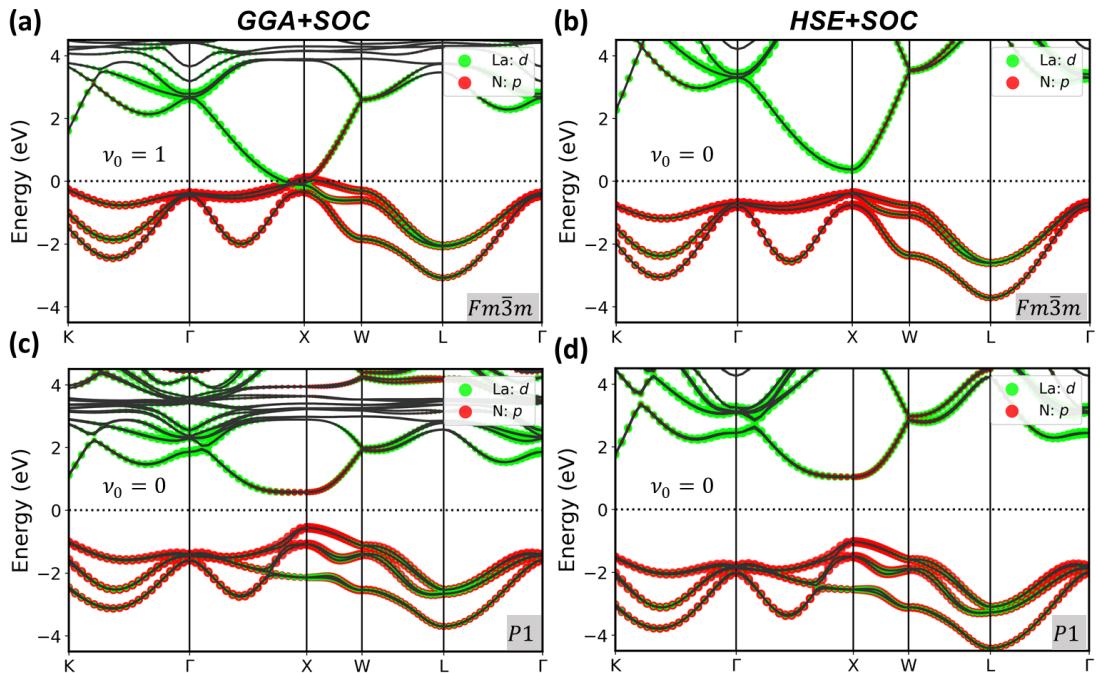


Figure 28: Electronic band structures with spin-orbit coupling (SOC) and Z_2 topological indices for rock-salt LaN, calculated with (a) the PBE-GGA functional and (b) the HSE06 hybrid functional. (c)-(d) are similar plots as (a)-(b), but for $P1$ LaN. The labelings of high-symmetry points are based on the fcc lattice. The Fermi level (horizontal dotted line) is set to 0 eV . GGA typically underestimates the band gap, which can be corrected by a more advanced hybrid functional.

We next shift our focus to electronic and topological properties. In a 3D system with time reversal symmetry (TRS), band topology can be characterized by the Z_2 indices ($\nu_0; \nu_1, \nu_2, \nu_3$), if an energy gap exists at each k -point in the BZ [71]. A strong topological insulator (STI) has a strong index $\nu_0 = 1$ with an odd number of Dirac cones on any surface. In contrast, the weak TI phase has $\nu_0 = 0$, and it can be

regarded as a stacking of 2D quantum spin Hall [239] layers along the (ν_1, ν_2, ν_3) “mod 2” reciprocal lattice vector [71]. Therefore, the weak indices (ν_1, ν_2, ν_3) depend on the choice of unit cell, and here they are defined by the primitive crystal structure. To understand the Z_2 topology of a system with both TRS and inversion symmetries [68], one can study simply the product of the parity eigenvalues of the occupied bands at the time-reversal invariant momenta (TRIM). In the rock-salt structure, the TRIM include 1 Γ , 3 X , and 4 L , as shown in Fig. 24(b). Therefore, only a sign change in the parity eigenvalue at Γ or X point can lead to a change in ν_0 (i.e. a STI topological phase transition). On the other hand, the three weak indices depend on 2 (even number) X and 2 (even number) L points, and thereby they are always 0.

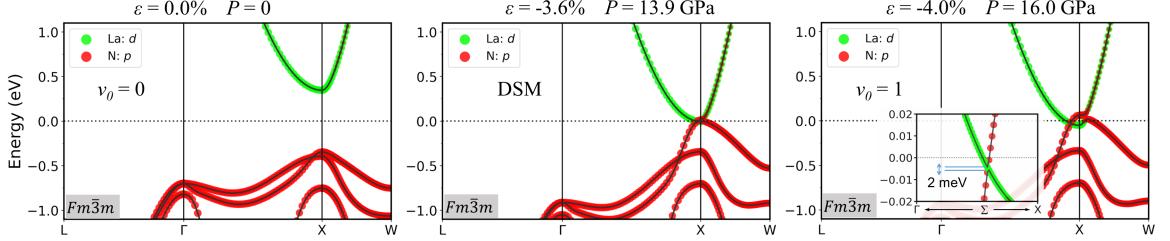


Figure 29: Electronic band structures and Z_2 topological indices of rock-salt LaN under different strain ε and corresponding pressure P . Left: $P = 0$, trivial insulator ($\nu_0 = 0$). Middle: $P = 13.9$ GPa, Dirac semimetal (DSM). Right: $P = 16.0$ GPa, strong topological insulator ($\nu_0 = 1$). The results show a pressure-driven topological phase transition in rock-salt LaN.

The electronic band topology of LaX was first investigated by Zeng *et al.* using the DFT+U method [240], which however does not fix the issue of underestimating LaX’s band gaps and may lead to incorrect topological classifications. Later DFT studies based on the modified Becke-Johnson (mBJ) meta-GGA functional [165, 166] or screened hybrid functionals [227] have addressed this issue, and thereby their topological classifications are potentially more appropriate [241, 197, 242]. Here, we study the electronic band structures using both PBE-GGA [162] and HSE06 [227]. For rock-salt LaN, the GGA band structure in Fig. 27(a) shows a semimetal and a topologically non-trivial STI phase ($\nu_0 = 1$) with band inversion at the X point. However,

this STI assignment is potentially incorrect, due to an underestimation of the overall gap between valence and conduction bands. In contrast, the more advanced HSE06 band structure in Fig. 28(b) shows that rock-salt LaN is a semiconductor with a direct band gap ~ 0.8 eV at the X point, and the material is topologically trivial ($\nu_0 = 0$). We also calculate the band structures of $P1$ -LaN. As shown in Figs. 27(c)-(d), the PBE-GGA and HSE06 results both show a semiconducting phase with a direct band gap ~ 1.1 and 2.2 eV, respectively. The orbitally-resolved band structures of $P1$ -LaN show contributions from nitrogen's p -orbitals near the conduction band minimum. However, we note that this apparent “band inversion” at the X point is not related to SOC. Instead, the orbital hybridization is caused by the lower-symmetry of the distorted $P1$ structure. Due to the absence of inversion symmetry, the topological nature cannot be simply determined from the parity eigenvalues of the TRIM. Instead, here ν_0 is calculated by tracking directly the Wannier charge centers, and $P1$ -LaN is found to be topologically trivial ($\nu_0 = 0$).

Finally, we study pressure-induced topological phase transitions (TPTs) in LaX . Recently, researchers have reported that the band topology of LaAs and LaSb can be controlled by pressure [221, 222]. External pressure can modify the band gap, leading to a band inversion between valence and conduction bands. Following a similar idea, here we calculate the HSE06 band structures of LaN under external pressure. As shown in the left panel of Fig. 29, the unstrained (ambient-condition) rock-salt LaN is topologically trivial ($\nu_0 = 0$). When a hydrostatic pressure of 13.9 GPa is applied, the energy gap is closed at the X point, and rock-salt LaN becomes a Dirac semimetal (middle panel of Fig. 29). When the external pressure is above 13.9 GPa, a band inversion occurs (right panel of Fig. 29), and the material becomes a STI ($\nu_0 = 1$). Since inversion is preserved across the transition, the Dirac semimetal appears exactly and only at the TPT, and must necessarily appear at a TRIM [243, 244], here the three X points in the BZ. We note that in the non-trivial STI phase, there is a very

small gap ~ 2 meV (the Σ point in the insert, right panel of Fig. 29) due to SOC. We also explore possible TPTs in $P1$ -LaN. However, for external pressure less than 18 GPa, the band gap remains finite, which means that $P1$ -LaN is always a trivial insulator before transforming into the PbO-like $P4/nmm$ tetragonal structure.

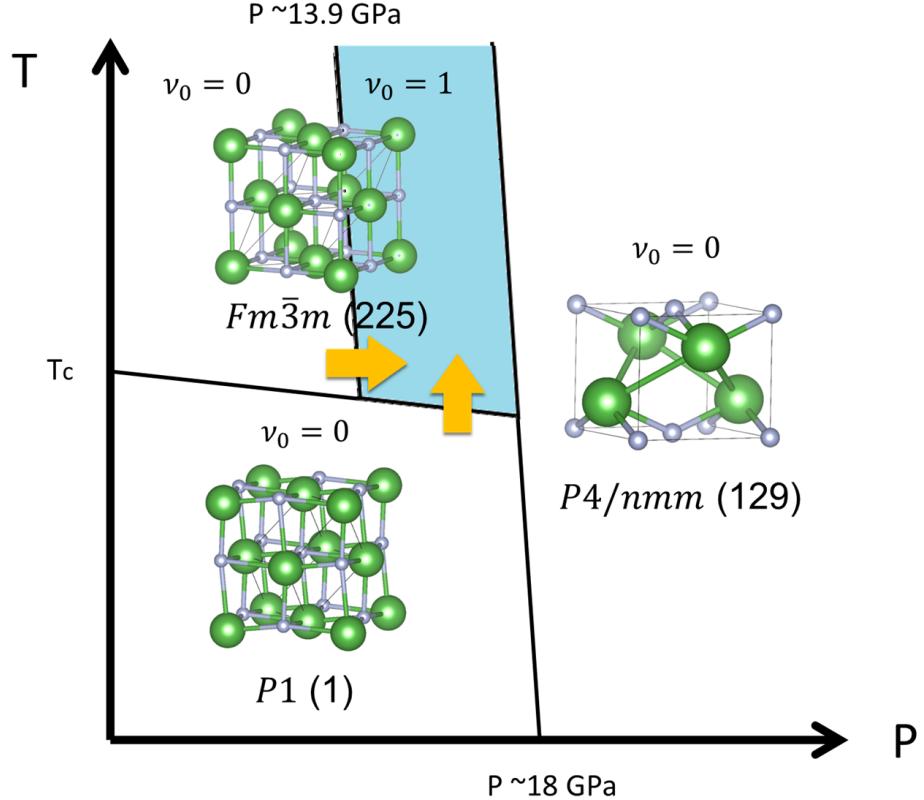


Figure 30: Schematic phase diagram of LaN. The two arrows indicate that the topological phase transition can be driven by either pressure or temperature.

Our study for LaN can be summarized by the schematic phase diagram shown in Fig. 30. When the pressure and temperature are low, we proposed that LaN has a distorted rock-salt structure ($P1$ symmetry) with spontaneous polarization. Under low temperature, LaN has a pressure-induced structural phase transition from $P1$ structure to $P4/nmm$ structure. Under low pressure, DFT predicts concurrent temperature-induced structural phase transition and ferroelectric transition. Besides, based on our study for topological phase transition, rock-salt LaN becomes topologically non-trivial when the pressure is between 13.9 and 18 GPa. Therefore, based on

the phase diagram, the topological phase transition can be achieved by both pressure and temperature as indicated by the arrows in Fig. 30.

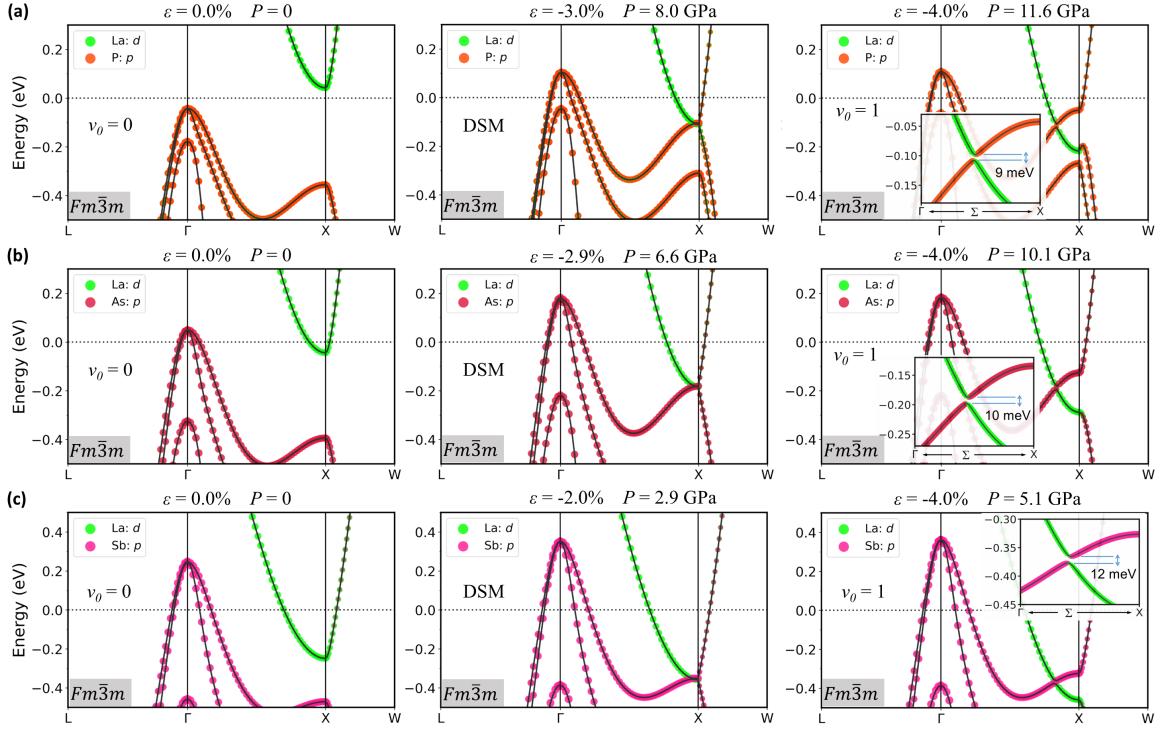


Figure 31: Electronic band structures and Z_2 topological indices of rock-salt (a) LaP, (b) LaAs, and (c) LaSb, under different strain ε and corresponding pressure P . Left: Trivial insulator ($\nu_0 = 0$). Middle: Dirac semimetal (DSM). Right: Strong topological insulator ($\nu_0 = 1$). The results show pressure-driven topological transitions in rock-salt lanthanum monipnictides..

Beside the topological phase transition of LaN, we also study pressure-induced topological phase transitions in rock-salt LaP, LaAs, and LaSb, by using a hybrid density functional. The electronic band structures of unstrained LaP, LaAs, and LaSb are shown in the left panels of Fig. 31, and they are all found to be topologically trivial ($\nu_0 = 0$) at zero pressure $P = 0$. The middle panels of Fig. 31 show the band structures of LaP, LaAs, and LaSb at the topological transition (Dirac semimetal), where the critical hydrostatic pressures are $P = 8.0$ GPa (with strain value $\varepsilon = -3.0\%$), 6.6 GPa ($\varepsilon = -2.9\%$), and 2.9 GPa ($\varepsilon = -2.0\%$), respectively. Finally, the right panels of Fig. 31 display the high-pressure band structures, showing band gaps at the

Σ point with gap sizes of 9, 10, and 12 meV, respectively for LaP, LaAs, and LaSb. This increasing gap size is expected with an increasing spin-orbital coupling strength as one goes down along the periodic table.

To conclude, using density functional theory and evolutionary crystal searches, we have discovered a new, low-temperature structure of LaN belonging to space group No. 1 with the *P1* symmetry. This lower-symmetry phase, which can be regarded as a distorted rock-salt structure, displays ferroelectricity with a spontaneous polarization of $36 \mu\text{C}/\text{cm}^2$. Using *ab initio* molecular dynamics, we found that when the temperature is increased, the *P1*-LaN will undergo simultaneous ferroelectric and structural transitions to the more symmetric rock-salt structure (space group 225, *Fm* $\bar{3}m$). In our enthalpy calculations, *P1*-LaN also could undergo a pressure-induced phase transition to the PbO-like *P4/nmm* tetragonal structure, when the external pressure is above 18 GPa. In our study of topological properties, the high-temperature rock-salt phase of LaN was found to undergo a pressure-induced transition from a trivial insulator to a strong topological insulator near 13.9 GPa. By increasing the temperature, we also have predicted another topological and structural transition from *P1*-LaN to rock-salt LaN, when the pressure is between 13.9 and ~ 18 GPa. These results demonstrate that LaN exhibits a rich phase diagram containing different ferroelectric, structural, and topological phases, and lanthanum monopnictides are intriguing quantum materials that can serve as a platform for exploring novel phase transitions induced by temperature and pressure.

This section is based on our manuscript on arXiv: arXiv:2104.12879 (2021) [245].

CHAPTER 5

CONCLUSION

In the first part of our studies, we focused on computational prediction for new superhard materials using machine learning, evolutionary algorithm, and density functional theory. We studied both B-C-N and B-N-O ternary compounds. For B-C-N materials, we first built machine learning models using existing online database, Materials Project. The models were then used to map out mechanical properties including bulk modulus, shear modulus and hardness. Our machine learning prediction showed that several B-C-N compounds with nearly 1:1 B-N ratio possess superhard property. Following the machine learning prediction, we conducted a series of evolutionary algorithm searches for B-C-N superhard compounds. Our results indicated that BC_{10}N has an extremely high hardness ~ 87 GPa. As for B-N-O materials, we generated our own machine learning data by evolutionary searches, and then we adopted a self-improving process to gradually locate the composition of superhard B-N-O compounds. We found that compositions with the formula $\text{B}_{x+2}\text{N}_x\text{O}_3$ are superhard and also thermodynamically more stable based on their formation energy and cohesive energy. Our studies of ternary superhard materials show that machine learning can efficiently accelerate the discovery of new functional materials.

Another part of this dissertation is topological materials, we found that ZrTe_5 has an extremely small bandgap ~ 60 meV, and it is in the vicinity of topological phase transition. By applying uniaxial strain along a -axis, ZrTe_5 undergoes a Z_2 topological phase transition from a strong topological insulator phase to a weak TI phase. In the case of LaN, we first discovered that the phonon dispersion of LaN is

dynamically unstable at 0 K, and we found a dynamically stable, low-temperature phase with $P1$ symmetry using evolutionary search. We next study the structural and topological phase transitions in LaN. We found LaN has a structural phase transition to rock-salt phase ($Fm\bar{3}m$ symmetry) when temperature is increased. By studying its phase diagram, our results indicate that high-temperature rock-salt LaN has a normal insulator to a strong topological insulator phase transition at $P \sim 14$ GPa, and its Z_2 transition can also be driven by temperature at P around 14 - 18 GPa. These results indicate that the control of topological phase transition can be effectively achieved by means of strain engineering and the hydrostatic pressure.

LIST OF PUBLICATIONS

Wei-Chih Chen, Joanna N. Schmidt, Da Yan, Yogesh K. Vohra, and Cheng-Chien Chen, “Machine Learning and Evolutionary Prediction of Superhard B-C-N Compounds”, *npj Comput Mater* **7**, 114 (2021).

Wei-Chih Chen, Chia-Min Lin, Joseph Maciejko, and Cheng-Chien Chen, “Structural and Topological Phase Transitions in LaN Driven by Pressure and temperature”, arXiv:2104.12879 (2021).

Chia-Min Lin, Wei-Chih Chen, and Cheng-Chien Chen, “First-Principles Study of Strain Effect on Thermoelectric Properties of LaP and LaAs”, arXiv:2008.06455 (2020).

Kaleb Burrage, Chia-Min Lin, Wei-Chih Chen, Cheng-Chien Chen, and Yogesh K. Vohra, “Electronic structure and anisotropic compression of Os₂B₃ to 358 GPa”, *J. Condens. Matter Phys.* **32**, 405703 (2020).

Kallol Chakrabarty, Wei-Chih Chen, Paul A. Baker, Vineeth M. Vijayan, Cheng-Chien Chen, and Shane A. Catledge, “Superhard Boron-Rich Boron Carbide with Controlled Degree of Crystallinity”, *Materials* **13**, 3622 (2020).

Paul A. Baker, Wei-Chih Chen, Cheng-Chien Chen, Shane A. Catledge, and Yogesh K. Vohra, “First-Principles Predictions and Synthesis of B₅₀C₂ by Chemical Vapor Deposition”, *Sci. Rep.* **10**, 4454 (2020).

Kaleb Burrage, Chia-Min Lin, Wei-Chih Chen, Cheng-Chien Chen, and Yogesh K. Vohra, “Experimental and Computational Studies on Superhard Material Rhenium Diboride to Ultrahigh Pressures”, *Materials* **13**, 1657 (2020).

Bijuan Chen, Ekaterina M. Pärschke, Wei-Chih Chen, Brandon Scoggins, Bing Li, Mahalingam Balasubramanian, Steve Heald, Jianbo Zhang, Hongshan Deng, Raimundas Sereika, Yesudhas Sorb, Xia Yin, Yan Bi, Ke Jin, Qiang Wu, Cheng-Chien Chen, Yang Ding, and Ho-kwang Mao “Probing Cerium 4f States across the Volume Collapse Transition by X-ray Raman Scattering”, *J. Phys. Chem. Lett.* **10**, 7890 (2019).

Joshua Mutch, Wei-Chih Chen, Preston Went, Tiema Qian, Ilham Zaky Wilson, Anton Andreev, Cheng-Chien Chen, and Jiun-Haw Chu, “Evidence for a strain tuned topological phase transition in $ZrTe_5$ ”, *Sci. Adv.* **5**, eaav9771 (2019).

Paul A. Baker, Shane A. Catledge, Sumner B. Harris, Kathryn J. Ham, Wei-Chih Chen, Cheng-Chien Chen, and Yogesh K. Vohra, “Computational Predictions and Microwave Plasma Synthesis of Superhard Boron-Carbon Materials”, *Materials* **11**, 1279 (2018).

REFERENCES

- [1] Stan Vepřek. The search for novel, superhard materials. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 17(5):2401–2420, 1999.
- [2] Natalia Dubrovinskaia, Leonid Dubrovinsky, Wilson Crichton, Falko Langenhorst, and Asta Richter. Aggregated diamond nanorods, the densest and least compressible form of carbon. *Applied Physics Letters*, 87(8):083106, 2005.
- [3] Richard B Kaner, John J Gilman, and Sarah H Tolbert. Designing superhard materials. *Science*, 308(5726):1268–1269, 2005.
- [4] Zhisheng Zhao, Bo Xu, and Yongjun Tian. Recent advances in superhard materials. *Annual Review of Materials Research*, 46:383–406, 2016.
- [5] Bo Xu and Yongjun Tian. Ultrahardness: measurement and enhancement. *The Journal of Physical Chemistry C*, 119(10):5633–5638, 2015.
- [6] Sergio Neves Monteiro, Ana Lúcia Diegues Skury, Márcia Giardinieri de Azevedo, and Guerold Sergueevitch Bobrovnitchii. Cubic boron nitride competing with diamond as a superhard engineering material—an overview. *Journal of Materials Research and Technology*, 2(1):68–74, 2013.
- [7] J Haines, JM Leger, and G Bocquillon. Synthesis and design of superhard materials. *Annual Review of Materials Research*, 31(1):1–23, 2001.
- [8] J Yu, Z Zheng, HC Ong, KY Wong, S Matsumoto, and WM Lau. Thermal

stability of cubic boron nitride films deposited by chemical vapor deposition. *The Journal of Physical Chemistry B*, 110(42):21073–21076, 2006.

- [9] Yinjuan Liu, Guodong David Zhan, Qiang Wang, Duanwei He, Jiawei Zhang, Akun Liang, Timothy E Moellendick, Le Zhao, and Xiao Li. Hardness of polycrystalline wurtzite boron nitride (wbn) compacts. *Scientific reports*, 9(1):1–6, 2019.
- [10] Jialin Sun, Jun Zhao, Zhifu Huang, Ke Yan, Xuehui Shen, Jiandong Xing, Yimin Gao, Yongxin Jian, Hejie Yang, and Bo Li. A review on binderless tungsten carbide: development and application. *Nano-Micro Letters*, 12(1):1–37, 2020.
- [11] Lailei Wu, Tianskai Yao, Yachun Wang, Jingwu Zhang, Furen Xiao, and Bo Liao. Understanding the mechanical properties of vanadium carbides: Nano-indentation measurement and first-principles calculations. *Journal of alloys and compounds*, 548:60–64, 2013.
- [12] ZD Kovziridze, Z Mestvirishvili, G Tabatadze, NS Nizharadze, M Mshvildadze, and E Nikoleishvili. Improvement of boron carbide mechanical properties in b₄c-tib₂ and b₄c-zrb₂ systems. 2013.
- [13] K Sairam, JK Sonber, TSR Ch Murthy, C Subramanian, RC Hubli, and AK Suri. Development of b4c–hfb2 composites by reaction hot pressing. *International Journal of Refractory Metals and Hard Materials*, 35:32–40, 2012.
- [14] Xiaopeng Jiao, Hua Jin, Fuyang Liu, Zhanhui Ding, Bin Yang, Fengguo Lu, Xudong Zhao, and Xiaoyang Liu. Synthesis of boron suboxide (b6o) with ball milled boron oxide (b2o3) under lower pressure and temperature. *Journal of Solid State Chemistry*, 183(7):1697–1703, 2010.

- [15] Chien-Min Sung and Michael Sung. Carbon nitride and other speculative superhard materials. *Materials chemistry and physics*, 43(1):1–18, 1996.
- [16] Wilson Acchar, Carlos Alberto Cairo, and Ana Maria Segadães. Tem study of a hot-pressed al₂o₃-nbc composite material. *Materials Research*, 8(1):109–112, 2005.
- [17] VL Solozhenko, OO Kurakevych, and AR Oganov. On the hardness of a new boron phase, orthorhombic γ -b 28. *Journal of superhard materials*, 30(6):428–429, 2008.
- [18] Paul A Baker, Wei-Chih Chen, Cheng-Chien Chen, Shane A Catledge, and Yogesh K Vohra. First-principles predictions and synthesis of b 50 c 2 by chemical vapor deposition. *Scientific reports*, 10(1):1–7, 2020.
- [19] Kallol Chakrabarty, Wei-Chih Chen, Paul A Baker, Vineeth M Vijayan, Cheng-Chien Chen, and Shane A Catledge. Superhard boron-rich boron carbide with controlled degree of crystallinity. *Materials*, 13(16):3622, 2020.
- [20] Michael T Yeung, Jialin Lei, Reza Mohammadi, Christopher L Turner, Yue Wang, Sarah H Tolbert, and Richard B Kaner. Superhard monoborides: hardness enhancement through alloying in w1- xtaxb. *Advanced Materials*, 28(32):6993–6998, 2016.
- [21] Lei Han, Shanmin Wang, Jinlong Zhu, Songbai Han, Wenmin Li, Bijuan Chen, Xiancheng Wang, Xiaohui Yu, Baochang Liu, Ruifeng Zhang, et al. Hardness, elastic, and electronic properties of chromium monoboride. *Applied Physics Letters*, 106(22):221902, 2015.
- [22] Ronald G Munro. Material properties of titanium diboride. *Journal of Research of the National institute of standards and Technology*, 105(5):709, 2000.

- [23] E Zapata-Solvias, DD Jayaseelan, Hua-Tay Lin, P Brown, and WE Lee. Mechanical properties of zrb2-and hfb2-based ultra-high temperature ceramics fabricated by spark plasma sintering. *Journal of the European Ceramic Society*, 33(7):1373–1386, 2013.
- [24] Lisa E Pangilinan, Christopher L Turner, Georgiy Akopov, Mackenzie Anderson, Reza Mohammadi, and Richard B Kaner. Superhard tungsten diboride-based solid solutions. *Inorganic chemistry*, 57(24):15305–15313, 2018.
- [25] Shanmin Wang, Xiaohui Yu, Jianzhong Zhang, Yi Zhang, Liping Wang, Kurt Leinenweber, Hongwu Xu, Dmitry Popov, Changyong Park, Wenge Yang, et al. Crystal structures, elastic properties, and hardness of high-pressure synthesized crb 2 and crb 4. *Journal of Superhard Materials*, 36(4):279–287, 2014.
- [26] Hsiu-Ying Chung, JM Yang, SH Tolbert, and RB Kaner. Anisotropic mechanical properties of ultra-incompressible, hard osmium diboride. *Journal of Materials Research*, 23(6):1797–1801, 2008.
- [27] Hsiu-Ying Chung, Michelle B Weinberger, Jonathan B Levine, Abby Kavner, Jenn-Ming Yang, Sarah H Tolbert, and Richard B Kaner. Synthesis of ultra-incompressible superhard rhenium diboride at ambient pressure. *Science*, 316(5823):436–439, 2007.
- [28] Andrew T Lech, Christopher L Turner, Jialin Lei, Reza Mohammadi, Sarah H Tolbert, and Richard B Kaner. Superhard rhenium/tungsten diboride solid solutions. *Journal of the American Chemical Society*, 138(43):14398–14408, 2016.
- [29] Reza Mohammadi, Andrew T Lech, Miao Xie, Beth E Weaver, Michael T Yehung, Sarah H Tolbert, and Richard B Kaner. Tungsten tetraboride, an inex-

- pensive superhard material. *Proceedings of the National Academy of Sciences*, 108(27):10958–10962, 2011.
- [30] Georgiy Akopov, Wai H Mak, Dimitrios Koumoulis, Hang Yin, Bryan Owens-Baird, Michael T Yeung, Mit H Muni, Shannon Lee, Inwhan Roh, Zachary C Sobell, et al. Synthesis and characterization of single-phase metal dodecaboride solid solutions: $Zr_{1-x}M_xB_{12}$ and $Zr_{1-x}U_xB_{12}$. *Journal of the American Chemical Society*, 141(22):9047–9062, 2019.
- [31] Teng Ma, Hui Li, Xu Zheng, Shanmin Wang, Xiancheng Wang, Huaizhou Zhao, Songbai Han, Jian Liu, Ruifeng Zhang, Pinwen Zhu, et al. Ultrastrong boron frameworks in ZrB_{12} : A highway for electron conducting. *Advanced Materials*, 29(3):1604003, 2017.
- [32] Reza Mohammadi, Christopher L Turner, Miao Xie, Michael T Yeung, Andrew T Lech, Sarah H Tolbert, and Richard B Kaner. Enhancing the hardness of superhard transition-metal borides: molybdenum-doped tungsten tetraboride. *Chemistry of Materials*, 28(2):632–637, 2016.
- [33] Georgiy Akopov, Michael T Yeung, Christopher L Turner, Reza Mohammadi, and Richard B Kaner. Extrinsic hardening of superhard tungsten tetraboride alloys with group 4 transition metals. *Journal of the American Chemical Society*, 138(17):5714–5721, 2016.
- [34] Georgiy Akopov, Michael T Yeung, Inwhan Roh, Zachary C Sobell, Hang Yin, Wai H Mak, Saeed I Khan, and Richard B Kaner. Effects of dodecaboride-forming metals on the properties of superhard tungsten tetraboride. *Chemistry of Materials*, 30(10):3559–3570, 2018.
- [35] R Mnatsakanyan, D Davtyan, T Minasyan, S Aydinyan, and I Hussainova.

Superhard b4c-reb2 composite by sps of microwave synthesized nanopowders. *Materials Letters*, 285:129163, 2021.

- [36] Joshua Gild, Yuanyao Zhang, Tyler Harrington, Sicong Jiang, Tao Hu, Matthew C Quinn, William M Mellor, Naixie Zhou, Kenneth Vecchio, and Jian Luo. High-entropy metal diborides: a new class of high-entropy materials and a new type of ultrahigh temperature ceramics. *Scientific reports*, 6(1):1–10, 2016.
- [37] Yan Zhang, Wei-Ming Guo, Ze-Bin Jiang, Qi-Qi Zhu, Shi-Kuan Sun, Yang You, Kevin Plucknett, and Hua-Tay Lin. Dense high-entropy boride ceramics with ultra-high hardness. *Scripta Materialia*, 164:135–139, 2019.
- [38] Pranab Sarker, Tyler Harrington, Cormac Toher, Corey Osse, Mojtaba Samiee, Jon-Paul Maria, Donald W Brenner, Kenneth S Vecchio, and Stefano Curtarolo. High-entropy high-hardness metal carbides discovered by entropy descriptors. *Nature communications*, 9(1):1–10, 2018.
- [39] Vladimir L. Solozhenko, Oleksandr O. Kurakevych, Denis Andrault, Yann Le Godec, and Mohamed Mezouar. Ultimate metastable solubility of boron in diamond: Synthesis of superhard diamondlike bc₅. *Phys. Rev. Lett.*, 102:015506, 2009.
- [40] Paul A Baker, Shane A Catledge, Sumner B Harris, Kathryn J Ham, Wei-Chih Chen, Cheng-Chien Chen, and Yogesh K Vohra. Computational predictions and microwave plasma synthesis of superhard boron-carbon materials. *Materials*, 11(8):1279, 2018.
- [41] Vladimir L Solozhenko, Sergey N Dub, and Nikolay V Novikov. Mechanical properties of cubic bc_{2n}, a new superhard phase. *Diamond and Related Mate-*

- rials*, 10(12):2228–2231, 2001.
- [42] Vladimir L Solozhenko, Denis Andrault, Guillaume Fiquet, Mohamed Mezouar, and David C Rubie. Synthesis of superhard cubic bc₂n. *Applied Physics Letters*, 78(10):1385–1387, 2001.
- [43] Y Zhao, DW He, LL Daemen, TD Shen, RB Schwarz, Y Zhu, DL Bish, J Huang, J Zhang, G Shen, et al. Superhard b–c–n materials synthesized in nanostructured bulks. *Journal of materials research*, 17(12):3139–3145, 2002.
- [44] Xiaobing Liu, Xiaopeng Jia, Zhuangfei Zhang, Ming Zhao, Wei Guo, Guofeng Huang, and Hong-an Ma. Synthesis and characterization of new “bcn” diamond under high pressure and high temperature conditions. *Crystal Growth & Design*, 11(4):1006–1014, 2011.
- [45] Hervé Hubert, Laurence AJ Garvie, Bertrand Devouard, and Paul F McMillan. High-pressure, high-temperature syntheses of super-hard α -rhombohedral boron-rich solids in the bcno. *MRS Online Proceedings Library (OPL)*, 499, 1997.
- [46] NB Bolotina, TI Dyuzheva, and NA Bendeliani. Atomic structure of boron suboxycarbide b (c, o) 0.155. *Crystallography Reports*, 46(5):734–740, 2001.
- [47] Yinwei Li, Quan Li, and Yanming Ma. B₂co: A potential superhard material in the bco system. *EPL (Europhysics Letters)*, 95(6):66006, 2011.
- [48] Chao Liu, Mingwei Chen, Julong He, Shuangshuang Yu, and Tongxiang Liang. Superhard b₂co phases derived from carbon allotropes. *RSC advances*, 7(82):52192–52199, 2017.
- [49] Meiguang Zhang, Haiyan Yan, Baobing Zheng, and Qun Wei. Influences of

- carbon concentration on crystal structures and ideal strengths of b₂c_xo compounds in the bco system. *Scientific reports*, 5(1):1–14, 2015.
- [50] Shengnan Wang, Artem R Oganov, Guangrui Qian, Qiang Zhu, Huafeng Dong, Xiao Dong, and M Mahdi Davari Esfahani. Novel superhard b–c–o phases predicted from first principles. *Physical Chemistry Chemical Physics*, 18(3):1859–1863, 2016.
- [51] Brad A Steele and Ivan I Oleynik. Ternary inorganic compounds containing carbon, nitrogen, and oxygen at high pressures. *Inorganic chemistry*, 56(21):13321–13328, 2017.
- [52] Qian Li, Jianyun Wang, Miao Zhang, Quan Li, and Yanming Ma. Superhard-driven search of the covalent network in the b₃no system. *RSC Advances*, 5(45):35882–35887, 2015.
- [53] Shrikant Bhat, Leonore Wiehl, Leopoldo Molina-Luna, Enrico Mugnaioli, Stefan Lauterbach, Sabrina Sicolo, Peter Kroll, Michael Duerrschnabel, Norimasa Nishiyama, Ute Kolb, et al. High-pressure synthesis of novel boron oxynitride b₆n₄o₃ with sphalerite type structure. *Chemistry of Materials*, 27(17):5907–5914, 2015.
- [54] K v Klitzing, Gerhard Dorda, and Michael Pepper. New method for high-accuracy determination of the fine-structure constant based on quantized hall resistance. *Physical review letters*, 45(6):494, 1980.
- [55] Klaus Von Klitzing. Developments in the quantum hall effect. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 363(1834):2203–2219, 2005.
- [56] David J Thouless, Mahito Kohmoto, M Peter Nightingale, and Marcel den Nijs.

- Quantized hall conductance in a two-dimensional periodic potential. *Physical review letters*, 49(6):405, 1982.
- [57] Michael Victor Berry. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 392(1802):45–57, 1984.
- [58] Michael Berry. Geometric phase memories. *Nature Physics*, 6(3):148–150, 2010.
- [59] Barry Simon. Holonomy, the quantum adiabatic theorem, and berry’s phase. *Physical Review Letters*, 51(24):2167, 1983.
- [60] F Duncan M Haldane. Model for a quantum hall effect without landau levels: Condensed-matter realization of the” parity anomaly”. *Physical review letters*, 61(18):2015, 1988.
- [61] Cui-Zu Chang, Jinsong Zhang, Xiao Feng, Jie Shen, Zuocheng Zhang, Minghua Guo, Kang Li, Yunbo Ou, Pang Wei, Li-Li Wang, et al. Experimental observation of the quantum anomalous hall effect in a magnetic topological insulator. *Science*, 340(6129):167–170, 2013.
- [62] Charles L Kane and Eugene J Mele. Quantum spin hall effect in graphene. *Physical review letters*, 95(22):226801, 2005.
- [63] B Andrei Bernevig and Shou-Cheng Zhang. Quantum spin hall effect. *Physical review letters*, 96(10):106802, 2006.
- [64] B Andrei Bernevig, Taylor L Hughes, and Shou-Cheng Zhang. Quantum spin hall effect and topological phase transition in hgte quantum wells. *science*, 314(5806):1757–1761, 2006.
- [65] Markus König, Steffen Wiedmann, Christoph Brüne, Andreas Roth, Hart-

- mut Buhmann, Laurens W Molenkamp, Xiao-Liang Qi, and Shou-Cheng Zhang. Quantum spin hall insulator state in hgte quantum wells. *Science*, 318(5851):766–770, 2007.
- [66] M Zahid Hasan and Charles L Kane. Colloquium: topological insulators. *Reviews of modern physics*, 82(4):3045, 2010.
- [67] Liang Fu and Charles L Kane. Time reversal polarization and a z 2 adiabatic spin pump. *Physical Review B*, 74(19):195312, 2006.
- [68] Liang Fu and Charles L Kane. Topological insulators with inversion symmetry. *Physical Review B*, 76(4):045302, 2007.
- [69] Takahiro Fukui and Yasuhiro Hatsugai. Quantum spin hall effect in three dimensional materials: Lattice computation of z2 topological invariants and its application to bi and sb. *Journal of the Physical Society of Japan*, 76(5):053702–053702, 2007.
- [70] David Vanderbilt. *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators*. Cambridge University Press, 2018.
- [71] Liang Fu, C. L. Kane, and E. J. Mele. Topological Insulators in Three Dimensions. *Phys. Rev. Lett.*, 98(10):106803, 2007.
- [72] Joel E Moore and Leon Balents. Topological invariants of time-reversal-invariant band structures. *Physical Review B*, 75(12):121306, 2007.
- [73] Rahul Roy. Topological phases and the quantum spin hall effect in three dimensions. *Physical Review B*, 79(19):195322, 2009.
- [74] Douglas R Hartree. The wave mechanics of an atom with a non-coulomb central

- field. part i. theory and methods. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 89–110. Cambridge university press, 1928.
- [75] Vladimir Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik*, 61(1-2):126–148, 1930.
- [76] John C Slater. A simplification of the hartree-fock method. *Physical review*, 81(3):385, 1951.
- [77] Llewellyn H Thomas. The calculation of atomic fields. In *Mathematical proceedings of the Cambridge philosophical society*, volume 23, pages 542–548. Cambridge University Press, 1927.
- [78] Enrico Fermi. Eine statistische methode zur bestimmung einiger eigenschaften des atoms und ihre anwendung auf die theorie des periodischen systems der elemente. *Zeitschrift für Physik*, 48(1-2):73–79, 1928.
- [79] Paul AM Dirac. Note on exchange phenomena in the thomas atom. In *Mathematical proceedings of the Cambridge philosophical society*, volume 26, pages 376–385. Cambridge University Press, 1930.
- [80] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.
- [81] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.
- [82] John P Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10):5048, 1981.

- [83] John P Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical review B*, 45(23):13244, 1992.
- [84] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of physics*, 58(8):1200–1211, 1980.
- [85] David M Ceperley and Berni J Alder. Ground state of the electron gas by a stochastic method. *Physical review letters*, 45(7):566, 1980.
- [86] John P Perdew, Kieron Burke, and Yue Wang. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Physical review B*, 54(23):16533, 1996.
- [87] John P Perdew, Matthias Ernzerhof, and Kieron Burke. Rationale for mixing exact exchange with density functional approximations. *The Journal of chemical physics*, 105(22):9982–9985, 1996.
- [88] Jochen Heyd, Gustavo E Scuseria, and Matthias Ernzerhof. Hybrid functionals based on a screened coulomb potential. *The Journal of chemical physics*, 118(18):8207–8215, 2003.
- [89] Aliaksandr V Krukau, Oleg A Vydrov, Artur F Izmaylov, and Gustavo E Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of chemical physics*, 125(22):224106, 2006.
- [90] OH Nielsen and Richard M Martin. Quantum-mechanical theory of stress and force. *Physical Review B*, 32(6):3780, 1985.
- [91] Artem R Oganov. *Modern methods of crystal structure prediction*. John Wiley

& Sons, 2011.

- [92] Artem R Oganov, Andriy O Lyakhov, and Mario Valle. How evolutionary crystal structure prediction works— and why. *Accounts of chemical research*, 44(3):227–237, 2011.
- [93] Artem R Oganov and Colin W Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *The Journal of chemical physics*, 124(24):244704, 2006.
- [94] Colin W Glass, Artem R Oganov, and Nikolaus Hansen. Uspex - evolutionary crystal structure prediction. *Computer physics communications*, 175(11–12):713–720, 2006.
- [95] Andriy O Lyakhov, Artem R Oganov, Harold T Stokes, and Qiang Zhu. New developments in evolutionary structure prediction algorithm uspex. *Computer Physics Communications*, 184(4):1172–1182, 2013.
- [96] Jonathan Schmidt, Mário RG Marques, Silvana Botti, and Miguel AL Marques. Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials*, 5(1):1–36, 2019.
- [97] Teng Zhou, Zhen Song, and Kai Sundmacher. Big data creates new opportunities for materials research: A review on methods and applications of machine learning for materials design. *Engineering*, 5(6):1017–1026, 2019.
- [98] Lauri Himanen, Amber Geurts, Adam Stuart Foster, and Patrick Rinke. Data-driven materials science: status, challenges, and perspectives. *Advanced Science*, 6(21):1900808, 2019.
- [99] Siwar Chibani and François-Xavier Coudert. Machine learning approaches for

- the prediction of materials properties. *APL Materials*, 8(8):080701, 2020.
- [100] James E Saal, Anton O Oliynyk, and Bryce Meredig. Machine learning in materials discovery: confirmed predictions and their underlying approaches. *Annual Review of Materials Research*, 50:49–69, 2020.
 - [101] Jiazen Cai, Xuan Chu, Kun Xu, Hongbo Li, and Jing Wei. Machine learning-driven new material discovery. *Nanoscale Advances*, 2(8):3115–3130, 2020.
 - [102] Bryce Meredig, Ankit Agrawal, Scott Kirklin, James E Saal, Jeff W Doak, Alan Thompson, Kunpeng Zhang, Alok Choudhary, and Christopher Wolverton. Combinatorial screening for new materials in unconstrained composition space with machine learning. *Physical Review B*, 89(9):094104, 2014.
 - [103] Al’ona Furmanchuk, Ankit Agrawal, and Alok Choudhary. Predictive analytics for crystalline materials: bulk modulus. *RSC advances*, 6(97):95246–95251, 2016.
 - [104] Maarten De Jong, Wei Chen, Randy Notestine, Kristin Persson, Gerbrand Ceder, Anubhav Jain, Mark Asta, and Anthony Gamst. A statistical learning framework for materials science: application to elastic moduli of k-nary inorganic polycrystalline compounds. *Scientific reports*, 6(1):1–11, 2016.
 - [105] Olexandr Isayev, Corey Osse, Cormac Toher, Eric Gossett, Stefano Curtarolo, and Alexander Tropsha. Universal fragment descriptors for predicting properties of inorganic crystals. *Nature communications*, 8(1):1–12, 2017.
 - [106] Jack D Evans and François-Xavier Coudert. Predicting the mechanical properties of zeolite frameworks by machine learning. *Chemistry of Materials*, 29(18):7833–7839, 2017.

- [107] Aria Mansouri Tehrani, Anton O Oliynyk, Marcus Parry, Zeshan Rizvi, Samantha Couper, Feng Lin, Lowell Miyagi, Taylor D Sparks, and Jakoah Brgoch. Machine learning directed search for ultraincompressible, superhard materials. *Journal of the American Chemical Society*, 140(31):9844–9853, 2018.
- [108] Patrick Avery, Xiaoyu Wang, Corey Oses, Eric Gossett, Davide M Proserpio, Cormac Toher, Stefano Curtarolo, and Eva Zurek. Predicting superhard materials via a machine learning informed evolutionary structure search. *npj Computational Materials*, 5(1):1–11, 2019.
- [109] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1):011002, 2013.
- [110] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L Chevrier, Kristin A Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.
- [111] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.
- [112] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.

- [113] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54(16):11169, 1996.
- [114] Stefano Curtarolo, Wahyu Setyawan, Gus LW Hart, Michal Jahnatek, Roman V Chepulskii, Richard H Taylor, Shidong Wang, Junkai Xue, Kesong Yang, Ohad Levy, et al. Aflow: An automatic framework for high-throughput materials discovery. *Computational Materials Science*, 58:218–226, 2012.
- [115] Claudia Draxl and Matthias Scheffler. The nomad laboratory: from data sharing to artificial intelligence. *Journal of Physics: Materials*, 2(3):036001, 2019.
- [116] James E Saal, Scott Kirklin, Muratahan Aykol, Bryce Meredig, and Christopher Wolverton. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *Jom*, 65(11):1501–1509, 2013.
- [117] Logan Ward, Ankit Agrawal, Alok Choudhary, and Christopher Wolverton. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials*, 2(1):1–7, 2016.
- [118] Logan Ward, Alexander Dunn, Alireza Faghaninia, Nils ER Zimmermann, Saurabh Bajaj, Qi Wang, Joseph Montoya, Jiming Chen, Kyle Bystrom, Maxwell Dylla, et al. Matminer: An open source toolkit for materials data mining. *Computational Materials Science*, 152:60–69, 2018.
- [119] Wei-Chih Chen, Joanna N. Schmidt, Da Yan, Yogesh K. Vohra, and Cheng-Chien Chen. Github. https://github.com/weichihuab/ML_B-C-N, 2021.
- [120] Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:832–844,

1998.

- [121] Yali Amit and Donald Geman. Shape quantization and recognition with randomized trees. *Neural computation*, 9(7):1545–1588, 1997.
- [122] Faming Gao, Julong He, Erdong Wu, Shimin Liu, Dongli Yu, Dongchun Li, Siyuan Zhang, and Yongjun Tian. Hardness of covalent crystals. *Physical review letters*, 91(1):015502, 2003.
- [123] Antonín Šimnek and Jiří Vackář. Hardness of covalent and ionic crystals: first-principle calculations. *Physical review letters*, 96(8):085501, 2006.
- [124] Keyan Li, Xingtao Wang, Fangfang Zhang, and Dongfeng Xue. Electronegativity identification of novel superhard materials. *Physical Review Letters*, 100(23):235504, 2008.
- [125] Andriy O Lyakhov and Artem R Oganov. Evolutionary search for superhard materials: Methodology and applications to forms of carbon and tio 2. *Physical Review B*, 84(9):092103, 2011.
- [126] Haiyang Niu, Shiwei Niu, and Artem R Oganov. Simple and accurate model of fracture toughness of solids. *Journal of Applied Physics*, 125(6):065105, 2019.
- [127] Efim Mazhnik and Artem R Oganov. A model of hardness and fracture toughness of solids. *Journal of Applied Physics*, 126(12):125109, 2019.
- [128] Xing-Qiu Chen, Haiyang Niu, Dianzhong Li, and Yiyi Li. Modeling hardness of polycrystalline materials and bulk metallic glasses. *Intermetallics*, 19(9):1275–1281, 2011.
- [129] Yongjun Tian, Bo Xu, and Zhisheng Zhao. Microscopic theory of hardness and design of novel superhard crystals. *International Journal of Refractory Metals*

and Hard Materials, 33:93–106, 2012.

- [130] Yanchao Wang and Yanming Ma. Perspective: Crystal structure prediction at high pressures. *The Journal of chemical physics*, 140(4):040901, 2014.
- [131] Jake Graser, Steven K Kauwe, and Taylor D Sparks. Machine learning and energy minimization approaches for crystal structure predictions: A review and new horizons. *Chemistry of Materials*, 30(11):3601–3612, 2018.
- [132] Artem R Oganov, Chris J Pickard, Qiang Zhu, and Richard J Needs. Structure prediction drives materials discovery. *Nature Reviews Materials*, 4(5):331–348, 2019.
- [133] Hendrik J Monkhorst and James D Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188, 1976.
- [134] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, 1994.
- [135] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758–1775, 1999.
- [136] Yvon Le Page and Paul Saxe. Symmetry-general least-squares extraction of elastic data for strained materials from ab initio calculations of stress. *Physical Review B*, 65(10):104104, 2002.
- [137] W Voigt. Lehrbuch der kristallphysik. vol. 1. *Aufl. Teubner. Berlin/Leipzig*, 1928.
- [138] András Reuß. Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 9(1):49–

58, 1929.

- [139] Richard Hill. The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A*, 65(5):349, 1952.
- [140] Atsushi Togo and Isao Tanaka. First principles phonon calculations in materials science. *Scripta Materialia*, 108:1–5, 2015.
- [141] S.F. Pugh. Xcii. relations between the elastic moduli and the plastic properties of polycrystalline pure metals. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 45(367):823–843, 1954.
- [142] Kaleb C Burrage, Chia-Min Lin, Wei-Chih Chen, Cheng-Chien Chen, and Yogenesh K Vohra. Experimental and computational studies on superhard material rhenium diboride under ultrahigh pressures. *Materials*, 13(7):1657, 2020.
- [143] Kaleb C Burrage, Chia-Min Lin, Wei-Chih Chen, Cheng-Chien Chen, and Yogenesh K Vohra. Electronic structure and anisotropic compression of os₂b₃ to 358 gpa. *Journal of Physics: Condensed Matter*, 32(40):405703, 2020.
- [144] Marc Harper et al. python-ternary: Ternary plots in python. *Zenodo* 10.5281/zenodo.594435.
- [145] Koichi Momma and Fujio Izumi. Vesta 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Cryst.*, 44(6):1272–1276, 2011.
- [146] Xiaoguang Luo, Xiang-Feng Zhou, Zhongyuan Liu, Julong He, Bo Xu, Dongli Yu, Hui-Tian Wang, and Yongjun Tian. Refined crystal structure and mechanical properties of superhard bc₄n crystal: first-principles calculations. *The Journal of Physical Chemistry C*, 112(25):9516–9519, 2008.

- [147] Lulu Liu, Ziyuan Zhao, Tong Yu, Shoutao Zhang, Jianyan Lin, and Guochun Yang. Hexagonal bc₂n with remarkably high hardness. *The Journal of Physical Chemistry C*, 122(12):6801–6807, 2018.
- [148] Yansun Yao, S Tse John, and Dennis D Klug. Crystal and electronic structure of superhard bc 5: First-principles structural optimizations. *Physical Review B*, 80(9):094106, 2009.
- [149] Lifang Xu, Zhisheng Zhao, Li-Min Wang, Bo Xu, Julong He, Zhongyuan Liu, and Yongjun Tian. Prediction of a three-dimensional conductive superhard material: diamond-like bc₂. *The Journal of Physical Chemistry C*, 114(51):22688–22690, 2010.
- [150] Shiyou Chen, XG Gong, and Su-Huai Wei. Superhard pseudocubic bc 2 n superlattices. *Physical review letters*, 98(1):015502, 2007.
- [151] Félix Mouhat and François-Xavier Coudert. Necessary and sufficient elastic stability conditions in various crystal systems. *Physical review B*, 90(22):224104, 2014.
- [152] Quan Li, Hui Wang, Yongjun Tian, Yang Xia, Tian Cui, Julong He, Yanming Ma, and Guangtian Zou. Superhard and superconducting structures of bc 5. *Journal of Applied Physics*, 108(2):023507, 2010.
- [153] Mark E Eberhart and Travis E Jones. Cauchy pressure and the generalized bonding model for nonmagnetic bcc transition metals. *Physical Review B*, 86(13):134106, 2012.
- [154] Shivakumar I Ranganathan and Martin Ostoja-Starzewski. Universal elastic anisotropy index. *Physical review letters*, 101(5):055504, 2008.

- [155] Muratahan Aykol, Shyam S Dwaraknath, Wenhao Sun, and Kristin A Persson. Thermodynamic limit for synthesis of metastable inorganic materials. *Science advances*, 4(4):eaaq0148, 2018.
- [156] Wei-Chih Chen, Joanna N Schmidt, Da Yan, Yogesh K Vohra, and Cheng-Chien Chen. Machine learning and evolutionary prediction of superhard bcn compounds. *npj Computational Materials*, 7(1):1–8, 2021.
- [157] Ravi Shankar, Sofia Marchesini, Erich A Muller, and Camille Petit. Gas adsorption in amorphous porous boron oxynitride: Grand canonical monte carlo simulations and experimental determination. 2021.
- [158] Junkai Ren, Luca Malfatti, Stefano Enzo, Carlo Maria Carbonaro, Laura Calvillo, Gaetano Granozzi, and Plinio Innocenzi. Boron oxynitride two-colour fluorescent dots and their incorporation in a hybrid organic-inorganic film. *Journal of colloid and interface science*, 560:398–406, 2020.
- [159] Ying Peng Xie, Gang Liu, Gao Qing Max Lu, and Hui-Ming Cheng. Boron oxynitride nanoclusters on tungsten trioxide as a metal-free cocatalyst for photocatalytic oxygen evolution from water splitting. *Nanoscale*, 4(4):1267–1270, 2012.
- [160] Corey L Arnold, Chukwudi E Iheomamere, Maddox Dockins, Spencer Gellerup, Nicholas R Glavin, Christopher Muratore, Nigel D Shepherd, and Andrey A Voevodin. Composition, dielectric breakdown, and bandgap of ultra-thin amorphous boron oxynitride produced by magnetron sputtering. *Vacuum*, 188:110211, 2021.
- [161] Marc Dussauze, Efstratios I Kamitsos, Patrik Johansson, Aleksandar Matic, Christos-Platon E Varsamis, Dominique Cavagnat, Philippe Vinatier, and

- Yohann Hamon. Lithium ion conducting boron-oxynitride amorphous thin films: synthesis and molecular structure by infrared spectroscopy and density functional theory modeling. *The Journal of Physical Chemistry C*, 117(14):7202–7213, 2013.
- [162] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.
- [163] Sobhit Singh, Irais Valencia-Jaime, Olivia Pavlic, and Aldo H. Romero. Elastic, mechanical, and thermodynamic properties of bi-sb binaries: Effect of spin-orbit coupling. *Phys. Rev. B*, 97:054108, 2018.
- [164] Sobhit Singh, Logan Lang, Viviana Dovale-Farelo, Uthpala Herath, Pedram Tavadze, François-Xavier Coudert, and Aldo H. Romero. Mechelastic: A python library for analysis of mechanical and elastic properties of bulk and 2d materials. arxiv 2012.04758, 2021.
- [165] Axel D. Becke and Erin R. Johnson. A simple effective potential for exchange. *The Journal of Chemical Physics*, 124(22):221101, 2006.
- [166] Fabien Tran and Peter Blaha. Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation potential. *Phys. Rev. Lett.*, 102:226401, 2009.
- [167] MI Sayyed, G Lakshminarayana, M Moghaddasi, IV Kityk, and MA Mahdi. Physical properties, optical band gaps and radiation shielding parameters exploration for dy 3+-doped alkali/mixed alkali multicomponent borate glasses. *Glass Physics and Chemistry*, 44(4):279–291, 2018.
- [168] Shuichi Murakami and Shun-ichi Kuga. Universal phase diagrams for the quantum spin hall systems. *Physical Review B*, 78(16):165313, 2008.

- [169] Shuichi Murakami. Phase transition between the quantum spin hall and insulator phases in 3d: emergence of a topological gapless phase. *New Journal of Physics*, 9(9):356, 2007.
- [170] Hongming Weng, Xi Dai, and Zhong Fang. Transition-metal pentatelluride zrte 5 and hfte 5: A paradigm for large-gap quantum spin hall insulators. *Physical review X*, 4(1):011002, 2014.
- [171] RY Chen, SJ Zhang, JA Schneeloch, C Zhang, Q Li, GD Gu, and NL Wang. Optical spectroscopy study of the three-dimensional dirac semimetal zrte 5. *Physical Review B*, 92(7):075107, 2015.
- [172] Yanwen Liu, Xiang Yuan, Cheng Zhang, Zhao Jin, Awadhesh Narayan, Chen Luo, Zhigang Chen, Lei Yang, Jin Zou, Xing Wu, et al. Zeeman splitting and dynamical mass generation in dirac semimetal zrte 5. *Nature communications*, 7(1):1–9, 2016.
- [173] Nityan L Nair, Philipp T Dumitrescu, Sanyum Channa, Sinéad M Griffin, Jeffrey B Neaton, Andrew C Potter, and James G Analytis. Thermodynamic signature of dirac electrons across a possible topological transition in zrte 5. *Physical Review B*, 97(4):041111, 2018.
- [174] Qiang Li, Dmitri E Kharzeev, Cheng Zhang, Yuan Huang, I Pletikosić, AV Fedorov, RD Zhong, JA Schneeloch, GD Gu, and T Valla. Chiral magnetic effect in zrte 5. *Nature Physics*, 12(6):550–554, 2016.
- [175] Xiang-Bing Li, Wen-Kai Huang, Yang-Yang Lv, Kai-Wen Zhang, Chao-Long Yang, Bin-Bin Zhang, YB Chen, Shu-Hua Yao, Jian Zhou, Ming-Hui Lu, et al. Experimental observation of topological edge states at the surface step edge of the topological insulator zrte 5. *Physical review letters*, 116(17):176803, 2016.

- [176] Zhi-Guo Chen, RY Chen, RD Zhong, John Schneeloch, C Zhang, Y Huang, Fanming Qu, Rui Yu, Q Li, GD Gu, et al. Spectroscopic evidence for bulk-band inversion and three-dimensional massive dirac fermions in zrte5. *Proceedings of the National Academy of Sciences*, 114(5):816–821, 2017.
- [177] Y Jiang, ZL Dun, HD Zhou, Z Lu, K-W Chen, S Moon, Tiglet Besara, TM Siegrist, RE Baumbach, D Smirnov, et al. Landau-level spectroscopy of massive dirac fermions in single-crystalline zrte 5 thin flakes. *Physical Review B*, 96(4):041101, 2017.
- [178] H Xiong, JA Sobota, S-L Yang, H Soifer, A Gauthier, M-H Lu, Y-Y Lv, S-H Yao, D Lu, M Hashimoto, et al. Three-dimensional nature of the band structure of zrte 5 measured by high-momentum-resolution photoemission spectroscopy. *Physical Review B*, 95(19):195119, 2017.
- [179] Bing Xu, LX Zhao, P Marsik, Evgenila Sheveleva, F Lyzwa, YM Dai, GF Chen, XG Qiu, and Christian Bernhard. Temperature-driven topological phase transition and intermediate dirac semimetal phase in zrte 5. *Physical review letters*, 121(18):187401, 2018.
- [180] Su-Yang Xu, Y Xia, LA Wray, S Jia, F Meier, JH Dil, J Osterwalder, B Slomski, A Bansil, H Lin, et al. Topological phase transition and texture inversion in a tunable topological insulator. *Science*, 332(6029):560–564, 2011.
- [181] Liang Wu, M Brahlek, R Valdés Aguilar, AV Stier, CM Morris, Y Lubashevsky, LS Bilbro, N Bansal, Seongshik Oh, and NP Armitage. A sudden collapse in the transport lifetime across the topological phase transition in (bi 1- x in x) 2 se 3. *Nature Physics*, 9(7):410–414, 2013.
- [182] P Dziawa, BJ Kowalski, K Dybko, R Buczko, A Szczerbakow, M Szot,

- E Łusakowska, T Balasubramanian, Bastian M Wojek, MH Berntsen, et al. Topological crystalline insulator states in pb_{1-x}sn_xse. *Nature materials*, 11(12):1023–1027, 2012.
- [183] Su-Yang Xu, Chang Liu, N Alidoust, M Neupane, D Qian, I Belopolski, JD Denlinger, YJ Wang, H Lin, LA a Wray, et al. Observation of a topological crystalline insulator phase and topological phase transition in pb_{1-x}sn_xte. *Nature communications*, 3(1):1–11, 2012.
- [184] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. Quantum espresso: a modular and open-source software project for quantum simulations of materials. *Journal of physics: Condensed matter*, 21(39):395502, 2009.
- [185] H Fjellvåg and A Kjekshus. Structural properties of zrte5 and hfte5 as seen by powder diffraction. *Solid State Communications*, 60(2):91–93, 1986.
- [186] Arash A. Mostofi, Jonathan R. Yates, Giovanni Pizzi, Young-Su Lee, Ivo Souza, David Vanderbilt, and Nicola Marzari. An updated version of wannier90: A tool for obtaining maximally-localised wannier functions. *Computer Physics Communications*, 185(8):2309–2310, 2014.
- [187] QuanSheng Wu, ShengNan Zhang, Hai-Feng Song, Matthias Troyer, and Alexey A. Soluyanov. Wanniertools : An open-source software package for novel topological materials. *Computer Physics Communications*, 224:405 – 416, 2018.
- [188] Timo Thonhauser, Valentino R Cooper, Shen Li, Aaron Puzder, Per Hyldgaard, and David C Langreth. Van der waals density functional: Self-consistent poten-

- tial and the nature of the van der waals bond. *Physical Review B*, 76(12):125112, 2007.
- [189] Kristian Berland, Valentino R Cooper, Kyuho Lee, Elsebeth Schröder, T Thonhauser, Per Hyldgaard, and Bengt I Lundqvist. van der waals forces in density functional theory: a review of the vdw-df method. *Reports on Progress in Physics*, 78(6):066501, 2015.
- [190] A Otero-De-La-Roza and Erin R Johnson. Van der waals interactions in solids using the exchange-hole dipole moment model. *The Journal of chemical physics*, 136(17):174109, 2012.
- [191] Alexey A. Soluyanov and David Vanderbilt. Computing topological invariants without inversion symmetry. *Phys. Rev. B*, 83:235401, 2011.
- [192] Zongjian Fan, Qi-Feng Liang, YB Chen, Shu-Hua Yao, and Jian Zhou. Transition between strong and weak topological insulator in zrte 5 and hfte 5. *Scientific reports*, 7(1):1–7, 2017.
- [193] Joshua Mutch, Wei-Chih Chen, Preston Went, Tiema Qian, Ilham Zaky Wilson, Anton Andreev, Cheng-Chien Chen, and Jiun-Haw Chu. Evidence for a strain-tuned topological phase transition in zrte5. *Science Advances*, 5(8), 2019.
- [194] F. F. Tafti, Q. D. Gibson, S. K. Kushwaha, N. Haldolaarachchige, and R. J. Cava. Resistivity plateau and extreme magnetoresistance in lasb. *Nature Physics*, 12(3):272–277, 2016.
- [195] X. H. Niu, D. F. Xu, Y. H. Bai, Q. Song, X. P. Shen, B. P. Xie, Z. Sun, Y. B. Huang, D. C. Peets, and D. L. Feng. Presence of exotic electronic surface states in labi and lasb. *Phys. Rev. B*, 94:165163, 2016.

- [196] Fazel Fallah Tafti, Quinn Gibson, Satya Kushwaha, Jason W. Krizan, Neel Haldolaarachchige, and Robert Joseph Cava. Temperature-field phase diagram of extreme magnetoresistance. *Proceedings of the National Academy of Sciences*, 113(25):E3475–E3481, 2016.
- [197] H.-Y. Yang, T. Nummy, H. Li, S. Jaszewski, M. Abramchuk, D. S. Dessau, and Fazel Tafti. Extreme magnetoresistance in the topologically trivial lanthanum monopnictide laas. *Phys. Rev. B*, 96:235128, 2017.
- [198] Yu Zhou, Wang-Li Tao, Zhao-Yi Zeng, Xiang-Rong Chen, and Qi-Feng Chen. Thermoelectric properties of topological insulator lanthanum phosphide via first-principles study. *Journal of Applied Physics*, 125(4):045107, 2019.
- [199] Chia-Min Lin, Wei-Chih Chen, and Cheng-Chien Chen. First-principles study of strain effect on thermoelectric properties of lap and laas. arxiv 2008.06455, 2020.
- [200] Heng Gao, Jörn W.F. Venderbos, Youngkuk Kim, and Andrew M. Rappe. Topological semimetals from first principles. *Annual Review of Materials Research*, 49(1):153–183, 2019.
- [201] Binghai Yan and Claudia Felser. Topological materials: Weyl semimetals. *Annu. Rev. Condens. Matter Phys.*, 8(1):337–354, 2017.
- [202] N.P. Armitage, E.J. Mele, and Ashvin Vishwanath. Weyl and Dirac semimetals in three-dimensional solids. *Rev. Mod. Phys.*, 90(1):015001, 2018.
- [203] Chen Fang, Hongming Weng, Xi Dai, and Zhong Fang. Topological nodal line semimetals. *Chin. Phys. B*, 25(11):117106, 2016.
- [204] Alexey A. Soluyanov, Dominik Gresch, Zhijun Wang, QuanSheng Wu, Matthias

- Troyer, Xi Dai, and B. Andrei Bernevig. Type-II Weyl semimetals. *Nature*, 527(7579):495–498, 2015.
- [205] Chen Fang, Matthew J. Gilbert, Xi Dai, and B. Andrei Bernevig. Multi-Weyl Topological Semimetals Stabilized by Point Group Symmetry. *Phys. Rev. Lett.*, 108(26):266802, 2012.
- [206] Benjamin J. Wieder, Youngkuk Kim, A.M. Rappe, and C.L. Kane. Double Dirac Semimetals in Three Dimensions. *Phys. Rev. Lett.*, 116(18):186402, 2016.
- [207] Barry Bradlyn, Jennifer Cano, Zhijun Wang, M. G. Vergniory, C. Felser, R. J. Cava, and B. Andrei Bernevig. Beyond Dirac and Weyl fermions: Unconventional quasiparticles in conventional crystals. *Science*, 353(6299), 2016.
- [208] M. Z. Hasan and C. L. Kane. Colloquium: Topological insulators. *Rev. Mod. Phys.*, 82:3045–3067, 2010.
- [209] Xiao-Liang Qi and Shou-Cheng Zhang. Topological insulators and superconductors. *Rev. Mod. Phys.*, 83(4):1057–1110, 2011.
- [210] Yun Wu, Tai Kong, Lin-Lin Wang, D. D. Johnson, Daixiang Mou, Lunan Huang, Benjamin Schrunk, S. L. Bud'ko, P. C. Canfield, and Adam Kaminski. Asymmetric mass acquisition in labi: Topological semimetal candidate. *Phys. Rev. B*, 94:081108, 2016.
- [211] Jayita Nayak, Shu-Chun Wu, Nitesh Kumar, Chandra Shekhar, Sanjay Singh, Jörg Fink, Emile E. D. Rienks, Gerhard H. Fecher, Stuart S. P. Parkin, Binghai Yan, and Claudia Felser. Multiple dirac cones at the surface of the topological metal labi. *Nature Communications*, 8(1):13942, 2017.
- [212] R. Lou, B.-B. Fu, Q. N. Xu, P.-J. Guo, L.-Y. Kong, L.-K. Zeng, J.-Z. Ma,

- P. Richard, C. Fang, Y.-B. Huang, S.-S. Sun, Q. Wang, L. Wang, Y.-G. Shi, H. C. Lei, K. Liu, H. M. Weng, T. Qian, H. Ding, and S.-C. Wang. Evidence of topological insulator state in the semimetal labi. *Phys. Rev. B*, 95:115140, 2017.
- [213] Baojie Feng, Jin Cao, Meng Yang, Ya Feng, Shilong Wu, Botao Fu, Masashi Arita, Koji Miyamoto, Shaolong He, Kenya Shimada, Youguo Shi, Taichi Okuda, and Yugui Yao. Experimental observation of node-line-like surface states in labi. *Phys. Rev. B*, 97:155153, 2018.
- [214] Thomas J. Nummy, Justin A. Waugh, Stephen P. Parham, Qihang Liu, Hung-Yu Yang, Haoxiang Li, Xiaoqing Zhou, Nicholas C. Plumb, Fazel F. Tafti, and Daniel S. Dessau. Measurement of the atomic orbital composition of the near-fermi-level electronic states in the lanthanum monopnictides labi, lasb, and laas. *npj Quantum Materials*, 3(1):24, 2018.
- [215] L.-K. Zeng, R. Lou, D.-S. Wu, Q. N. Xu, P.-J. Guo, L.-Y. Kong, Y.-G. Zhong, J.-Z. Ma, B.-B. Fu, P. Richard, P. Wang, G. T. Liu, L. Lu, Y.-B. Huang, C. Fang, S.-S. Sun, Q. Wang, L. Wang, Y.-G. Shi, H. M. Weng, H.-C. Lei, K. Liu, S.-C. Wang, T. Qian, J.-L. Luo, and H. Ding. Compensated semimetal lasb with unsaturated magnetoresistance. *Phys. Rev. Lett.*, 117:127204, 2016.
- [216] Zhijun Wang, Yan Sun, Xing-Qiu Chen, Cesare Franchini, Gang Xu, Hongming Weng, Xi Dai, and Zhong Fang. Dirac semimetal and topological phase transitions in $A_3\text{bi}$ ($a = \text{Na}, \text{k}, \text{rb}$). *Phys. Rev. B*, 85:195320, 2012.
- [217] Bartomeu Monserrat, Joseph W Bennett, Karin M Rabe, and David Vanderbilt. Antiferroelectric topological insulators in orthorhombic $a\text{mgb}_3$ compounds ($a = \text{li}, \text{na}, \text{k}$). *Physical review letters*, 119(3):036802, 2017.

- [218] Edbert J Sie, Clara M Nyby, CD Pemmaraju, Su Ji Park, Xiaozhe Shen, Jie Yang, Matthias C Hoffmann, BK Ofori-Okai, Renkai Li, Alexander H Reid, et al. An ultrafast symmetry switch in a weyl semimetal. *Nature*, 565(7737):61–66, 2019.
- [219] Chirag Vaswani, L-L Wang, Dinusha Herath Mudiyanselage, Q Li, PM Lozano, GD Gu, Di Cheng, Boqun Song, Liang Luo, Richard HJ Kim, et al. Light-driven raman coherence as a nonthermal route to ultrafast topology switching in a dirac semimetal. *Physical Review X*, 10(2):021013, 2020.
- [220] Jonathan M. DeStefano and Lin-Lin Wang. Pressure effect on band inversion in $aecd_2as_2$ ($a = \text{Ca}, \text{Sr}, \text{Ba}$). *Phys. Rev. B*, 103:115207, 2021.
- [221] Shoaib Khalid, Fernando P. Sabino, and Anderson Janotti. Topological phase transition in laas under pressure. *Phys. Rev. B*, 98:220102, 2018.
- [222] Peng-Jie Guo, Huan-Cheng Yang, Kai Liu, and Zhong-Yi Lu. Theoretical study of the pressure-induced topological phase transition in lasb. *Phys. Rev. B*, 96:081112, 2017.
- [223] Mohamed Hacene, Ani Anciaux-Sedrakian, Xavier Rozanska, Diego Klahr, Thomas Guignon, and Paul Fleurat-Lessard. Accelerating vasp electronic structure calculations using graphic processing units. *Journal of Computational Chemistry*, 33(32):2581–2589, 2012.
- [224] M. Hutchinson and M. Widom. Vasp on a gpu: Application to exact-exchange calculations of the stability of elemental boron. *Computer Physics Communications*, 183(7):1422 – 1426, 2012.
- [225] Olle Hellman, Peter Steneteg, I. A. Abrikosov, and S. I. Simak. Temperature dependent effective potential method for accurate free energy calculations of

- solids. *Phys. Rev. B*, 87:104111, 2013.
- [226] T Tadano, Y Gohda, and S Tsuneyuki. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *Journal of Physics: Condensed Matter*, 26(22):225402, 2014.
- [227] Aliaksandr V. Krukau, Oleg A. Vydrov, Artur F. Izmaylov, and Gustavo E. Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of Chemical Physics*, 125(22):224106, 2006.
- [228] Dominik Gresch, Gabriel Autès, Oleg V. Yazyev, Matthias Troyer, David Vanderbilt, B. Andrei Bernevig, and Alexey A. Soluyanov. Z2pack: Numerical implementation of hybrid wannier centers for identifying topological materials. *Phys. Rev. B*, 95:075146, 2017.
- [229] Sebastian B. Schneider, Dominik Baumann, Ashkan Salamat, and Wolfgang Schnick. Reversible high-pressure phase transition in lan. *Journal of Applied Physics*, 111(9):093503, 2012.
- [230] Bärbel Krause, Dmitry S. Kuznetsov, Andrey E. Yakshin, Shyjumon Ibrahimkutty, Tilo Baumbach, and Fred Bijkerk. *In situ* and real-time monitoring of structure formation during non-reactive sputter deposition of lanthanum and reactive sputter deposition of lanthanum nitride. *Journal of Applied Crystallography*, 51(4):1013–1020, 2018.
- [231] Yvon Le Page and Paul Saxe. Symmetry-general least-squares extraction of elastic data for strained materials from ab initio calculations of stress. *Phys. Rev. B*, 65:104104, 2002.
- [232] Félix Mouhat and François-Xavier Coudert. Necessary and sufficient elastic

- stability conditions in various crystal systems. *Phys. Rev. B*, 90:224104, 2014.
- [233] Gökhan Gökoğlu and Aytaç Erkişi. Lattice dynamics and elastic properties of lanthanum monopnictides. *Solid state communications*, 147(5-6):221–225, 2008.
- [234] D. Mukherjee, B.D. Sahoo, K.D. Joshi, and Satish C. Gupta. Stabilization of tetragonal phase in lan under high pressure via peierls distortion. *High Pressure Research*, 33(3):563–571, 2013.
- [235] Sebastian B. Schneider, Dominik Baumann, Ashkan Salamat, and Wolfgang Schnick. Reversible high-pressure phase transition in lan. *Journal of Applied Physics*, 111(9):093503, 2012.
- [236] Hulei Yu, Dexiang Gao, Xiancheng Wang, Xueyan Du, Xiaohuan Lin, Wenhan Guo, Ruqiang Zou, Changqing Jin, Kuo Li, and Yue Chen. Unraveling a novel ferroelectric gese phase and its transformation into a topological crystalline insulator under high pressure. *NPG Asia Materials*, 10(9):882–887, 2018.
- [237] Yi Xia and Maria K. Y. Chan. Anharmonic stabilization and lattice heat transport in rocksalt -gete. *Applied Physics Letters*, 113(19):193902, 2018.
- [238] R. D. King-Smith and David Vanderbilt. Theory of polarization of crystalline solids. *Phys. Rev. B*, 47:1651–1654, 1993.
- [239] Joseph Maciejko, Taylor L. Hughes, and Shou-Cheng Zhang. The Quantum Spin Hall Effect. *Annu. Rev. Condens. Matter Phys.*, 2(1):31–53, 2011.
- [240] Minggang Zeng, Chen Fang, Guoqing Chang, Yu-An Chen, Timothy Hsieh, Arun Bansil, Hsin Lin, and Liang Fu. Topological semimetals and topological insulators in rare earth monopnictides. arxiv 1504.03492, 2015.
- [241] Peng-Jie Guo, Huan-Cheng Yang, Bing-Jing Zhang, Kai Liu, and Zhong-Yi Lu.

- Charge compensation in extremely large magnetoresistance materials lasb and labi revealed by first-principles calculations. *Phys. Rev. B*, 93:235142, 2016.
- [242] Urmimala Dey, Monodeep Chakraborty, A. Taraphder, and Sumanta Tewari. Bulk band inversion and surface dirac cones in lasb and labi: Prediction of a new topological heterostructure. *Scientific Reports*, 8(1):14867, 2018.
- [243] Shuichi Murakami. Phase transition between the quantum spin Hall and insulator phases in 3D: emergence of a topological gapless phase. *New J. Phys.*, 9(9):356, 2007.
- [244] Shuichi Murakami and Shun-ichi Kuga. Universal phase diagrams for the quantum spin Hall systems. *Phys. Rev. B*, 78(16):165313, 2008.
- [245] Wei-Chih Chen, Chia-Min Lin, Joseph Maciejko, and Cheng-Chien Chen. Lan structural and topological transitions driven by temperature and pressure. arxiv 2104.12879, 2021.