








Structural phase transition of monochalcogenides investigated with machine learning

J. Zhang ^{1,2}, F. Zhang ^{1,2}, D. Wei ^{1,2}, L. Liu ³, X. Liu,⁴ D. Fang ⁵, G. X. Zhang ⁶, X. Chen,⁷ and D. Wang ^{1,2,*}

¹*School of Microelectronics & State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China*

²*Key Lab of Micro-Nano Electronics and System Integration of Xi'an City, Xi'an Jiaotong University, Xi'an 710049, China*

³*Guangxi Key Laboratory of Optical and Electronic Materials and Devices, College of Materials Science and Engineering, Guilin University of Technology, Guilin 541004, China*

⁴*Engineering Research Center of Nanoelectronic Integration and Advanced Equipment, Ministry of Education, School of Physics and Electronic Science, East China Normal University, Shanghai 200062, China*

⁵*MOE Key Laboratory for Nonequilibrium Synthesis and Modulation of Condensed Matter, School of Physics, Xi'an Jiaotong University, Xi'an 710049, China*

⁶*School of Chemistry and Chemical Engineering, Harbin Institute of Technology, Harbin 150001, China*

⁷*Department of Applied Physics, Aalto University, Espoo 00076, Finland*



(Received 21 October 2021; revised 4 March 2022; accepted 7 March 2022; published 28 March 2022)

As machine learning becomes increasingly important in science and engineering, it holds the promise to provide a universal approach applicable to various systems to investigate their crystalline phase transitions. Here, we build and train accurate artificial neural networks that can distinguish tiny energy difference, which is crucial to predict the crystalline phase transitions. Employing the trained artificial neural networks in Monte Carlo simulations as the surrogate energy function, we apply this approach to monochalcogenides, including bulk and two-dimensional monolayer SnTe and GeTe, investigating their crystalline phase transitions. The machine-learning approach, when viewed as providing a universal mathematical structure, can be transferred to the investigation of other materials when the training data set generated with *ab initio* methods are available. Moreover, the machine-learning approach resolves the difficulties associated with constructing the effective Hamiltonian for monochalcogenides, showing great potential with its accuracy and efficiency.

DOI: [10.1103/PhysRevB.105.094116](https://doi.org/10.1103/PhysRevB.105.094116)

I. INTRODUCTION

When atoms in a crystal are displaced globally, a structural phase transition occurs, which changes its crystallographic structure. Since it is usually accompanied by a change in some macroscopic properties such as the dielectric susceptibility, the structural phase transition is an important phenomenon [1]. A particular class of material, that is, the ferroelectric materials, are well known for their crystalline phase transitions that are often associated with the change from the paraelectric to the ferroelectric phase where ion displacements play a crucial role [2]. Since ferroelectric materials have spontaneous polarization that can be reversed by an external electric field, they are important functional materials for many applications, e.g., ultrafast switches, phased-array radar, and dynamic random-access memories [3]. Interestingly, it has been discovered that many monochalcogenides, including SnTe, SnSe, GeTe, are also ferroelectric materials [4,5] with structural changes accompanying their phase transitions. In this work, we propose to apply the machine-learning (ML) approach to investigate the structural phase transition of monochalcogenides.

The crystalline phase transitions are usually described by a phenomenological theory, most notably the Landau-type

theory, which have been used for more than 50 years for ferroelectric materials [6]. With the advent of first-principles calculation, microscopic theory was proposed. For ferroelectric materials, it was the effective Hamiltonian approach proposed in the late 1990s [7–9] which critically depends on the concept of soft mode [10,11]. For ferroelectric perovskites, the effective Hamiltonian has been a popular method to investigate their static and dynamic properties and achieved great successes [12–18]. In practice, the effective Hamiltonian approach needed to identify the most important degrees of freedom in a system, and then construct the internal energy of the system as a function of these degrees of freedom and their interactions based on symmetry arguments. The coefficients entering the effective Hamiltonian energy are typically obtained by first-principles computations [9].

Since the effective Hamiltonians were mostly designed for ferroelectric perovskites, for other types of materials with structural phase transition, it may need to be constructed again from scratch. When it involves many degrees of freedom, the construction process can be difficult since all the dynamic variables need to be specified and proper formulas, along with the coefficients, need to be proposed, neither of which is trivial. In addition, for this first-principles-based approach, it is hard to achieve a transferability since each system may require a different formula to describe the energy. We note that, while the effective Hamiltonian is popular for perovskites, there is no such a thing for ferroelectric pyrochlore ($A_2B_2O_7$). A more

*dawei.wang@xjtu.edu.cn

universal approach, not limited to perovskites, is therefore desired. For instance, in recent years, some new force fields, e.g., the ReaxFF, have been applied to ferroelectrics, even for defective systems [19–21].

The need to have a more unified approach was also demonstrated in recent years with monochalcogenides. In 2016, it was reported that SnTe thin films with a thickness of 1 unit cell (UC, with two layers of atoms) can have stable spontaneous polarization up to 270 K, and 2-UC to 4-UC SnTe films also have strong ferroelectric properties at room temperature [22,23]. The ferroelectricity of this two-dimensional (2D) material makes it a promising candidate for applications in devices such as high-density memory nanosensors [24,25]. Monochalcogenides have also been used for their photovoltaic effect at room temperature [26,27] or as phase change materials [28]. Its phase-transition sequence and the corresponding phase-transition temperatures are also important information as tuning the phase transition to room temperature may enhance its performance [22]. Since accurately predicting the phase-transition temperature requires simulating large systems with hundreds of atoms, it is usually not possible with a pure *ab initio* approach, such as the density functional theory (DFT) [29], which is computationally intensive [30]. To overcome this limitation, empirical formulas need to be proposed to fit the interatomic potential energy (or force) with DFT results using small systems. As one example, the effective Hamiltonian approach was adopted for the investigation of SnTe thin films recently [31].

While the effective Hamiltonian approach was initially developed for perovskites, its variations, including the Landau-theory-based and the patched effective Hamiltonian [31,32], have been applied to the investigation of SnTe, GeTe, and SnSe [5,31,33,34]. Given the abundance of perovskites and monochalcogenides, a more unified approach, which is applicable to diversified systems, is sought after. However, considerable challenges remain as the electronic properties of monochalcogenides are unique [35] and different from most perovskites. The difficulties of constructing such approach and the underlying reason will be discussed in Sec. IV C.

Recent advances in ML have provided the opportunity to construct, in a unified fashion, the energy function of a given system using artificial neural networks (ANNs) [36,37]. After training with a data set obtained from electronic structure calculations with DFT [38,39], the ANNs can accurately predict the potential-energy surface (PES) of a given system. The well-trained ANN has the advantage that it can be employed for the simulation of different systems with minimal modification because it is a universal mathematical structure [40]. With such an approach, there is no need to reconstruct the formulas for different systems, while the accuracy of the ANNs and the generated PES remain satisfactory. We note that, for the effective Hamiltonian approach, while the parameters depend on DFT calculations, the energy terms (in other words the model) need to be built beforehand. In contrast, the use of ANNs will combine the advantages of both approaches, i.e., the accuracy of DFT and the efficiency of explicit formulas in obtaining the PES [41].

While ML approaches have been applied to many interesting systems [42–45], the crystalline phase transitions induced by atomic displacements, which are important in ferroelec-

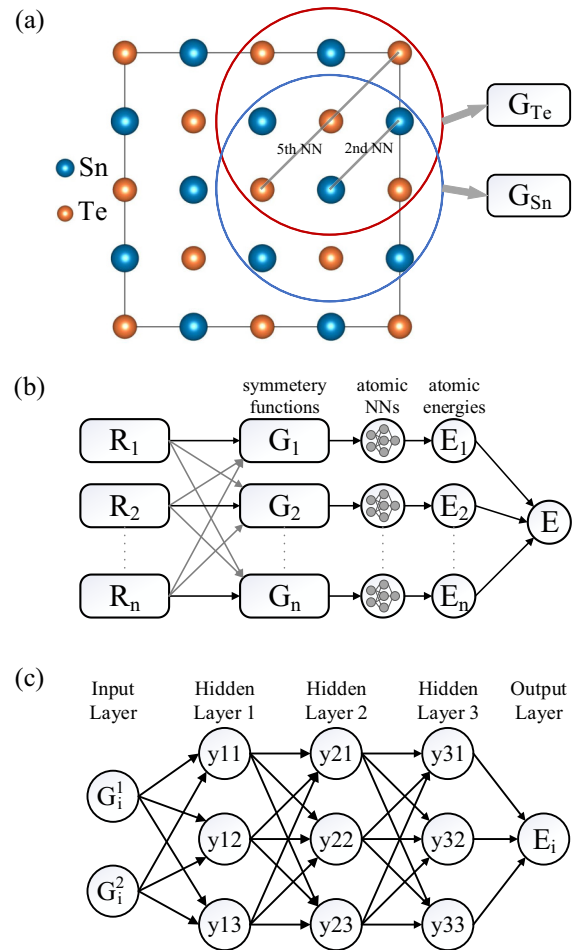


FIG. 1. (a) The 2D monolayer SnTe shows alternating Sn and Te atoms along the x and y axes. The blue (red) ring indicates the chemical environments encircling the Sn (Te) atom at the center. (b) The topology of the whole structure is built to predict the total energy, which contains the ANNs that are schematically shown in (c).

tric materials, have not been dealt with. One difficulty could be that the crystalline phase transitions are often associated with tiny energy difference [<1 meV/atom, see Fig. 4(a)], therefore requiring the ANNs to be accurate. In this work, we develop an accurate neural-network-based approach to treat the crystalline phase transitions [30,46,47], focusing on the structural phase transitions in crystals induced by ion displacement whereas most previous studies have focused on discovering new structures in liquid and amorphous phases [48–53]. In particular, highly accurate energy prediction in the vicinity of the double potential well, which is crucial for crystalline phase transitions, has not been thoroughly investigated using ML, particularly, no such studies on SnTe.

In this work, using the ML approach, we investigate both SnTe and GeTe, including monolayer systems and bulks. Practically, we build ANNs suitable for two different types of atoms (Sn/Ge and Te), and take DFT results as training data for the supervised learning of the ANNs. In this way, we have successfully constructed the necessary PES (with respect to the displacement of atoms) and employed it in Monte Carlo (MC) simulations to find their crystalline phase transitions. We demonstrate that the setup employing ANN (see Fig. 1) is

quite general that it can be transferred to investigate different systems. Particularly, we show that a monolayer SnTe has a phase transition when all the atoms are confined to 2D and the structural phase transition occurring at ~ 200 K.

This paper is organized as follows. In Sec. II, we provide details about the approach adopted in this work. In Sec. III, we present the results for SnTe and GeTe, showing the quality of the ANNs and their crystalline phase transitions. In Sec. IV, we discuss the results and the potentials of using ANNs as the surrogate energy function. Finally, in Sec. V, we briefly summarize this work.

II. METHOD

The key ingredient in constructing the PES is to build and train an ANN that can efficiently and accurately predict the energy of a given atomic configuration, similar to what an effective Hamiltonian does. As an example, a monolayer SnTe has the structure as shown in Fig. 1(a) where Sn and Te atoms alternate with each other along both the [100] and [010] directions. When an Sn or Te atom moves, the ANNs shall be able to predict the corresponding energy change.

To fulfill this mission, the whole architecture for training the ANNs has the topology shown in Fig. 1(b), which is created by Behler and Parrinello [38,54]. The total energy E of the system is the sum of atomic contributions E_i [55]. The mapping of atomic energy $E = \sum_i E_i$ proved to be feasible [56], where E_i is the energy imposed on the i th atom by its neighboring environment, which will be determined by the ANN we build. It is realized that the interatomic potential energy decays rapidly with distance (which is termed the “nearsightedness” [57,58]), therefore, a cutoff function is usually used to limit the interaction between atoms to an appropriate range. We have found that, considering the interaction up to the second nearest neighbors (NNs) and the correlation up to the fifth NNs [see Fig. 1(a)] will produce satisfactory results. The input to each ANN in Fig. 1(b) is determined by the coordinates of the i th atom and its eight neighboring atoms, which are encircled in Fig. 1(a) where two situations (Sn in the center and Te in the center) are indicated. For a 2D structure, the cutoff radius contains a total of 9 atoms up to the second NNs with each atom moving in two directions, resulting in a vector with 18 elements as the input.

To properly and adequately represent the local chemical environment around an atom, we use atom-centered symmetry functions G as descriptors, which are a series of functions of atom positions [38]. As indicated in Ref. [38], the number of symmetry functions describing a given structure should be greater than the degrees of freedom of the described system so that all information is fully recorded. In Fig. 1(b), the column named “atomic NNs” contains identical ANNs that take the chemical environment of an atom, which are encoded in G as shown by the column named “symmetry functions,” as inputs and outputs the energy E_i indicated by the column named “atomic energies.” Finally, E_i is summed to give the total energy E .

The core components of the whole structure are the ANNs as shown schematically in Fig. 1(c), which can be built with

much freedom. For instance, we can use a simple neural network with back-propagation [59,60] or some deep neural networks [61]. Here, given the relatively simple chemical environments, we constructed an ANN with three hidden layers, each layer containing 40 nodes. In addition, since there are two types of atoms (Sn and Te) in the system, two separate ANNs, which have the same structure but different weights inside its nodes, were established to calculate the two types of energies E_{Sn} and E_{Te} imposing on Sn and Te atoms, respectively.

We use the supervised learning as implemented in PYTORCH [62] to train the ANNs, where the DFT calculations are employed to obtain the training data set. For 2D SnTe, it has the lattice constant $a_0 = 6.165$ Å and a vacuum layer of 20 Å along the z direction as shown in Fig. 1(a). *Ab initio* molecular dynamics (MD) simulations with GPAW [63] are performed to simulate a 2×2 system (16 atoms) from 500 to 1 K to generate training samples. Additional configurations with random displacements of atoms are also used to train the model so that it can cope with more complex situations. In these calculations, GPAW uses plane waves with a cutoff energy of 900 eV, different Brillouin-zone sampling grid (for 2D SnTe, for 2×2 unit cell with a $4 \times 4 \times 1k$ mesh) [64] and different exchange-correlation functionals respectively (e.g., the Perdew-Burke-Ernzerhof (PBE) [65]), depending on the system. Because the monovalent characteristics of monochalcogenides (see Sec. IV B), we have also tested and used more advanced corrections, including DFT-D3 [66,67] and strongly constrained and appropriately normed semilocal density functional (SCAN) [68,69], to obtain the training samples (see the Appendix).

In order to train the ANNs, a sufficiently large number of configurations need to be calculated, and the number depends on the actual material. For instance, 2D SnTe used 4000 samples, of which 3900 were used to optimize the ANN, and 100 were used as a preliminary test of the predictive ability of the ANNs. For each of the configurations, its energy is calculated and the energy of the original configuration E_0 , where none of the atoms are displaced, is used as the energy reference. Figure 2 compares the values predicted by the ANNs and calculated by GPAW, which shows a good agreement where the maximum difference is within 8 meV per atom.

The use of the ANNs depends on its ability to reproduce the PES with the accuracy of DFT. We note that the energies of the chosen configurations used in Fig. 2 are all higher than the reference energy E_0 where no atom moves. It has been pointed that while achieving good prediction for random samples is relatively easy, the energies of the ground state or low-energy states are more difficult to accurately predict [70]. In order to assess the PES generated with the ANNs, we generated special configurations where all the Sn or Ge atoms are synchronously displaced along a particular direction to a distance d as shown in Fig. 3(a). Since the configurations are independent of the training data set, they constitute further independent tests to the accuracy of the trained ANNs. In the following sections, we have verified that the prediction by the ANNs for the low-in-energy configurations agrees well with GPAW [e.g., see Figs. 3(c) and 3(d)].

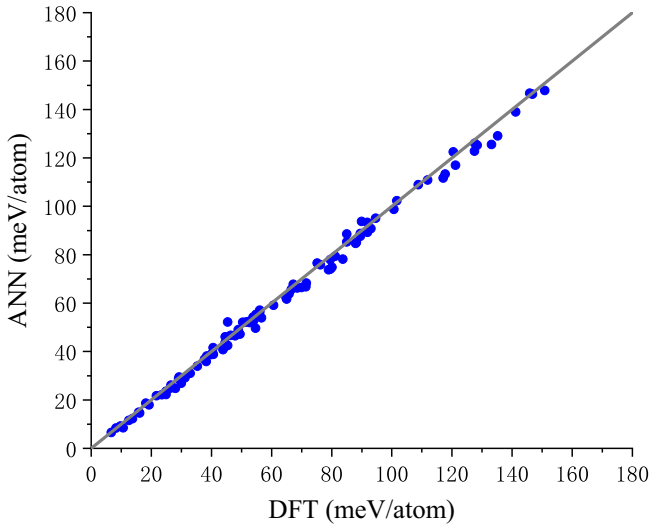


FIG. 2. Comparison of the energy predicted by the trained ANN and generated by DFT (using GPAW) where for each data point, its x -axis value is from GPAW while its y -axis value is the ANN prediction.

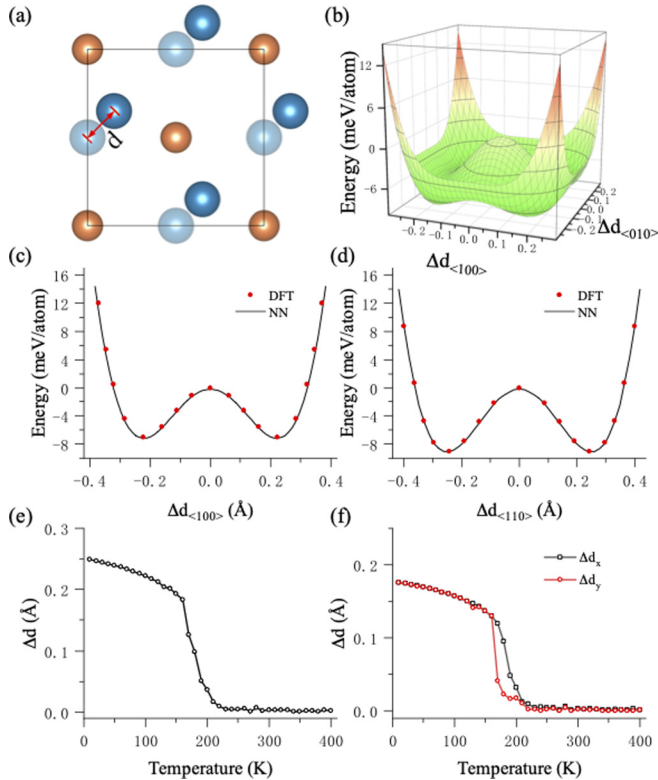


FIG. 3. (a) All the Sn atoms in a unit cell move synchronously to a distance d . (b) The PES is generated by the ANN, showing multiple energy minima. (c), (d) Compare the values predicted by the ANN along the $\langle 100 \rangle$ and the $\langle 110 \rangle$ directions to the GPAW results. (e) Average relative displacement Δd between Sn and Te atoms versus temperature [see Eq. (1)]. (f) Averaged displacement components Δd_x and Δd_y versus temperature.

III. RESULTS

With the trained ANNs acting as the energy function, the phase transition of SnTe can be investigated now. We first check how accurate the ANNs are, and then use it in MC simulations to find the atom displacements with respect to temperature, which discloses information regarding the phase transition of SnTe.

A. 2D monolayer SnTe

For the 2D SnTe, we use the DFT-D3 to compute the training samples, the applicability of which is discussed in the Appendix, Sec. A 2. In order to check the PES generated with the ANNs, we generated special configurations where all the Sn atoms are displaced as shown in Fig. 3(a). Figure 3(b) shows the predicted energies of these configurations and the generated PES. It is interesting to see that the ANNs have generated a smooth PES with multiple local-energy minima. To quantitatively check the predictions of the ANNs, we have sampled along the $\langle 100 \rangle$ and the $\langle 110 \rangle$ directions and compared the energy predicted by the ANN to those calculated with GPAW results in Figs. 3(c) and 3(d). For the region of interest ($d \leq 0.4$ Å), the accuracy provided by the ANN is very high. Moreover, the ANN has successfully reproduced the double-well potential that is a critical indication of possible crystalline phase transitions [71]. We note that the soft mode, which is critical to understand the behavior of a ferroelectric, can be correctly predicted with the ANN since its (imaginary) frequency is associated with the curvature of double-well potential curve at $\Delta d = 0$. This is remarkable when one recalls that the ANN has a universal internal structure and is trained using configurations with essentially random atom displacements.

Having constructed and trained the ANNs, we now employ them in MC simulations as the energy calculator. While Figs. 3(c) and 3(d) show that the ANN can fit the PES very well along the $\langle 100 \rangle$ and the $\langle 110 \rangle$ directions, one needs to keep in mind that the power of the ANNs lies in their ability to predict the energy of *any* configuration, and Figs. 3(b)–3(d) just show some special cases. Such power enables it to be used for MC simulation. In MC simulations, the ANNs are used to predict how much energy change arises when an atom is displaced. To find the phase-transition temperature, we set up a 2D 12×12 supercell (576 atoms), and gradually cool down the system from 400 to 10 K with a step of 10 K. At each temperature, we sweep the system 80 000 times, in each sweep all atoms are displaced randomly and such moves are accepted or rejected depending on the incurred change of energy. We have also tried different number of steps and found that using 80 000 steps for each temperature already ensures the convergence.

Figure 3(e) shows the temperature evolution of the average relative displacement Δd , which is determined by the displacements of Sn and Te within the same unit cell, i.e.,

$$\Delta d = \frac{1}{N} \sum_{i=1}^N \langle |d_{\text{Te}} - d_{\text{Sn}}| \rangle, \quad (1)$$

where $\langle \dots \rangle$ indicates the supercell average, and N is the number of MC sweeps used for the final average. As we can

see from Fig. 3(e), a phase transition occurs at around 200 K, where the average displacement starts to increase with the decreasing temperature, gradually reaching a value around 0.244 Å. Figure 3(f) shows that the averaged displacement component Δd_x and Δd_y separately, indicating that the relative displacements are along the $\langle 110 \rangle$ direction. These results are consistent with the energy minimum at $\Delta d = 0.244$ Å as shown in Fig. 3(d). Interestingly, Fig. 3(f) also indicates the existence of a second phase in a narrow temperature range (~ 160 – 210 K), which will be discussed in Sec. IV B.

B. Bulk SnTe

For bulk SnTe, there was considerable discussion about the best way to perform the DFT calculation as we discuss in Sec. IV B. Practically, we follow the work of Ye *et al.* [31] and choose to use the SCAN exchange-correlation functional which has been shown to be capable of predicting accurate properties for diversely bonded systems [72]. We find that a $6 \times 6 \times 6$ k mesh can achieve satisfactory results. For instance, the relaxed cubic lattice constant was found to be 6.34 Å, agreeing with experiments [22]. Further details will be shown in the Appendix, Sec. A 2. Using a $1 \times 1 \times 2$ supercell, 6000 different configurations are generated as the training set by randomly displacing the atoms. The sample energies were calculated by Vienna *ab initio* simulation package (VASP) [73] with an cutoff energy of 900 eV, a $6 \times 6 \times 3$ Brillouin-zone sampling grid, and the SCAN exchange-correlation functional.

After obtaining the samples, the ANN-based approach can be used to treat bulk SnTe without additional setup. Using the same ANN structure as in 2D, we retrained the neural network using three-dimensional (3D) samples, and show the test against the potential well in Fig. 4(a), where the red, green, and blue curves represent the potential well corresponding to shifting the Sn atoms in the supercell along the $\langle 100 \rangle$, $\langle 110 \rangle$, and $\langle 111 \rangle$ directions. As the energy difference between different displacement directions is as small as 0.3 meV (close to the energy minimum), the bulk SnTe poses a big challenge to ANNs' accuracy. Figure 4(a) shows that the ANNs can indeed distinguish such small energy differences.

After obtaining the neural network potential, it is used as the energy calculator for the subsequent MC simulation with a $12 \times 12 \times 12$ supercell (13 824 atoms), where the temperature was cooled from 100 to 5 K at a 5 K intervals, and each temperature was scanned 80 000 times. Figure 4(b) show that the phase-transition temperature of 3D SnTe is around 35 K, which is lower than the reported phase-transition temperature of 98 K [22]. The reason and possible solution will be discussed in Sec. IV D.

C. GeTe

The approach we have used in Secs. III A and III B can be easily transferred to another system. To demonstrate this advantage, we first apply it to 2D monolayer GeTe. For this purpose, the DFT calculation with GPAW adopts a cutoff energy of 900 eV, a $4 \times 4 \times 1$ k mesh, and PBE. The system has the in-plane lattice constant $a_0 = 5.74$ Å, and for the first-principle calculations we use a vacuum layer of 20 Å along

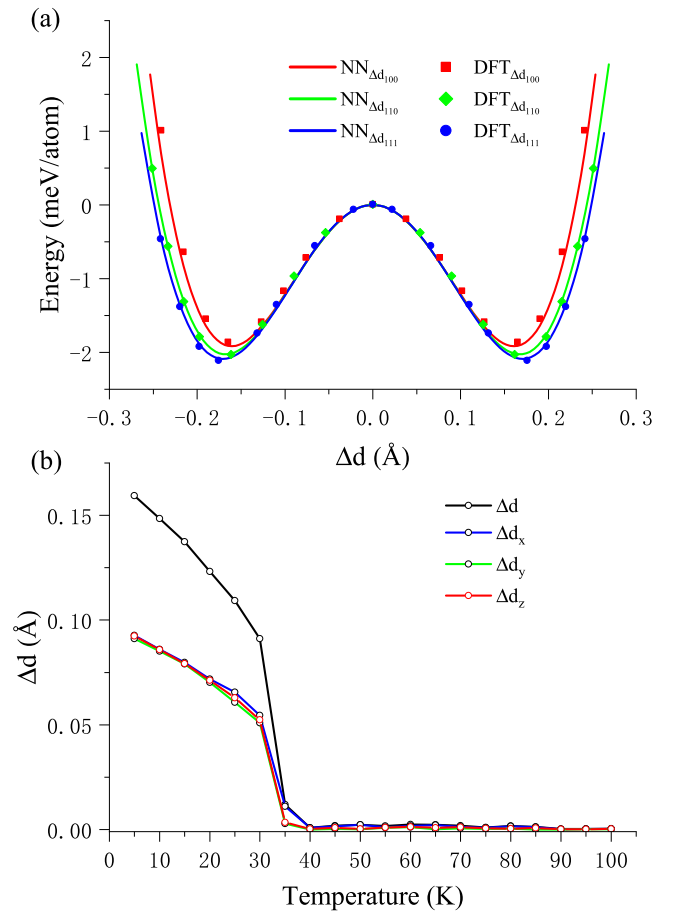


FIG. 4. (a) The potential well of bulk SnTe with atom displaced in different high-symmetry directions. (b) Average relative displacement (Δd) and displacement component (Δd_x , Δd_y , and Δd_z) between Sn and Te atoms versus temperature.

the z direction similar to SnTe. As atoms only move inside the plane, a total of 8000 random samples are obtained and used for training the ANNs.

With the same procedure as that for SnTe, the trained ANNs for GeTe show excellent agreement with DFT results [Figs. 5(a) and 5(b)] and that the double-well potential is deeper than that of 2D SnTe. We again conducted MC simulations, also using a 2D 12×12 supercell (576 atoms), and gradually cooled down the system from 600 to 20 K with a step of 20 K with 80 000 MC sweeps at each temperature. It was found that GeTe, similar to SnTe, also has a structural phase transition with $T_C \sim 300$ K as shown in Figs. 5(c) and 5(d). The phase-transition temperature is higher than that of SnTe, consistent with the deeper potential well of GeTe.

In addition, we also apply the same procedure to bulk GeTe, which was first investigated using the effective Hamiltonian approach [9]. Using a $1 \times 1 \times 2$ supercell of bulk GeTe, 10 000 different structures are generated as the training set. The samples are calculated with GPAW using a $8 \times 8 \times 4$ k mesh and PBE. As shown in Fig. 6(a), the ANN is adequate to predict the DFT energies. MC simulation results in Fig. 6(b) show that the phase-transition temperature of bulk GeTe is ~ 250 K, which is lower than the experimentally

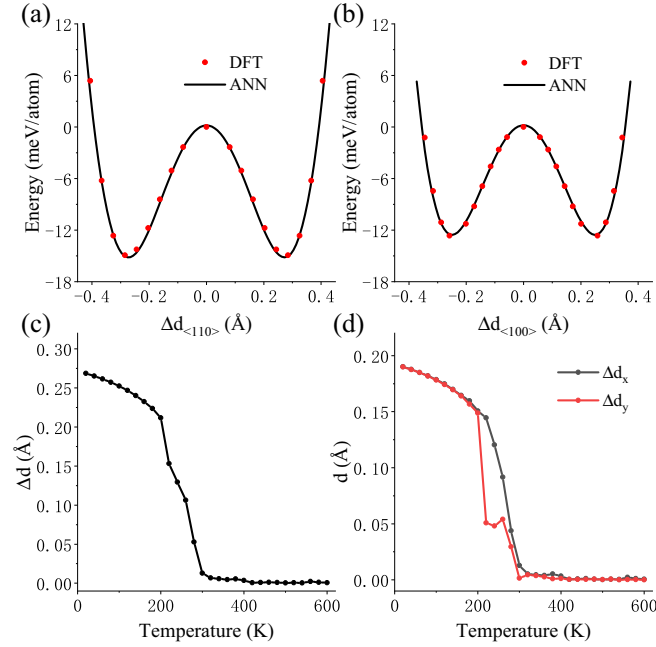


FIG. 5. (a), (b) The ANN's prediction results for 2D GeTe. (c) Average relative displacement Δd between Ge and Te atoms versus temperature. (d) Averaged displacement component Δd_x and Δd_y versus temperature.

reported 600–700 K. In addition, Fig. 5(d) also indicates that there exists a second phase (atoms shifting along the $\langle 100 \rangle$ direction) in a narrow temperature range between 220 and 250 K, similar to 2D SnTe and GeTe, but does not appear in bulk SnTe. The discrepancy in the transition temperature and the second phase will be discussed in Sec. IV D.

IV. DISCUSSION

In this work, we have reached several goals, successfully applying the ML approach to the investigation of ferroelectric phase transitions, which was mostly the arena of effective Hamiltonians. Using this approach, we have constructed 2D and 3D potential surface functions for SnTe and GeTe with ANNs, achieving an accuracy that is capable even for 3D SnTe where the depth of the potential well can be as small as 2 meV/atom. While such accuracy is necessary [Fig. 4(a) shows how challenging the requirement could be], it has not been achieved for the crystalline ferroelectric phase transition. In addition, as SnTe has attracted much interest in the past few years [31–33], this work with the ANNs provides an alternative method to explore its ferroelectric phase transitions.

In this section, we discuss some important aspects of this ML approach that can address crystalline phase transitions, including the unique features of SnTe and GeTe and the advantages of the ML techniques. With its unique electronic properties, SnTe has posed particular challenges that can be resolved by ML which can be adapted to more systems with minimal modification. We will also address the limitations of the current approach that need further investigation in the future.

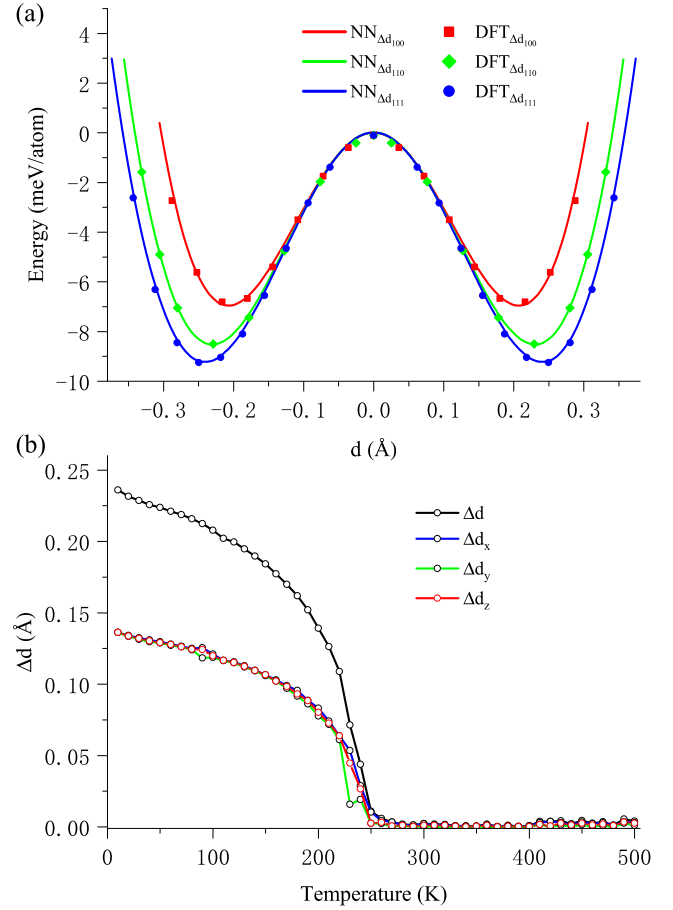


FIG. 6. (a) The potential well of bulk GeTe with atom displaced in different high-symmetry directions. (b) Average relative displacement (Δd) and displacement component (Δd_x , Δd_y , and Δd_z) between Ge and Te atoms versus temperature.

A. Machine learning

This work has heavily depended on ML techniques to investigate the structural phase transition of monochalcogenides using SnTe and GeTe as examples. While the overall procedure is based on supervised learning, each stage (e.g., the training of ANNs and the descriptors) warrants its individual investigation to ensure the accuracy and efficiency of the training and later simulations. Practically, we use the ANN as a unified surrogate energy function, the success of which crucially depends on two factors: (i) The power of ANNs to decompose the total energy of a given configuration into the contribution from each atom [55,56] and (ii) the construction of the symmetric function G as good descriptors as they exhibit the invariance of translation, rotation, and rearrangement of the same elements, as well as the smooth decay of the contribution for atoms away from the center [74].

In practice, ML requires building a sample set to train the ANNs. For 2D SnTe and GeTe, about 4000 and 8000 samples are generated respectively, each containing a configuration of 16 atoms (2×2 supercell) and the corresponding energy. There are many investigations on how to generate the samples in an efficient way [75,76]. The actual number of samples needed could be even smaller if the samples are properly

chosen, e.g., focusing on the low-energy configurations. However, we have found that using MD to generate all the samples may not have the best effects. In this work we used the combination of MD simulation and random sampling to obtain the training set. In the random sampling part, in order to cover a large sample space, we have fixed the positions of some atoms and randomly shifted the remaining atoms.

Regarding sample generation, it is also worth noting that 3D SnTe and GeTe require about 6000 and 10 000 samples, respectively, with their energies calculated, each sample employing a configuration of 16 atoms ($1 \times 1 \times 2$ supercell). While the number of samples seems large, the computational cost of the samples is insignificant comparing to that of direct DFT computations needed to find the phase transitions. For instance, for the $12 \times 12 \times 12$ (13 824 atoms) supercell, in the MC simulations for each temperature, we use 80 000 sweeps, in each sweep every atom is displaced, which means that a total of 1.1×10^9 DFT calculations for a supercell with 13 824 atoms need to be done, just for one temperature among the several tens needed for the phase transition. This magnitude of computation required by direct DFT is clearly inaccessible, explaining why the ML approach is more advantageous.

In the training of the ANNs, the loss function is the mean-square error (MSE), comparing the predicted values and true values (provided by DFT calculations). For random samples, the energy difference could be as large as a few eV. However, the energy difference that is most relevant to the structural phase transition could be as small as 1 meV. Given the nature of crystalline phase transitions, this work has focused on the low-energy regions where important energy difference can be less than 1 meV, which naturally requires a high numerical accuracy from the ML approach. It has been shown that this requirement can be fulfilled by the ANNs. In addition, ML provides a model-free method, which can directly predict the energy of the system using the positions of the atoms. In essence, it replaces the energy (or atomic force) calculation with electrons (as in DFT) to achieve a much higher efficiency. The accuracy and efficiency, along with its transferability to various systems, constitute the prevailing advantage of the ML approach proposed here.

On the other hand, although 2D and 3D use the same network structure (including the number of inputs), their meanings are completely different although the descriptors compress the inputs into the same dimensions. Moreover, DFT itself has some difficulties treating 2D and 3D systems in the same way (as we discuss in Sec. IV B). These facts make the reuse of the trained ANNs for systems with different dimensions quite difficult.

In the MC simulation, it turns out that the calculation of the atom-centered symmetry functions is quite time consuming, which needs to be optimized. In fact, the computational cost increases sharply with the truncation radius R_c , as well as the number of three-body descriptors that consider the angles among three atoms [74]. We note that the proper selection of hyperparameters and the optimization of descriptors have been an important topic of investigation [77–79].

Finally, let us consider the universal nature of this ML approach. We note the approach itself (see Fig. 1), not the ANNs, is transferable in the sense that the whole setup [see Figs. 1(b) and 1(c)] can be used for different types of materi-

als, similar to what DFT can do, without having to build a new model for each different material, benefiting from the fact that the ANN can act as a general mathematical approximation. The ML approach contrasts with the effective Hamiltonian method, which may need very different models for different systems, in that it does not need much customization, but only needs appropriate data to train the ANNs. From this point of view, the ML approach is indeed more “universal.”

B. SnTe and GeTe

Many monochalcogenides, including SnTe and GeTe, have unique bonding mechanisms that are termed metavalent bonding (MVB) by researchers [35,80], which is believed to be different from metallic, ionic, or covalent bonding. The electronic structure instability plays an important role in MVB, reflecting a competition between the electron delocalization (as in metals) and the electron localization (as in ionic crystals or covalent compounds) [35]. As a comparison, ferroelectric oxides have the structural instability that usually arises from electric dipole-dipole interactions, but lack the electronic instability as their bonds are ionic [35]. This difference has profound implication for the construction of an effective Hamiltonian as we discuss in Sec. IV C.

When bulk SnTe or GeTe is cooled from high temperature, the Peierls-type rhombic distortion occurs, resulting in the $R3m$ symmetry, forming covalent bonds [35,81,82] and reaching a stable state. The potential well for GeTe is much deeper than SnTe (see Fig. 6), consistent with the fact that GeTe has stronger electron sharing (more covalent) than SnTe [cf. Fig. 8(b) of Ref. [35]]. The shallow potential well of SnTe is also consistent with the smaller displacement of Sn to its equilibrium position (0.17 Å).

The 2D monolayer SnTe is special in that reducing the dimensions will impede the electron delocalization, altering the electronic structure that affects the bond type, leading to large property changes. In a recent paper, Kooi and Wuttig [35] pointed out that this example probably shows the destruction of MVB to a mixed ionic-covalent bonding. The change to ionic or the covalent bonding will increase the interaction between atoms. For instance, ionic or covalent bonding may introduce stronger interaction that increases the ion displacements [from 0.17 Å in Fig. 4(a) to 0.244 Å in Fig. 3(d)], which may explain the increased phase-transition temperature. Such bond-type change is an interesting and unique phenomenon with SnTe, which is evidenced by the similarity (in terms of potential well depth and the magnitude of displacement) of the double-well potential of 2D SnTe (Fig. 3) and GeTe (Fig. 5) while their bulk differs substantially (Fig. 6). We note that GeTe conforms to the prediction of the finite-size scaling theory for ferroelectrics, that is when the ferroelectric thin-film material becomes thinner, its phase-transition temperature will become lower than that of the bulk [83,84]. The reason could be that bulk GeTe has strong covalent bond (unlike bulk SnTe where the bonding is weaker) which made the bulk phase transition already high [47].

The MVB theory effectively explains many unique phenomena associated with monochalcogenides, and SnTe in particular, such as the increase of phase-transition temperature as it gets thinner. However, the MVB also makes the construc-

tion of an effective Hamiltonian difficult, which we discuss next in Sec. IV C. In addition, to address the unique properties of SnTe, our DFT calculations have employed different schemes to deal with SnTe, especially bulk SnTe [31,47]. For instance, we have found that relaxation of bulk using SCAN will obtain a lattice constant (6.34 Å) that is most close to experimental value [22,31] than PBE (6.40 Å) or PBE with van der Waals (vdW) (6.345 Å). Our own tests in the Appendix, Sec. A 2, show that SCAN is a good choice in terms of both accuracy and efficiency.

Finally, our results of the crystalline phase transitions have indicated the existence of a second phase. As shown in Figs. 3(f) and 5(d), 2D monolayer SnTe and GeTe have another structural phase in a narrow temperature range, where the average displacement of atoms is along the $\langle 100 \rangle$ direction. To understand its origin, we checked snapshots of atom displacements of 2D SnTe during the simulations, which indicate that, in this temperature range, atoms shift along different $\langle 110 \rangle$ directions, resulting in a second ferroelectric phase with average displacement along the $\langle 100 \rangle$ direction. This is similar to what happens in BaTiO₃, where while most dipoles are in the $\langle 111 \rangle$ direction, at certain temperatures, the averaged polarization could be along $\langle 100 \rangle$ or $\langle 110 \rangle$ directions [85]. This phenomenon is likely caused by the competition between energy minima as shown in Fig. 3(b). The energy difference between the saddle point and the energy minima seems to be the key factor that determines if the other phase exists. For instance, the signature of the extra phase appears [see Figs. 3(f), 5(d), and 6(b)] when the value is large (1.92 meV for 2D SnTe, 2.58 meV for 2D GeTe, and 2.45 meV for bulk GeTe), but does not exist when the value is small (0.25 meV for bulk SnTe) as seen in Fig. 4(b).

C. Effective Hamiltonian

An effective Hamiltonian provides an economic description of a given system and a surrogate energy function with the abstracted atom displacements (e.g., local mode) as the input. The effective Hamiltonian approach has been most successful in dealing with ferroelectric perovskites [9], which possesses structural instability often associated with the dipole-dipole interaction as perovskites are ionic oxides. In contrast, MVB materials have structural instabilities associated with electronic instabilities [35]. This feature has profound implication for the construction of an effective Hamiltonian for monochalcogenides since it does not satisfy the necessary condition of a local mode.

For ferroelectric perovskites, the long-range dipole-dipole interaction is the driving force causing the structural instability. The local interactions are approximately harmonic with regard to the displacement of atoms (e.g., the value κ_2 for BaTiO₃ is positive in Table II of Ref. [9]), therefore giving rise to the local mode. In contrast, the instabilities of monochalcogenides come from the electronic structure which is a more localized feature. As a consequence, the relevant phonon mode is highly anharmonic with large Grüneisen parameter [80,86], making it hard to define a proper local mode that is valid for the displacement that spans the whole range of interest. This issue could be one reason that the effective Hamiltonian approach predicts a phase-transition temperature

that is away from experimental results. Moreover, according to the MVB theory, the bonding type could change with the displacement and the dimension, which is demonstrated by the unusually high phase-transition temperature of 2D SnTe. One way to address the problem is to patch the effective Hamiltonian [31,32], where the idea was to introduce new terms to compensate for the change in the order parameter when the dimension changes from 3D to 2D as the SnTe becomes thinner [31]. While the prediction of the transition temperature has improved (e.g., 147 K versus experimental 98 K for bulk SnTe), the patch is hard to conceive, in addition to the fact that it will cause fragmentation of the effective Hamiltonian approach in the long run, therefore losing its elegance in simplicity and “ease of use.”

Given the above discussion, it may be concluded that the difficulty with monochalcogenides hinges on the facts that (i) the proper local mode is hard to determine [34] and (ii) the local mode (order parameter) can change substantially as the dimension of the system is reduced. The ML approach, on the other hand, does not need to determine the local mode in the first place and it does not care if the dimension is reduced or not, therefore, no patch is needed when the thickness changes. It is important to note that our ANN-based approach has treated the 2D and bulk SnTe in a unified fashion, requiring no separate determination of the local mode or patching the effective Hamiltonian [31,32]. In addition, the method does not have to start all over again from the very beginning when a similar but different system is treated.

With the machine-learning-based approach, the determination of the local mode, the potential modification of the effective Hamiltonian, and the calculation of the coefficients can be avoided. The only requirements are the generation of new training data sets. Therefore, the current approach can be extended conveniently to various systems.

D. Limitation

In our simulations, we have assumed the lattices of SnTe or GeTe are fixed in the clamped state. For 2D systems, this can be realized using epitaxial strain, whereas for 3D system it may not be entirely realistic. From Landau-Devoshire theory, it is known that the phase-transition temperature is indeed affected by strain [87]. We note that the inclusion of strain, ideally embedded in the ANNs (see Fig. 1), is not a trivial job and will be done in a future work. Especially when such precision needs to be maintained, the addition of strain will greatly increase the degree of freedom of the samples. The phase transitions of bulk SnTe and GeTe we predicted are therefore lower than known experimental values.

In order to address this issue and correctly predict the phase-transition temperature, we performed additional calculations to estimate the effects of strain with DFT. In Fig. 7, the “cubic phase” indicates the lattice structure obtained by relaxing the unit cell while keeping all the atoms in their ideal positions (therefore a cubic phase); the “trigonal phase” indicates the relaxed lattice structure by first moving the Sn/Te (or Ge/Te) atoms along one $\langle 111 \rangle$ direction and allowing the positions of the atoms to be relaxed as well. Figures 7(a) and 7(b) show the potential well of bulk SnTe and GeTe in their cubic and trigonal phases, respectively, where the atom displace-

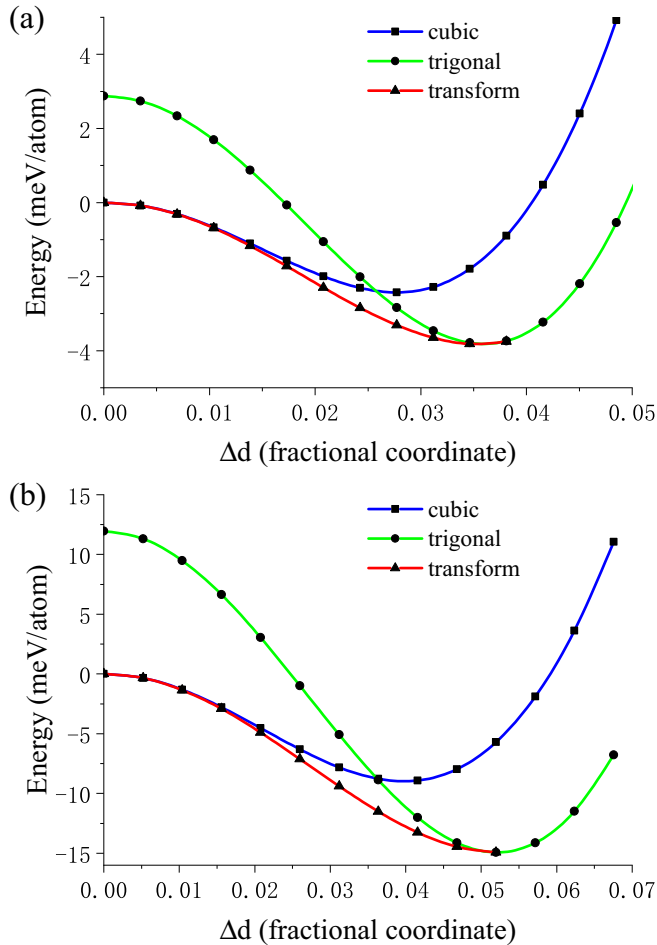


FIG. 7. Energy versus atom displacement for the cubic (blue curve) and trigonal (green curve) phases of SnTe (a) and GeTe (b). The lattice structure for each remains the same during the shift of atoms. The red curve shows the obtained energy when we fix the fractional coordinates of the atoms and relax the unit cell, which demonstrates an interpolating path with minimal energy between the cubic and the trigonal phases.

ment is along the $\langle 111 \rangle$ direction. The two figures compare the energies of the cubic and trigonal phases when the atom positions are the same (in fractional coordinates). The results indicate that, when strain is included, the actual PES will most likely become a smooth surface bridging the cubic and trigonal phase as shown by the red line with triangles in Fig. 7, which show that, for bulk SnTe, the potential well is -3.82 meV, giving rise to a T_C of 60 K by considering the T_C without strain obtained with MC [Fig. 4(b)], while for bulk GeTe, $T_C = 420$ K. As a matter of fact, if we only consider using the potential well of the trigonal phase (the green curve with black circles), the phase-transition temperatures for bulk SnTe and GeTe are 96 and 753 K, respectively, close to experimental values. Without doubt, MC simulations will provide consistent phase-transition temperatures when the strain as new degrees of freedom are included.

We finally note that, by including the influence from atoms up to the fifth (for the 2D case) or the seventh (for the bulk case) NNs, the ML approach can take into account of the

local electronic structure change due to ion shift [88]. In addition, the long-range Coulomb interaction may also need to be included via the Ewald method [89]. However, the current results indicate that this part of the energy will likely be much smaller than that of typical perovskites such as BaTiO_3 or PbTiO_3 .

V. CONCLUSION

In this work, we have developed a ML approach to investigate the crystalline phase transitions of SnTe and GeTe. The constructed energy functions of SnTe and GeTe using ANN are employed in MC simulations to obtain the phase transitions, where the energy difference between different ferroelectric phases can be smaller than 1 meV/atom. Unlike other approaches, no specific model or formula, which often requires *a priori* knowledge or a good understanding of the given system, is necessary to approximate the PES. This approach therefore also makes it unnecessary to determine the coefficients appearing in the model or the formula, essentially resulting in a model-free method while achieving the accuracy of *ab initio* calculations. This feature is advantageous than other more complex methods when the real difficulty of obtaining the phase-transition temperature lies in the construction of a sufficiently accurate model PES. In addition, since the ANNs can work as a universal mathematical structure, it can approximate various systems according to the *universal approximation theorem* (see, e.g., Chap. 4 of Ref. [60]) within the range covered by the sampling space. Such virtues, along with its accuracy and efficiency, make the ML approach very flexible to investigate crystalline phase transitions of various systems.

ACKNOWLEDGMENTS

We are grateful to Professor W. Zhang for fruitful discussion. This work is financially supported by the National Natural Science Foundation of China (NSFC), Grants No. 11974268 and No. 12111530061. X.C. thanks the financial support from Academy of Finland Project No. 308647. D.W. thanks the support from the Guangxi Key Laboratory of Optical and Electronic Materials and Devices (20KF-13), L.L. thanks the support from the Natural Science Foundation of Guangxi (Grants No. FA198015, No. GA245006, No. BA245069, No. FA198015, No. BA297029, and No. 20AA-7). D.F. thanks the support from the Natural Science Foundation of Shaanxi Province (Grant No. 2019JQ-240). G.Z. thanks the support from NSFC (Grant No. 21503057) and the Fundamental Research Funds for the Central Universities (Grant No. HIT.NSRIF.2017032). X.C. and D.W. also thank the support from CSC (IT Center for Science, Finland), Project No. 2001447, for providing computation resources.

APPENDIX: CONVERGENCE

Training sample generation is one important step in the process of building the necessary ANNs to predict the phase-transition temperature of monochalcogenides. In order to obtain accurate results, we have various tests to validate the

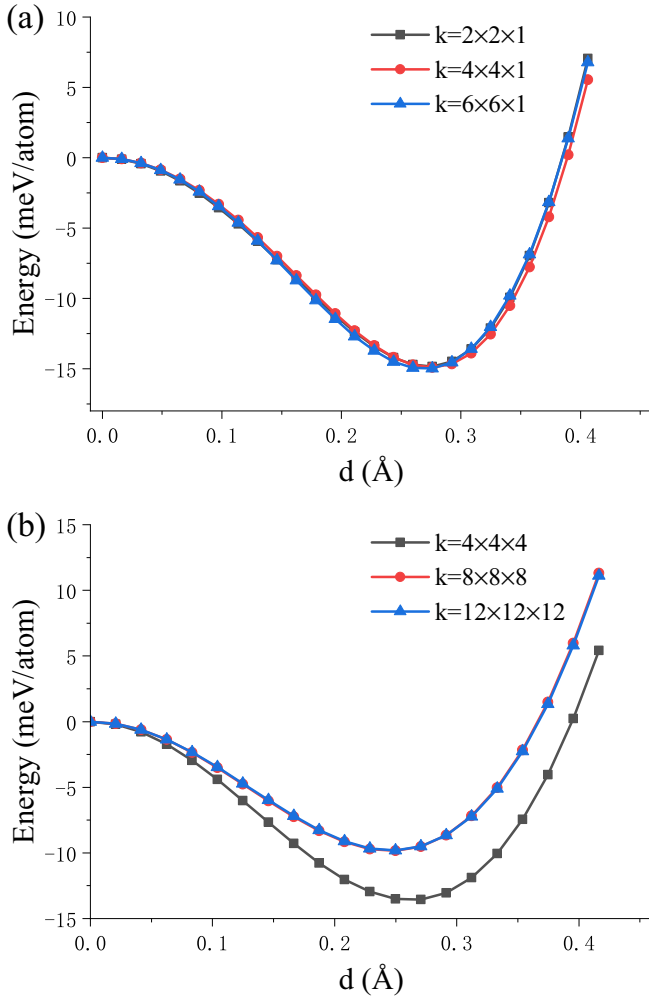


FIG. 8. The energy-versus-displacement curves are obtained with different k -point setup for (a) 2D monolayer GeTe of a 2×2 unit cell and (b) Bulk GeTe of a $1 \times 1 \times 1$ unit cell.

exchange-correlation functional, as well as the setups used in first-principles calculations. Since the phase-transition temperature is closely related to the energy landscape with respect to atom displacement, we choose the convergence of the double-well potential of Sn(Ge) atoms moving in different high-symmetry directions ($\langle 111 \rangle$ direction for 3D and $\langle 110 \rangle$ direction for 2D) as key tests.

1. GeTe

For GeTe, we find that the first-principles calculation results are insensitive to the choice of the exchange-correlation functionals. The use of PBE with a moderately dense k mesh is adequate as shown in Fig. 8, which checks the convergence using the energy-versus-displacement curve. For the 2D monolayer GeTe, we used a 2×2 unit cell (16 atoms and a 20-Å vacuum layer) and found in Fig. 8(a) that an adequate convergence is achieved with a $2 \times 2 \times 1$ k mesh. However, also considering the double potential in other directions (e.g., the $\langle 100 \rangle$ direction) and the accuracy in the lattice constant, the $4 \times 4 \times 1$ k mesh is used in the sample generation. For bulk GeTe, we used a $1 \times 1 \times 1$ unit cell (8 atoms) to check

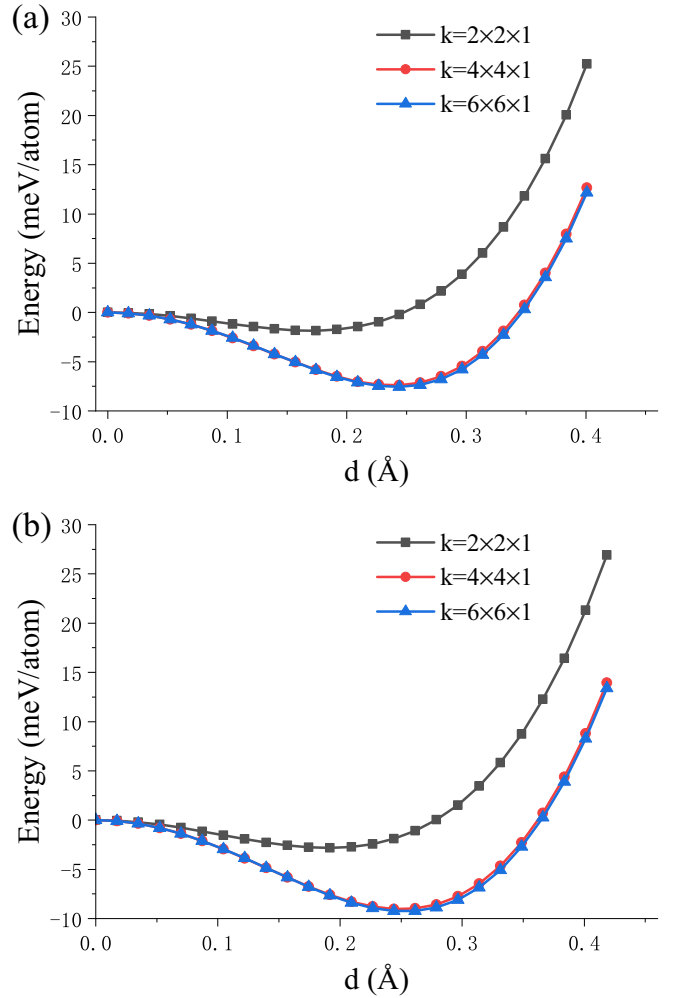


FIG. 9. The energy-versus-displacement curves for 2D SnTe are obtained with a 2×2 unit cell and different k -point setups using (a) PBE and (b) DFT-D3

the convergence, which is achieved with an $8 \times 8 \times 8$ k mesh as evidenced in Fig. 8(b).

2. SnTe

We found that the sample generation for SnTe is difficult, which is likely related to its unique electronic properties as explained in Sec. IV B. For 2D monolayer SnTe, we used a 2×2 unit cell and tested PBE and DFT-D3 with different k -mesh setups. Figure 9(a) shows that the potential well deepens as the k mesh becomes denser, achieving a convergence at $4 \times 4 \times 1$. A similar conclusion can be made using the DFT-D3 method as shown in Fig. 9(b).

As discussed in Sec. IV B, the metavalent status makes first-principles calculations of bulk SnTe tricky. Therefore, we tested a few different approaches, including PBE, DFT-D3, and SCAN, where GPAW was used for the first two calculations while VASP was used for the SCAN calculation. It has been shown that SCAN can perform better than other semilocal density functionals [90] to accurately predict the geometries and energies of diversely bonded systems (including covalent, metallic, ionic, hydrogen, and van der Waals bonds)

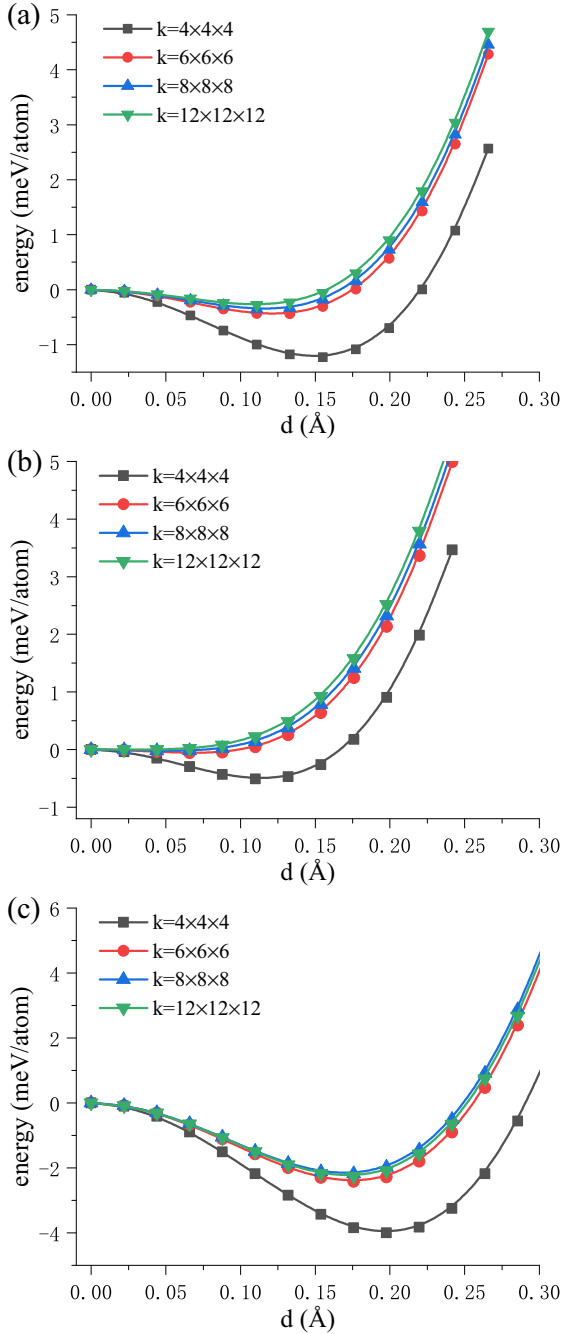


FIG. 10. The energy-versus-displacement curves of bulk SnTe using different k -point setups and different exchange-correlation functionals, which are (a) PBE, (b) DFT-D3, and (c) SCAN.

[91], making it likely also suitable for SnTe with unique electronic properties.

Figure 10(a) shows that, with PBE, the results do not converge even with a very dense k mesh ($12 \times 12 \times 12$) where the potential well continues to become shallower, conflicting with the known experimental fact that bulk SnTe does undergo crystalline phase transitions with decreasing temperature at ~ 100 K. To address this problem, we also tried using DFT-D3, but with no obvious improvement [see Fig. 10(b)], which reflects the difficulty in treating the vdW interactions in MVB systems. Finally, using SCAN we find that (i) the relaxed

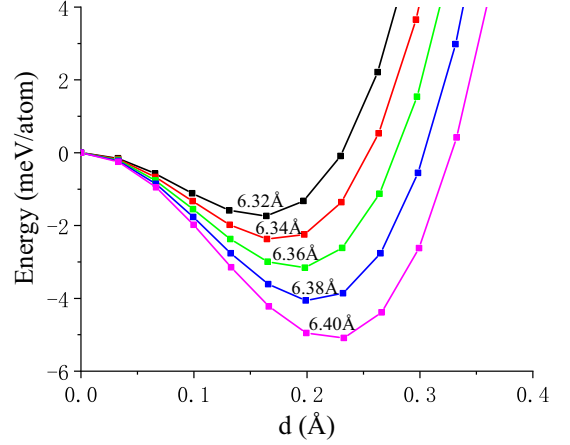


FIG. 11. Potential well of bulk SnTe at different lattice constants.

lattice constant (6.34 \AA) is most close to the experimental value of 6.32 \AA and (ii) the first-principles calculation results converge at the $8 \times 8 \times 8 k$ mesh where Fig. 10(c) indicates the $6 \times 6 \times 6$ and $8 \times 8 \times 8 k$ meshes produce very similar results, which also encloses the result obtained with the k mesh of $12 \times 12 \times 12$. This allows us to conclude that SCAN can better describe SnTe among the methods considered here. This can be understood by the fact that SCAN represents a major step forward in the nonempirical semilocal approximations [68,90], capturing mid-range vdW interactions to a certain extent, as can be seen from the improved performance [92].

Therefore, in the sample generation we adopted SCAN and the $6 \times 6 \times 6 k$ mesh considering the balance between accuracy and the cost of the calculations. We note that the SCAN results here are consistent with those from Ref. [31]. In addition, we find that for bulk SnTe, the potential well is affected by the lattice constants as shown in Fig. 11, again showing that the phase transition is affected by strain as discussed in Sec. IV D.

3. Discussion

Regarding the accuracy of the DFT calculation, we have used adequate k -points sampling and large plane-wave cutoff. For the sample generation with the $1 \times 1 \times 2$ SnTe supercell, we use the $6 \times 6 \times 3 k$ -point sampling and 900 eV as the cutoff, which are quite large numbers that achieve the desired DFT accuracy. It shall be noted that, after the ANNs are trained, there is no need to run DFT calculations on large supercells (e.g., the $12 \times 12 \times 12$ supercell used in the MC simulations) since following calculations employ the ANNs instead of DFT calculations. Therefore, the large k -points sampling would not be too much a burden.

It is worth noting that the accuracy of the ANNs mentioned in the text should not be confused with the accuracy of DFT calculation. Since the energy difference between ferroelectric phases is very small, as shown in Fig. 4(a), it is important for the ANNs to capture the tiny difference to correctly predict the ferroelectric phase at low temperatures, therefore, the predicted value of the ANN needs to be as close to the DFT result as possible. The accuracy of the ANNs mentioned in the text refers to the extremely small difference between the ANN predictions and the DFT results.

- [1] R. A. Cowley, Structural phase transitions I. Landau theory, *Adv. Phys.* **29**, 1 (1980).
- [2] M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Oxford University Press, Oxford 1977).
- [3] J. F. Scott, Applications of modern ferroelectrics, *Science* **315**, 954 (2007).
- [4] G. S. Pawley, W. Cochran, R. A. Cowley, and G. Dolling, Diatomic Ferroelectrics, *Phys. Rev. Lett.* **17**, 753 (1966).
- [5] R. Fei, W. Kang, and L. Yang, Ferroelectricity and Phase Transitions in Monolayer Group-IV Monochalcogenides, *Phys. Rev. Lett.* **117**, 097601 (2016).
- [6] K. M. Rabe, C. H. Ahn, and J. Triscone, *Physics of Ferroelectrics: A Modern Perspective* (Springer, Berlin, 2007).
- [7] W. Zhong, D. Vanderbilt, and K. Rabe, Phase Transitions in BaTiO₃ from First Principles, *Phys. Rev. Lett.* **73**, 1861 (1994).
- [8] R. King-Smith and D. Vanderbilt, First-principles investigation of ferroelectricity in perovskite compounds, *Phys. Rev. B* **49**, 5828 (1994).
- [9] W. Zhong, D. Vanderbilt, and K. M. Rabe, First-principles theory of ferroelectric phase transitions for perovskites: The case of BaTiO₃, *Phys. Rev. B* **52**, 6301 (1995).
- [10] W. Cochran, Crystal Stability and the Theory of Ferroelectricity, *Phys. Rev. Lett.* **3**, 412 (1959).
- [11] W. Cochran, Soft modes, a personal perspective, *Ferroelectrics* **35**, 3 (1981).
- [12] L. Bellaïche and D. Vanderbilt, Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites, *Phys. Rev. B* **61**, 7877 (2000).
- [13] I. A. Kornev, L. Bellaïche, P.-E. Janolin, B. Dkhil, and E. Suard, Phase Diagram of Pb(Zr, Ti)O₃ Solid Solutions from First Principles, *Phys. Rev. Lett.* **97**, 157601 (2006).
- [14] D. Wang, A. A. Bokov, Z.-G. Ye, J. Hlinka, and L. Bellaïche, Subterahertz dielectric relaxation in lead-free Ba(Zr, Ti)O₃ relaxor ferroelectrics, *Nat. Commun.* **7**, 11014 (2016).
- [15] A. Al-Barakaty, S. Prosandeev, D. Wang, B. Dkhil, and L. Bellaïche, Finite-temperature properties of the relaxor PbMg_{1/3}Nb_{2/3}O₃ from atomistic simulations, *Phys. Rev. B* **91**, 214117 (2015).
- [16] J. Zhang, Y. J. Wang, J. Liu, J. Xu, D. Wang, L. Wang, X. L. Ma, C. L. Jia, and L. Bellaïche, Origin of sawtooth domain walls in ferroelectrics, *Phys. Rev. B* **101**, 060103(R) (2020).
- [17] J. Liu, L. Liu, J. Zhang, L. Jin, D. Wang, J. Wei, Z. G. Ye, and C. L. Jia, Charge effects in donor doped perovskite ferroelectrics, *J. Am. Ceram. Soc.* **103**, 5392 (2020).
- [18] S. Prosandeev, S. Prokhorenko, Y. Nahas, A. Al-Barakaty, L. Bellaïche, P. Gemeiner, D. Wang, A. A. Bokov, Z.-G. Ye, and B. Dkhil, Evidence for Goldstone-like and Higgs-like structural modes in the model PbM_{1/3}Nb_{2/3}O₃ relaxor ferroelectric, *Phys. Rev. B* **102**, 104110 (2020).
- [19] K. P. Kelley, D. E. Yilmaz, L. Collins, Y. Sharma, H. N. Lee, D. Akbarian, A. C. T. van Duin, P. Ganesh, and R. K. Vasudevan, Thickness and strain dependence of piezoelectric coefficient in BaTiO₃ thin films, *Phys. Rev. Materials* **4**, 024407 (2020).
- [20] D. Akbarian, D. E. Yilmaz, Y. Cao, P. Ganesh, I. Dabo, J. Munro, R. V. Ginheved, and A. C. T. van Duin, Understanding the influence of defects and surface chemistry on ferroelectric switching: a ReaxFF investigation of BaTiO₃, *Phys. Chem. Chem. Phys.* **21**, 18240 (2019).
- [21] K. P. Kelley, A. N. Morozovska, E. A. Eliseev, V. Sharma, D. E. Yilmaz, A. C. T. van Duin, P. Ganesh, A. Borisevich, S. Jesse, P. Maksymovych, N. Balke, S. V. Kalinin, and R. K. Vasudevan, Oxygen vacancy injection as a pathway to enhancing electromechanical response in ferroelectrics, *Adv. Mater.* **34**, 2106426 (2022).
- [22] K. Chang, J. Liu, H. Lin, N. Wang, K. Zhao, A. Zhang, F. Jin, Y. Zhong, X. Hu, W. Duan, Q. Zhang, L. Fu, Q. Xue, X. Chen, and S. Ji, Discovery of robust in-plane ferroelectricity in atomic-thick SnTe, *Science* **353**, 274 (2016).
- [23] K. Liu, J. Lu, S. Picozzi, L. Bellaïche, and H. Xiang, Intrinsic Origin of Enhancement of Ferroelectricity in SnTe Ultrathin Films, *Phys. Rev. Lett.* **121**, 027601 (2018).
- [24] L. W. Martin and A. M. Rappe, Thin-film ferroelectric materials and their applications, *Nat. Rev. Mater.* **2**, 16087 (2017).
- [25] M. Dawber, K. M. Rabe, and J. F. Scott, Physics of thin-film ferroelectric oxides, *Rev. Mod. Phys.* **77**, 1083 (2005).
- [26] Z. Li, Y. Guo, F. Zhao, C. Nie, H. Li, J. Shi, X. Liu, J. Jiang, and S. Zuo, Effect of film thickness and evaporation rate on co-evaporated SnSe thin films for photovoltaic applications, *RSC Adv.* **10**, 16749 (2020).
- [27] Z. Li, Y. Guo, F. Zhao, C. Nie, J. Shi, X. Liu, S. Zuo, and J. Jiang, The effect of substrate temperature and Sn/Se mass ratio on the co-evaporated SnSe thin film for photovoltaic application, *Vacuum* **177**, 109343 (2020).
- [28] G. Bruns, P. Merkelbach, C. Schlockermann, M. Salinga, M. Wuttig, T. D. Happ, J. B. Philipp, and M. Kund, Nanosecond switching in GeTe phase change memory cells, *Appl. Phys. Lett.* **95**, 043108 (2009).
- [29] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* **140**, A1133 (1965).
- [30] S. Barraza-Lopez, T. P. Kaloni, S. P. Poudel, and P. Kumar, Tuning the ferroelectric-to-paraelectric transition temperature and dipole orientation of group-IV monochalcogenide monolayers, *Phys. Rev. B* **97**, 024110 (2018).
- [31] Q.-J. Ye, Z.-Y. Liu, Y. Feng, P. Gao, and X.-Z. Li, Ferroelectric Problem beyond the Conventional Scaling Law, *Phys. Rev. Lett.* **121**, 135702 (2018).
- [32] Q. Ye, X. Zhang, and X. Li, The phonon-related effective Hamiltonian method for displacive ferroelectric materials, *Electron. Struct.* **1**, 044006 (2019).
- [33] C. Xu, Y. Nahas, S. Prokhorenko, H. Xiang, and L. Bellaïche, Berezinskii-Kosterlitz-Thouless phase in two-dimensional ferroelectrics, *Phys. Rev. B* **101**, 241402(R) (2020).
- [34] K. M. Rabe and J. D. Joannopoulos, Theory of the structural phase transition of GeTe, *Phys. Rev. B* **36**, 6631 (1987).
- [35] B. J. Kooi and M. Wuttig, Chalcogenides by design, functionality through metavalent bonding and confinement, *Adv. Mater.* **32**, 1908302 (2020).
- [36] A. Singraber, T. Morawietz, J. Behler, and C. Dellago, Parallel multistream training of high-dimensional neural network potentials, *J. Chem. Theory Comput.* **15**, 3075 (2019).
- [37] H. Wang, L. Zhang, J. Han, and W. E, DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics, *Comput. Phys. Commun.* **228**, 178 (2018).
- [38] J. Behler, Perspective: Machine learning potentials for atomistic simulations, *J. Chem. Phys.* **145**, 170901 (2016).

- [39] V. L. Deringer, M. Caro, and G. Csányi, Machine learning interatomic potentials as emerging tools for materials science, *Adv. Mater.* **31**, 1902765 (2019).
- [40] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control, Math. Control Signal Syst.* **2**, 303 (1989).
- [41] Q. Tong, P. Gao, H. Liu, Y. Xie, J. Lv, Y. Wang, and J. Zhao, Combining machine learning potential and structure prediction for accelerated materials design and discovery, *J. Phys. Chem. Lett.* **11**, 8710 (2020).
- [42] J. Behler, R. Martonak, D. Donadio, and M. Parrinello, Metadynamics Simulations of the High-Pressure Phases of Silicon Employing a High-Dimensional Neural Network Potential, *Phys. Rev. Lett.* **100**, 185501 (2008).
- [43] J. S. Smith, O. Isayev, and A. E. Roitberg, ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost, *Chem. Sci.* **8**, 3192 (2017).
- [44] V. L. Deringer, N. Bernstein, G. Csányi, C. B. Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold, and S. R. Elliott, Origins of structural and electronic transitions in disordered silicon, *Nature (London)* **589**, 59 (2021).
- [45] J. Weinreich, A. Römer, M. L. Paleico, and J. Behler, Properties of α -brass nanoparticles. neural network potential energy surface, *J. Phys. Chem. C* **124**, 12682 (2020).
- [46] T. P. Kaloni, K. Chang, B. J. Miller, Q.-K. Xue, X. Chen, S.-H. Ji, S. S. P. Parkin, and S. Barraza-Lopez, From an atomic layer to the bulk: Low-temperature atomistic structure and ferroelectric and electronic properties of SnTe films, *Phys. Rev. B* **99**, 134108 (2019).
- [47] S. P. Poudel and S. Barraza-Lopez, Metastable piezoelectric group-IV monochalcogenide monolayers with a buckled honeycomb structure, *Phys. Rev. B* **103**, 024107 (2021).
- [48] F. C. Mocanu, K. Konstantinou, T. H. Lee, N. Bernstein, V. L. Deringer, G. Csányi, and S. R. Elliott, Modeling the phase-change memory material, $\text{Ge}_2\text{Sb}_2\text{Te}_5$, with a machine-learned interatomic potential, *J. Phys. Chem. B* **122**, 8998 (2018).
- [49] F. Zipoli and A. Curioni, Reactive potential for the study of phase-change materials: GeTe, *New J. Phys.* **15**, 123006 (2013).
- [50] F. Prudente, P. H. Acioli, and J. J. S. Neto, The fitting of potential energy surfaces using neural networks: Application to the study of vibrational levels of H_3^+ , *J. Chem. Phys.* **109**, 8801 (1998).
- [51] D. F. R. Brown, M. N. Gibbs, and D. C. Clary, Combining ab initio computations, neural networks, and diffusion Monte Carlo: An efficient method to treat weakly bound molecules, *J. Chem. Phys.* **105**, 7597 (1996).
- [52] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Machine Learning a General-Purpose Interatomic Potential for Silicon, *Phys. Rev. X* **8**, 041048 (2018).
- [53] G. Sosso, G. Miceli, S. Caravati, J. Behler, and Marco Bernasconi, Neural network interatomic potential for the phase change material GeTe, *Phys. Rev. B* **85**, 174103 (2012).
- [54] J. Behler, Representing potential energy surfaces by high-dimensional neural network potentials, *J. Phys.: Condens. Matter* **26**, 183001 (2014).
- [55] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [56] D. Yoo, K. Lee, W. Jeong, D. Lee, S. Watanabe, and S. Han, Atomic energy mapping of neural network potential, *Phys. Rev. Materials* **3**, 093802 (2019).
- [57] E. Prodan and W. Kohn, Nearsightedness of electronic matter, *Proc. Natl. Acad. Sci. USA* **102**, 11635 (2005).
- [58] S. Fias, F. Heidar-Zadeh, P. Geerlings, and P. W. Ayers, Chemical transferability of functional groups follows from the nearsightedness of electronic matter, *Proc. Natl. Acad. Sci.* **114**, 11633 (2017).
- [59] R. Hecht-Nielsen, Theory of the backpropagation neural network, in *Proceedings of the International 1989 Joint Conference on Neural Networks* (IEEE, Piscataway, NJ, 1989).
- [60] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. (Pearson, New York, 2008).
- [61] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, A survey of deep neural network architectures and their applications, *Neurocomputing* **234**, 11 (2017).
- [62] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, *PyTorch: An Imperative Style, HighPerformance Deep Learning Library* (NeurIPS, 2019).
- [63] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Duřak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen *et al.*, Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method, *J. Phys.: Condens. Matter* **22**, 253202 (2010).
- [64] H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B* **13**, 5188 (1976).
- [65] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [66] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.* **132**, 154104 (2010).
- [67] M. Dion, H. Rydberg, E. Schroder, D. C. Langreth, and B. I. Lundqvist, Van der Waals Density Functional for General Geometries, *Phys. Rev. Lett.* **92**, 246401 (2004).
- [68] J. Sun, A. Ruzsinszky, and J. P. Perdew, Strongly Constrained and Appropriately Normed Semilocal Density Functional, *Phys. Rev. Lett.* **115**, 036402 (2015).
- [69] A. P. Bartók and J. R. Yates, Regularized SCAN functional, *J. Chem. Phys.* **150**, 161101 (2019).
- [70] K. Toyoura, D. Hirano, A. Seko, M. Shiga, A. Kuwabara, M. Karasuyama, K. Shitara, and I. Takeuchi, Machine-learning-based selective sampling procedure for identifying the low-energy region in a potential energy surface: A case study on proton conduction in oxides, *Phys. Rev. B* **93**, 054112 (2016).
- [71] R. Cohen, Origin of ferroelectricity in perovskite oxides, *Nature (London)* **358**, 136 (1992).
- [72] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, Performance of various density-functional approximations for cohesive properties of 64 bulk solids, *New J. Phys.* **20**, 063020 (2018).
- [73] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).

- [74] J. Behler, Atom-centered symmetry functions for constructing highdimensional neural network potentials, *J. Chem. Phys.* **134**, 074106 (2011).
- [75] J. Vandermause S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events, *npj Comput. Mater.* **6**, 20 (2020).
- [76] W. Jeong, K. Lee, D. Yoo, D. Lee, and S. Han, Toward reliable and transferable machine learning potentials: Uniform training by overcoming sampling bias, *J. Phys. Chem. C* **122**, 22790 (2018).
- [77] L. Li, H. Li, I. D. Seymour, L. Koziol, and G. Henkelman, Pair-distribution-function guided optimization of fingerprints for atom-centered neural network potentials, *J. Chem. Phys.* **152**, 224102 (2020).
- [78] M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzsényi, and P. Marquetand, wACSF—Weighted atom-centered symmetry functions as descriptors in machine learning potentials, *J. Chem. Phys.* **148**, 241709 (2018).
- [79] K. Zhang, L. Yin, and G. Liu, Physically inspired atom-centered symmetry functions for the construction of high dimensional neural network potential energy surfaces, *Comput. Mater. Sci.* **186**, 110071 (2021).
- [80] M. Wuttig, V. L. Deringer, X. Gonze, C. Bichara, and J.-Y. Raty, Incipient metals: Functional materials with a unique bonding mechanism, *Adv. Mater.* **30**, 1803777 (2018).
- [81] J.-V. Raty and M. Wuttig, The interplay between Peierls distortions and metavalent bonding in IV–VI compounds: Comparing GeTe with related monochalcogenides, *J. Phys. D: Appl. Phys.* **53**, 234002 (2020).
- [82] J.-P. Gaspard, Vanishing-harmonicity and phase-change materials, *Phys. Status Solidi RRL* **15**, 2000536 (2021).
- [83] M. E. Fisher and M. N. Barber, Scaling Theory for Finite-Size Effects in the Critical Region, *Phys. Rev. Lett.* **28**, 1516 (1972).
- [84] G. A. T. Allan, Critical temperatures of ising lattice films, *Phys. Rev. B* **1**, 352 (1970).
- [85] J. Liu, L. Jin, Z. Jiang, L. Liu, L. Himanen, J. Wei, N. Zhang, D. Wang, and C.-L. Jia, Understanding doped perovskite ferroelectrics with defective dipole model, *J. Chem. Phys.* **149**, 244122 (2018).
- [86] C. W. Li, J. Hong, A. F. May, D. Bansal, S. Chi, T. Hong, G. Ehlers, and O. Delaire, Orbital driven giant phonon anharmonicity in SnSe, *Nat. Phys.* **11**, 1063 (2015).
- [87] Z. Jiang, R. Zhang, D. Wang, D. Sichuga, C.-L. Jia, and L. Bellaïche, Strain-induced control of domain wall morphology in ultrathin PbTiO₃ films, *Phys. Rev. B* **89**, 214113 (2014).
- [88] J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, *Int. J. Quantum Chem.* **115**, 1032 (2015).
- [89] D. Wang, J. Liu, J. Zhang, S. Raza, X. Chen, and C.-L. Jia, Ewald summation for ferroelectric perovskites with charges and dipoles, *Comput. Mater. Sci.* **162**, 314 (2019).
- [90] Y. Yao and Y. Kanai, Plane-wave pseudopotential implementation and performance of SCAN meta-GGA exchange-correlation functional for extended systems, *J. Chem. Phys.* **146**, 224105 (2017).
- [91] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, Accurate first-principles structures and energies of diversely bonded systems from an efficient density functional, *Nat. Chem.* **8**, 831 (2016).
- [92] H. Peng, Z. Yang, J. P. Perdew, and J. Sun, Versatile van der Waals Density Functional Based on a Meta-Generalized Gradient Approximation, *Phys. Rev. X* **6**, 041005 (2016).