Dual adaptive sampling and machine learning interatomic potentials for modeling materials with chemical bond hierarchy

Hongliang Yang,[1,2,3] Yifan Zhu,[1,2] Erting Dong,[3,4] Yabei Wu,[3,4] Jiong Yang,[5,*] and Wenqing Zhang[3,4,*]

[1]*State Key Laboratory of High Performance Ceramics and Superfine Microstructure, Shanghai Institute of Ceramics, Chinese Academy of Sciences, Shanghai 200050, China*

[2]*University of Chinese Academy of Sciences, Beijing 100049, China*

[3]*Department of Physics and Shenzhen Institute for Quantum Science & Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

[4]*Guangdong Provincial Key Lab for Computational Science and Materials Design, and Shenzhen Municipal Key-Lab for Advanced Quantum Materials and Devices, Southern University of Science and Technology, Shenzhen 518055, China*

[5]*Materials Genome Institute, Shanghai University, Shanghai 200444, China,*

(Received 13 May 2021; revised 7 August 2021; accepted 9 September 2021; published 27 September 2021)

The development of reliable and flexible machine learning based interatomic potentials (ML-IPs) is becoming increasingly important in studying the physical properties of complex condensed matter systems. Besides the structure descriptor model for total energy decomposition, the trial-and-error approach used in the design of the training dataset makes the ML-IP hardly improvable and reliable for modeling materials with chemical bond hierarchy. In this work, a dual adaptive sampling (DAS) method with an on the fly ambiguity threshold was developed to automatically generate an effective training dataset covering a wide temperature range or a wide spectrum of thermodynamic conditions. The DAS method consists of an inner loop for exploring the local configuration space and an outer loop for covering a wide temperature range. We validated the developed DAS method by simulating thermal transport of complex materials. The simulation results show that even with a substantially small dataset, our approach not only accurately reproduces the energies and forces but also predicts reliably effective high-order force constants to at least fourth order. The lattice thermal conductivity and its temperature dependence were evaluated using the Green-Kubo simulations with ML-IP for $CoSb_3$ with up to third-order phonon scattering, and those for Mg_3Sb_2 with up to fourth-order phonon scattering, and all show good agreements with experiments. Our work provides an avenue to effectively construct a training dataset for ML-IP of complex materials with chemical bond hierarchy.

DOI: 10.1103/PhysRevB.104.094310

I. INTRODUCTION

Interatomic potentials with validated parametrization have important applications in simulating various physical properties in condensed matter physics, materials science, and even chemistry [1]. Density functional theory (DFT) calculations can accurately describe interatomic interactions, but usually require large computational resources and show nonlinear scaling of computational time which makes the simulation of large-scale complex systems a big challenge. Countless types of effectively parametrized potentials have been developed over the decades such as Tersoff potentials [2] and embedded atom method (EAM) potentials [3]. Although the calculations based on those empirical interatomic potentials could be fast, the reliability hardly meets the needs of complex materials simulation in scientific and industrial applications [4], let alone the severe limit and difficulty in universality of interatomic potential for complex systems with multiple compositions. Recently, machine learning based interatomic potential (ML-IP) has been proposed to compensate for the disadvantages of DFT and empirical force fields, which have been shown to be computationally efficient but with accuracy comparable to DFT results [5–7]. The ML-IP, also known as mathematical potential, has a large number of fitting parameters compared to empirical interatomic potentials [4]. These fitting parameters need to be learned from DFT computational data. The generation of the training dataset for ML-IP has attracted intense interest [8–10], since it is required for the training dataset to fully cover the configuration space for targeted application scenarios to guarantee the high accuracy of ML-IP. Consequently, a large training dataset will be generated if an inefficient sampling method is adopted, which will further hinder the application of ML-IP.

In order to realize efficient sampling, a class of active learning based sampling methods has recently been proposed to design a training dataset for ML-IP [8–10]. The active learning based methods usually explore the configuration space in an iterative manner. In each iteration loop, the structures are first sampled using molecular dynamics (MD) simulation under specific thermodynamic conditions, then some structures which are believed to substantially improve the accuracy of ML-IP are selected for DFT calculations, and finally the DFT calculation data are added to the training dataset to

*jiongy@t.shu.edu.cn; zhangwq@sustech.edu.cn

 ©2021 American Physical Society

refine the potential. Different query strategies, such as query by committee [8,10–12], predictive variance [13,14], and the *D*-optimality criterion [9,15–17], have been used to design training datasets on the fly. The predictive variance and the *D*-optimality criterion are adopted in Gaussian process based potentials [13,14] and moment tensor potentials [9,15–17], respectively. The query by committee strategy, which selects the training structures with high ambiguity in the energy and/or forces predicted by an ensemble of independently trained potentials, has been investigated in several neural network based potentials [8,10–12]. A threshold for ambiguity usually needs to be set, which is a quantity characterizing the extent of disagreement of predicted energy and/or forces over several independently trained potentials [8,10–12]. The setting of the threshold is of vital importance to the success of the sampling method [10]. If the threshold is too low, too many structures will be added into the training dataset, rendering an increase of computational cost; if the threshold is too high, the active learning iterations will be terminated prematurely, and the accuracy of ML-IP will be greatly affected. It is challenging to set a suitable threshold value, especially for complex applications covering a wide spectrum of thermodynamic conditions. For example, the threshold is recommended to be a value slightly higher than the training accuracy achievable by ML-IP [10]. However, the training accuracy of ML-IP is closely related to the training structures; e.g., the training accuracy of ML-IP decreases with increasing sampling temperature [16] or decreases with increasing disorder of the structures [9,18]. Therefore, an efficient sampling method that can adaptively update the ambiguity threshold for different thermodynamic conditions is of great importance for complex applications.

Simulating thermal transport is one of the typical applications of ML-IP, since thermal transport plays an important role in many fields such as thermoelectrics [19–21], and thermal management composites [22]. The quantitative prediction of the lattice thermal conductivity of complex materials, such as those with chemical bond hierarchy [23–27], remains challenging even for DFT calculations. Simulating phonon thermal transport has inherently multilevel complexity because it requires not only accurate phonon spectra from the second-order force constants but also reliable nonlinear phonon-phonon interactions from third-order and even fourth-order force constants [28–31] covering a wide temperature range. Therefore, thermal transport can be used to test the efficiency and accuracy of sampling methods. Recent studies of ML-IP in thermal transport have focused on materials with third-order anharmonicity where the lattice thermal conductivity follows the traditional T^{-1} relationship [16,18]. It may be appropriate to generate a training dataset at one arbitrarily given temperature for these materials [16,32,33] only if effective third-order force constants barely vary with temperature. However, it is not adequate to sample at a single temperature in materials with chemical bond hierarchy, as the displacement pattern of the atoms at high temperatures is completely different from that at low temperatures [23,27]. For example, these structures with large displacements which are rarely observed under MD at low temperatures must be added to the training dataset to achieve stable MD simulations [34]. One could also simply combine *ab initio* MD trajectories of different temperatures as the training dataset [18], but this has been found to have many drawbacks, such as a large number of training structures (e.g., 10 000 structures for crystalline Si [18]) and unstable dynamics [9,34].

In this work, we develop a dual adaptive sampling (DAS) method for generating ML-IP training datasets in materials with chemical bond hierarchy and apply this method in thermal transport to test its efficiency and reliability. Our method can accurately predict not only the energies and forces, but also the lattice thermal conductivity of complex materials, while reducing the number of training structures. The rest of the paper is organized as follows. In Sec. II, we give a detailed description of the DAS method and the training error estimation of the ML-IP. In Sec. III, we show a comparison of lattice thermal conductivity results calculated by the Boltzmann transport equation method and the Green-Kubo method with the obtained ML-IP for $CoSb_3$ and Mg_3Sb_2, and also a comparison of the theoretically calculated thermal conductivity with the experimental values. In Sec. IV, we discuss the advantages of the sampling method and numerically verify that second-, third-, and fourth-order force constants can be accurately predicted. A brief conclusion is given in Sec. V.

II. DUAL ADAPTIVE SAMPLING METHOD

In complex applications such as thermal transport modeling, the training structures should ideally cover the configuration space of the considered wide spectrum of thermodynamic conditions such as temperature. It is relatively easy to design a training dataset for materials with simple crystal structure. The difficulty for a sampling method to achieve reliable and flexible ML-IP comes from the intrinsic complex chemical bonding in complex materials, specifically, the fact that the response of atoms at a given temperature can vary substantially due to differences in chemical bond strength [23]. As the temperature increases, the displacement magnitude of different species from the equilibrium positions becomes larger and the corresponding distribution becomes broader as shown in Fig. 1(a). This is the physical reason why the extrapolation of a training dataset sampled at one arbitrarily given temperature to a wider temperature spectrum is usually safe for rigid materials, but leads to unstable dynamics for materials with weak chemical bonds [9,34]. Therefore, based on the fact that the spatial extent of configuration space explored by atoms varies depending on the chemical bond strength, we developed a DAS method, which consists of an inner adaptive loop and an outer loop. The inner loop deals with the local configuration space [i.e., sampling inside each peak in Fig. 1(a)], while the outer loop aims to cover a wide temperature range [i.e., covering all the peaks in Fig. 1(a)].

The schematic workflow of the outer loop of DAS is shown in Fig. 1(b). The initial training dataset contains several structures using a simple sampling (preliminary sampling), such as short-time *ab initio* MD or just random displacements of a few atoms. The preliminary dataset is used to train a rough potential to drive sampling iterations, and the dataset is to be expanded to explore more configuration space in an adaptive self-consistent manner (inner loop, which will be discussed shortly). Several temperatures, equally spaced over a wide temperature range, are selected as the center temperatures of

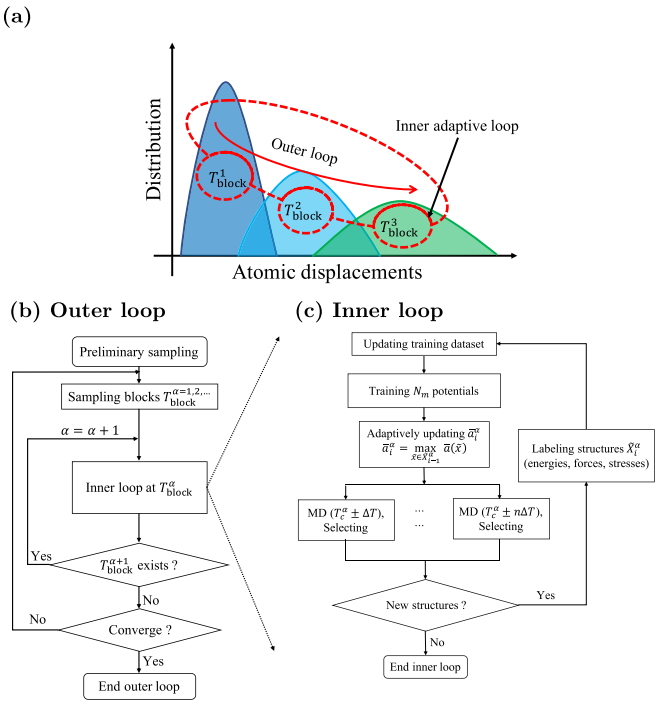

FIG. 1. Schematic explanation for the dual adaptive sampling method. (a) Schematic atomic displacement distribution pattern in the materials with complex chemical bonding under varying temperature. The outer loop (b) and inner loop (c) in the DAS workflow. The temperatures considered in the inner loop at sampling block $T^{\alpha}_{\text{block}}$ are $T^{\alpha}_c \pm \Delta T, \ldots, T^{\alpha}_c \pm n\Delta T$, where n is a positive integer; T^{α}_c and ΔT are the central temperature of the sampling block α and a small temperature deviation, respectively. The thermodynamic conditions are represented by temperature in the diagram for brevity. All the N_m potentials are used to calculate the ensemble ambiguity. The MD simulation for exploring the configuration space uses the potential with the lowest training loss. For any MD simulations of the ith iteration, these structures giving ambiguity greater than $\bar{a}^{\alpha}_i$ are selected for subsequent processing (see the main text for details).

the sampling blocks in the outer loop. Each sampling block represents one inner loop. To ensure the accuracy of the potentials, the convergence of the outer loop is checked at the end of all inner loops, and the criterion should be consistent with the applications of interest. In this work, the convergence criterion for the outer loop requires that the ML-IP based energies, forces, and phonon dispersions are in good agreements with the DFT results, and this is essential in reliably evaluating the lattice thermal transport.

Each inner loop is responsible for several temperatures close to the central temperature. Practically, it is efficient to cover those structures at close-by temperatures within a specific inner loop sampling block. As shown in Fig. 1(c), each inner loop of DAS consists of three main steps: (1) training N_m ML-IPs; (2) selecting new structures from MD simulations; (3) updating structure labels (energy, forces, and stresses) for the training dataset. At the beginning of each iteration of the inner loop, N_m potentials with the same functional form are trained independently with the same dataset but starting from different random initialization parameters, and then the one with the lowest loss on the training dataset is adopted to explore the configuration space by MD simulation. Structural sampling based on MD simulation has the advantage of sampling true physical states at a specific thermodynamic condition [35]. Those structures selected from MD simulations at a specific time interval are then scrutinized by the query by committee method based on the calculated ensemble ambiguity to pick those structures valuable to the training dataset. In the query by committee method [8], only those structures yielding ensemble ambiguity higher than a threshold value are selected as candidates, which can avoid selecting structures that have already been accurately predicted by the potential trained on the current dataset. As we have mentioned, previous studies have employed fixed threshold values for ensemble ambiguity [8,10–12]. In this work, we introduce an adaptive way to update the threshold of ensemble ambiguity on the fly.

The ensemble ambiguity $a(x)$ on an unlabeled input x measures the disagreement of the output over these independently trained potentials [36]. The ensemble ambiguity is a lower bound for the average error [36]. If the error on the training dataset for each potential model is low, then the ambiguity of the training points is also low, which ensures that the pattern giving high ambiguity is unlikely to be a neighbor of one of the training structures. Therefore, any new candidate structure yielding high ambiguity on force must contain one or more atomic local environments differing significantly from the existing training structures. In other words, assuming that all the structures generated using MD simulations give low ambiguity, the space spanned by the existing training structures could be believed to satisfactorily fully cover the configuration space at the given thermodynamic conditions. This inspired us to define a criterion for the convergence of the inner loop as

$$\bar{a}(x) \leqslant \bar{a}^{\alpha}_i = \max_{\tilde{x}\in\tilde{X}^{\alpha}_{i-1}} \bar{a}(\tilde{x}), \quad \text{for all} \quad x \in X^{\alpha}_i, \tag{1}$$

where $\bar{a}(x)$ is the ensemble ambiguity on an unlabeled structure x, X^{α}_i is the set of structures from MD simulations in the ith iteration of the sampling block α, and $\tilde{X}^{\alpha}_{i-1}$ is the set of structures added to the training dataset in the $(i-1)$th iteration. In this work, we employ atomic forces to estimate ensemble ambiguity, since forces are more sensitive to the atomic local environment [10,35]. The ensemble ambiguity $a(x)$ is calculated by [10]

$$\bar{a}(x) = \max_j \sqrt{\frac{1}{N_m}\sum_m \left\| \boldsymbol{f}_{j,m}(x) - \bar{\boldsymbol{f}}_j(x) \right\|^2} \quad \text{and}$$

$$\bar{\boldsymbol{f}}_j(x) = \frac{1}{N_m}\sum_m \boldsymbol{f}_{j,m}(x), \tag{2}$$

where N_m is the number of trained potentials and $\boldsymbol{f}_{j,m}$ is the force on atom j predicted by the potential m. The criterion (1) basically states that the ambiguity of the training dataset convergence to a value, and the accuracy of the ML-IP cannot be further improved by including new structures under the current thermodynamic condition. In fact, as the number of iterations increases, the description of the dynamics becomes more and more accurate due to the increased accuracy of the ML-IPs. Therefore, the structures added to the training dataset in the $(i-1)$th iteration becomes more suitable to check whether convergence is achieved at the ith iteration.

At the first iteration of each sampling block, the ambiguity threshold is set to a very small value, for example, 0.01 eV/Å in this work. In this way, the convergence criterion of the inner loop is not fixed but adaptively updated on the fly. If the convergence criterion is satisfied, the inner loop is terminated. Otherwise, these structures with $\bar{a}$ higher than $\bar{a}_i^{\alpha}$ are taken as candidate structures entering into the adapting procedure.

The local environment descriptors represent atomic local neighborhood environments with continuous atomic coordinates in space as the features to be input into various ML-IP models [37]. Many different types of local environment descriptors have been proposed, among them the atom-centered symmetry functions [5], smooth overlap of atomic positions [37], and the moment tensor potential [6], and some others [7,38] are well accepted and tested for different applications. The current work chose the moment tensor potential [6,39] to fit the training dataset generated by our sampling method, due to its computational efficiency and reasonable universality for different applications [40].

$CoSb_3$ skutterudite and Mg_3Sb_2 are chosen as the testing cases for our methodology. Both are promising intermediate-temperature thermoelectric materials. The lattice thermal conductivity of $CoSb_3$ was observed to follow a temperature scaling T^{-1}, consistent with the picture of third-order phonon scattering. This had been explained by many DFT-based calculations in Boltzmann transport theory [41,42]. Simulations in the Boltzmann transport equation (BTE) method using ML-IP also could reproduce lattice thermal conductivities of $CoSb_3$ at the temperature near or above its Debye temperature of 307 K [16]. However, the lattice thermal conductivity calculated by the Green-Kubo method is systematically underestimated compared to that calculated by the BTE method using the moment tensor potential as adopted in this work [16]. The Mg_3Sb_2-based compounds, consisting of ionic Mg^{2+} and covalent $(Mg_2Sb_2)^{2-}$ layers, also attracted intensive interest due to their excellent thermoelectric performance near room temperature [27,43,44]. Interestingly, the lattice thermal conductivity of Mg_3Sb_2 shows a temperature dependence deviating from T^{-1}, and this has been explained by temperature dependence of force constants [27]. In addition, the *ab initio* MD trajectories show that the displacement distribution in Mg_3Sb_2 at different temperatures differs significantly, approximately isotropic at 300 K but strongly anisotropic at 700 K for Mg ions in $(Mg_2Sb_2)^{2-}$ layers [27].

We implemented the DAS method in our homemade code [45]. The MD simulation with ML-IP was carried out by the LAMMPS package [46]. All *ab initio* calculations were carried out using the Vienna *ab initio* simulation package (VASP) [47] with projector-augmented wave (PAW) pseudopotentials [48] and the generalized approximation of Perdew-Burke-Ernzerhof form [49] for the exchange correlation energy functional. Unless otherwise specified, an energy convergence criterion of 10^{-5} eV and a plane-wave energy cutoff of 450 eV were adopted for all *ab initio* calculations. The time step of the MD simulation was set to 1.0 fs. The geometry optimization of the primitive cell was done with a $13 \times 13 \times 13$ and $15 \times 15 \times 15$ Monkhorst-Pack **k**-point mesh for $CoSb_3$ and Mg_3Sb_2, respectively. During the sampling process, all MD simulations were started from a supercell by replicating the DFT-relaxed unit cell. For $CoSb_3$ and Mg_3Sb_2, we used a $2 \times 2 \times 2$ supercell (256 atoms) by replicating the conventional unit cell and a $4 \times 4 \times 4$ supercell (320 atoms) by replicating the primitive unit cell, respectively. To improve the accuracy of the force, a Monkhorst-Pack **k** mesh of $2 \times 2 \times 2$ was used in calculating the energies, stresses, and forces.

The preliminary dataset was constructed by selecting a structure every two steps from a 20.0 fs *ab initio* MD simulation at 700 K. It is worth noting that those structures with strongly anisotropic displacement distribution of Mg atoms in Mg_3Sb_2 are not included in the preliminary dataset. In our tests, the preliminary dataset (the number of structures, sampling temperature, and sampling method) has insignificant effect on the accuracy of the potential as shown in Table S5 of the Supplemental Material [50]. First, four temperatures were taken at 200 K intervals below 800 K as the central temperatures of the sampling block. Then, these structures at four different temperatures around each central temperature, with 50 K as a step, were taken into one sampling block. More details on the parameter settings for each sampling block can be found in the Supplemental Material [50]. Considering the temperature dependence of lattice constants, samplings at volume compression (∼–10% for Mg_3Sb_2 and ∼–5% for $CoSb_3$) or expansion (∼10% for Mg_3Sb_2 and ∼5% for $CoSb_3$) at each temperature are used to increase transferability of potentials in volume space. It should be addressed that parallel sampling at several temperatures and different volumes is not the key to the success of the current adaptive sampling strategy, but it does speed up the efficiency and convergence structure selection process. In each inner loop, we simultaneously trained five potentials of the same functional form with random initialization. All MD samplings for any iteration of the inner loop started with random initial velocities at the given temperature and volume. Each MD simulation ran for a total of 10.0 ps, sampling one structure at 20 step intervals.

In order to avoid introducing too many correlated structures, when the number of structures yielding $\bar{a}(x)$ higher than $\bar{a}_i^{\alpha}$ in a MD simulation is greater than 1, we only pick a representative structure with $\bar{a}(x)$ close to the peak of the ambiguity distribution of these structures. Therefore, the maximum number of structures selected for one iteration in the inner loop is the number of MD simulations. If no structure is selected in one iteration of the inner loop, then the inner loop terminates. In the two systems considered in this work, the outer loop convergence criterion could be reached after all the sampling blocks have gone through only one cycle. The final training datasets of $CoSb_3$ and Mg_3Sb_2 contain 359 and 487 structures, respectively.

To validate the accuracy of the developed ML-IPs, the energies and forces of randomly selected structures calculated by ML-IP were compared with the results of DFT calculations. These test structures were selected from the MD trajectories at every 0.1 ns after 1.0 ns running. Figure 2 shows the forces evaluated in both ML-IP and DFT for test structures. It is found that the mean absolute errors of the energies are 5.7, 6.1, 6.3, and 6.7 meV/atom for $CoSb_3$ at 100, 300, 500, and 700 K, respectively, and the corresponding mean absolute errors in the forces are 24.6, 28.2, 32.0, and 37.4 meV/Å. The mean absolute errors of the energies are 1.0, 0.48, 1.04, and 0.36 meV/atom for Mg_3Sb_2 at 100, 300, 500, and 700 K and the

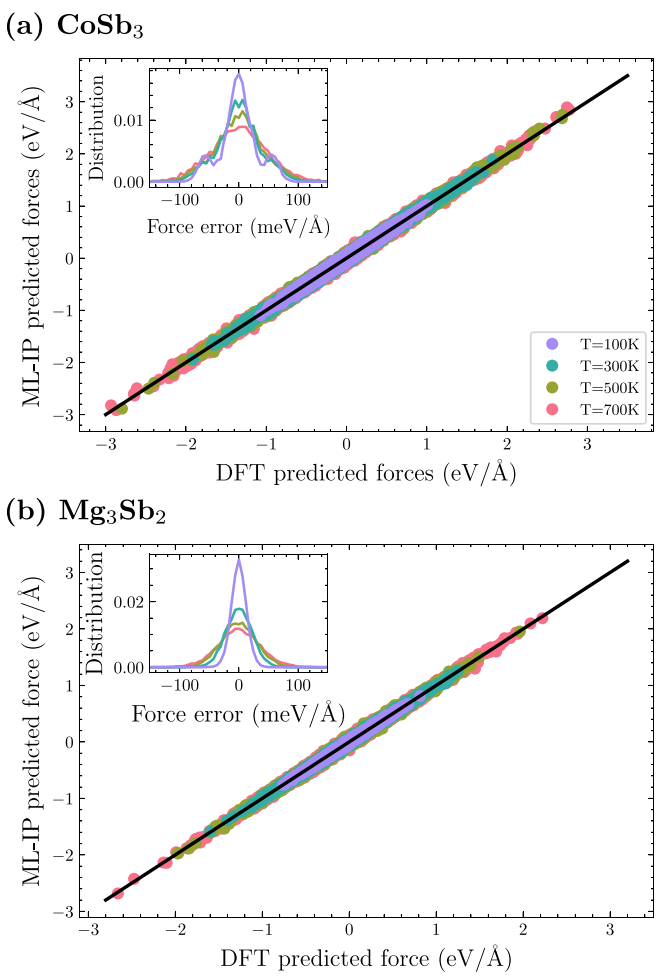

FIG. 2. Comparison of the DFT predicted forces and the ML-IP predicted forces (a) $CoSb_3$ and (b) Mg_3Sb_2. The distribution of absolute error in the forces is also shown as an inset.

corresponding mean absolute errors in the forces are 10.0, 17.9, 23.9, and 28.2 meV/Å, respectively. The error in forces is small and normally distributed around zero as shown in Fig. 2 (inset, upper left), indicating that the ML-IP can predict DFT forces with high accuracy. The error in forces of our ML-IP for $CoSb_3$ is similar to that of a recent study [16], but interestingly, the number of force components in our training structures is 275 712 for the 300–800 K temperature range, comparable to the number of force components of 275 328 in the training structures at 500 K of Ref. [16]. This clearly supports the sampling efficiency of the current DAS approach. The dot product between the normalized DFT forces and ML-IP forces ($\hat{\boldsymbol{F}}_{\text{DFT}} \cdot \hat{\boldsymbol{F}}_{\text{ML-IP}}$) is near unity (around 0.99) for all structures ($CoSb_3$ and Mg_3Sb_2). Overall, the energy and forces predicted by ML-IP agree very well with DFT calculations in both magnitude and direction.

III. APPLICATIONS IN LATTICE THERMAL CONDUCTIVITY

After modeling the two testing systems using the DAS method, we used the developed ML-IP to calculate the lattice thermal conductivities. We used two typical methods to calculate the temperature-dependent lattice thermal conductivity, the perturbation-based BTE method and the Green-Kubo method. The former accounts for up to third-order anharmonicity while the latter naturally includes all high-order anharmonicity [51,52]. Under the BTE method, the forces in the supercells with finite displacements were calculated by both DFT and ML-IP, in order to provide a benchmark for the developed methodology in this work. A higher energy convergence criterion of 10^{-8} eV was used in calculating force constants for Mg_3Sb_2 to reduce the noise in forces. The second- and third-order force constants of $CoSb_3$ were calculated in the $2 \times 2 \times 2$ supercell (128 atoms) with a Monkhorst-Pack **k** mesh of $2 \times 2 \times 2$. The second-order force constants of Mg_3Sb_2 were calculated in the $4 \times 4 \times 4$ supercell (320 atoms) with a Monkhorst-Pack **k** mesh of $2 \times 2 \times 2$, while the third-order force constants were calculated in the $3 \times 3 \times 3$ supercell (135 atoms) with a Monkhorst-Pack **k** mesh of $2 \times 2 \times 2$. The cutoff of interaction range for $CoSb_3$ and Mg_3Sb_2 were set to 5.0 and 6.6 Å, respectively. The lattice thermal conductivities of BTE were computed with the SHENGBTE package based on a full iterative solution [53]. The BTE was solved with a $15 \times 15 \times 15$ **q**-point grid for both systems.

The Green-Kubo method which relates the lattice thermal conductivity $\kappa_{\alpha\beta}(\alpha, \beta = x, y, z)$ with heat flux can be expressed as [51,52]

$$\kappa_{\alpha\beta} = \frac{1}{k_B T^2 V} \int_0^\infty \langle J_\alpha(0) J_\beta(t) \rangle dt, \tag{3}$$

where k_B is the Boltzmann constant, T is temperature, V is the volume of the simulation box, $\langle \cdots \rangle$ is the average over time, and $\boldsymbol{J}(t)$ is the heat flux. We used the direct method to estimate the integral in Eq. (3) [54,55], specifically, first specifying a range of the correlation time where the integral is converged, and then taking the mean value and standard error in that range as the average value and error estimate, respectively. The heat flux is expressed as [56,57]

$$\boldsymbol{J} \equiv \frac{d}{dt} \sum_i \boldsymbol{r}_i E_i = \sum_i \boldsymbol{v}_i E_i + \sum_i \boldsymbol{W}_i \cdot \boldsymbol{v}_i, \tag{4}$$

where $\boldsymbol{r}_i$, $\boldsymbol{v}_i$, $\boldsymbol{W}_i$, and E_i are the velocity, position, atomic virial stress, and total energy of atom i, respectively. The atomic virial stress is defined by

$$\boldsymbol{W}_i = \sum_{j \neq i} (\boldsymbol{r}_j - \boldsymbol{r}_i) \otimes \frac{\partial U_j}{\partial \boldsymbol{r}_i} = \sum_{j \neq i} \boldsymbol{r}_{ij} \otimes \boldsymbol{F}_{ij}, \tag{5}$$

where U_j is the potential energy of atom j and $\otimes$ is the Cartesian outer product operator. Equation (3) holds for the general ML-IP in periodic boundary conditions. According to Eq. (4), the heat flux can be reliably determined as long as energies and forces are accurately predicted. The Green-Kubo method based on DFT is good for complex materials but the extremely huge computation cost makes it infeasible in large-scale systems, which makes reliable and robust interatomic potentials, e.g., the MTP-based [16,17] and a few other ML-IPs [18,58], important for simulating complex materials.

The simulation cells of $CoSb_3$ and Mg_3Sb_2 under the Green-Kubo method contain 2048 and 2560 atoms, respectively. We did not consider isotope effects in our simulations

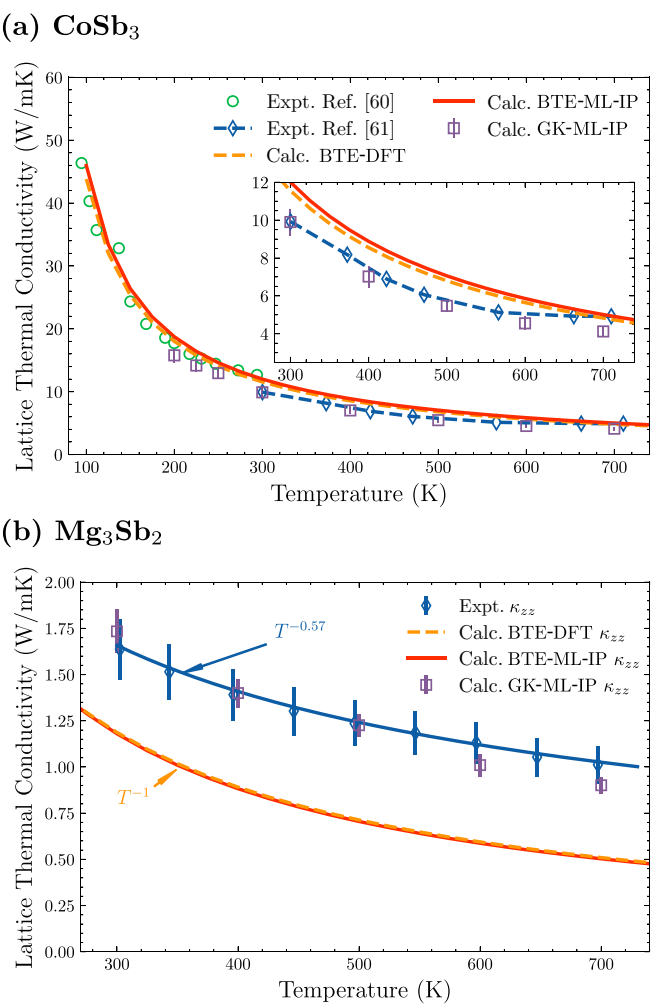

FIG. 3. Lattice thermal conductivities of theoretical and experimental data for (a) $CoSb_3$ and (b) Mg_3Sb_2. The lattice thermal conductivity obtained by the Green-Kubo method based on ML-IP is labeled as GK-ML-IP. The experimental data for $CoSb_3$ are taken from Morelli *et al.* [60] and Caillat *et al.* [61]. The experimental data for Mg_3Sb_2 are taken from Zhu *et al.* [27].

since their contribution to the lattice thermal conductivity is negligible, less than 1% for both systems in our BTE calculations. The experimental lattice constants were adopted for lattice thermal conductivity calculation at each temperature [43,59]. The convergence of lattice thermal conductivity with respect to simulation size was carefully checked as shown in Fig. S4 of the Supplemental Material [50]. At each temperature, we obtained the average heat flux autocorrelation function by 40 independent simulations starting from different initial velocities. Each simulation was first equilibrated for 100.0 ps in a *NVT* ensemble, and subsequently switched to a *NVE* ensemble first run for 100.0 ps, after which heat flux data were recorded for 2.0 ns every ten steps. The converged correlation times of $CoSb_3$ and Mg_3Sb_2 were [200 ps, 500 ps] and [50 ps, 100 ps], respectively.

Figure 3(a) shows the computational and experimental temperature-dependent lattice thermal conductivities for $CoSb_3$. The lattice thermal conductivities of $CoSb_3$ obtained by solving the BTE using the second- and third-order force constants calculated by DFT or ML-IP are in good agreement, with a difference of about 5.0%. The lattice thermal conductivity calculated by BTE in this work also differs very little from the previous DFT calculation [41], which is about 2.0%. The lattice thermal conductivities of $CoSb_3$ obtained by the Green-Kubo method are an excellent match to the experimental values of single crystals [35] as shown in Fig. 3(a), 9.90 ± 0.69 W/mK at 300 K, with an error of about 5.0% compared to the experimental measurement. Thus, the current ML-IP with the DAS method is reliable for predicting lattice thermal conductivities of $CoSb_3$, under both the BTE and Green-Kubo methods. Notice that the previous lattice thermal conductivity results [16] from the Green-Kubo method with ML-IP showed a systematic underestimation compared to experimental results, exceeding about 15% at 300 and 500 K. Furthermore, the elastic constants, i.e., c_{11}, c_{12}, and c_{44} of $CoSb_3$, are calculated with ML-IP, and the values are 179.5, 38.3, and 48.6 GPa, respectively. They are also in good agreement with the previous calculations [62,63].

The lattice thermal conductivities of pristine Mg_3Sb_2 have a relatively weak temperature dependence as experimentally observed [64–66]. The lattice thermal conductivities predicted by the Green-Kubo method with ML-IP agree very well with the experimental results on a single crystal as shown in Fig. 3(b). For example, the predicted value is 1.73 ± 0.12 W/mK at 300 K, very close to the experimental value of 1.69 ± 0.16 W/mK. Figure 3(b) also plots the lattice thermal conductivities from BTE using the second- and third-order force constants calculated by ML-IP or DFT; they follow a classical T^{-1} temperature-dependent relation. The difference between the lattice thermal conductivities calculated by the BTE method using the force constants from ML-IP and DFT is less than 1.0%, indicating that the accuracy of our developed ML-IP is comparable to DFT as a force calculator. However, the BTE method systematically underestimates the lattice thermal conductivities of Mg_3Sb_2 by about 30%, 40%, and 40% at 300, 500, and 700 K compared to the Green-Kubo method, respectively. This indicates clearly that the BTE method considering only third-order phonon scattering is not adequate for this material. These results imply that high-order anharmonicity is important for accurate prediction of lattice thermal conductivities of Mg_3Sb_2. A recent theoretical study [27] extracting the effective force constants from *ab initio* MD trajectories also predicted a weak temperature dependence of lattice thermal conductivity but systematically overestimated the lattice thermal conductivities compared to experimental results. Overall, the lattice thermal conductivities calculated by the present Green-Kubo method with ML-IP agree well with the experimental results.

We also tested the effect of the sampling temperature range on the calculated lattice thermal conductivities to demonstrate the importance of covering a wide range of temperature for the accuracy of ML-IP. We generated the training structures for both testing systems using the DAS method with the same setting as mentioned above, except that a narrow temperature range of 250–350 K was adopted. The training datasets for $CoSb_3$ and Mg_3Sb_2 in the narrow temperature range contain 120 and 132 structures, respectively. Figure S5 of the

Supplemental Material [50] shows the lattice thermal conductivities for a narrow temperature range of sampling, labeled as ML-IP (narrow). The calculated lattice thermal conductivities of $CoSb_3$ using the BTE with ML-IP (narrow) agree well with the DFT results. However, the error of lattice thermal conductivities obtained by BTE with ML-IP (narrow) reach about 9.0% in Mg_3Sb_2 comparing with the results from the BTE based on DFT. The lattice thermal conductivities calculated by the Green-Kubo method with ML-IP (narrow) have an error of 10%–15% compared to results from the wide temperature spectrum sampling for the two testing systems at 300 and 500 K. However, the MD simulations with ML-IP (narrow) are unstable for modeling Mg_3Sb_2 at 700 K due to the displacement pattern of Mg_3Sb_2 at 300 K being completely different from that at 700 K (see MD trajectories as shown in Fig. S7 for details). Therefore, covering a wide temperature spectrum when generating a training dataset for Mg_3Sb_2 is indispensable for the accuracy and transferability of ML-IP, at least for the Green-Kubo method.

The lattice thermal conductivities for Mg_3Sb_2 were also calculated using the Green-Kubo method with ML-IP but with a training dataset generated from a fixed value for the ambiguity threshold, i.e., a low threshold 0.05 eV/Å and a high threshold 0.1 eV/Å. The final lattice thermal conductivities were shown in Fig. S6 of the Supplemental Material [50]. The final training datasets of the low and high thresholds contain 671 and 399 structures, respectively. In the above two cases with different fixed thresholds, the lattice thermal conductivities are close to the experimental value at 300 K with an error of below 7%, but the temperature dependence of the lattice thermal conductivity cannot be correctly reproduced. Furthermore, the number of structures generated using the fixed threshold 0.05 eV/Å is about 40% higher than that of the DAS method using an on the fly threshold.

IV. DISCUSSIONS

The physics behind the reliability and transferability of DAS is the versatile response of materials with inherently multilevel complexity to varying thermodynamic conditions. Due to the coexistence of different chemical bonds in one material, the response pattern of different atoms could vary substantially at different temperatures [23,26,27], e.g., the obviously different displacement distributions at 300 and 700 K in Mg_3Sb_2 [27]. Figure 4(a) shows the distribution of displacements of the atoms away from their equilibrium positions. The large overlap of displacement distributions at different temperatures suggests that taking several temperatures near a central temperature and combining the corresponding structures into one sampling block can improve the sampling efficiency because those structures do have similar local configuration space to some extent. The peak position of the displacement distribution increases with increasing temperature for both $CoSb_3$ and Mg_3Sb_2. For example, the peak position at 700 K is about 1.5 times higher than that at 300 K. However, the full width at half maximum of the distribution of Mg_3Sb_2 is nearly twice that of $CoSb_3$, implying that the configuration space explored by atoms in Mg_3Sb_2 is much larger than that in $CoSb_3$. The larger local configuration space

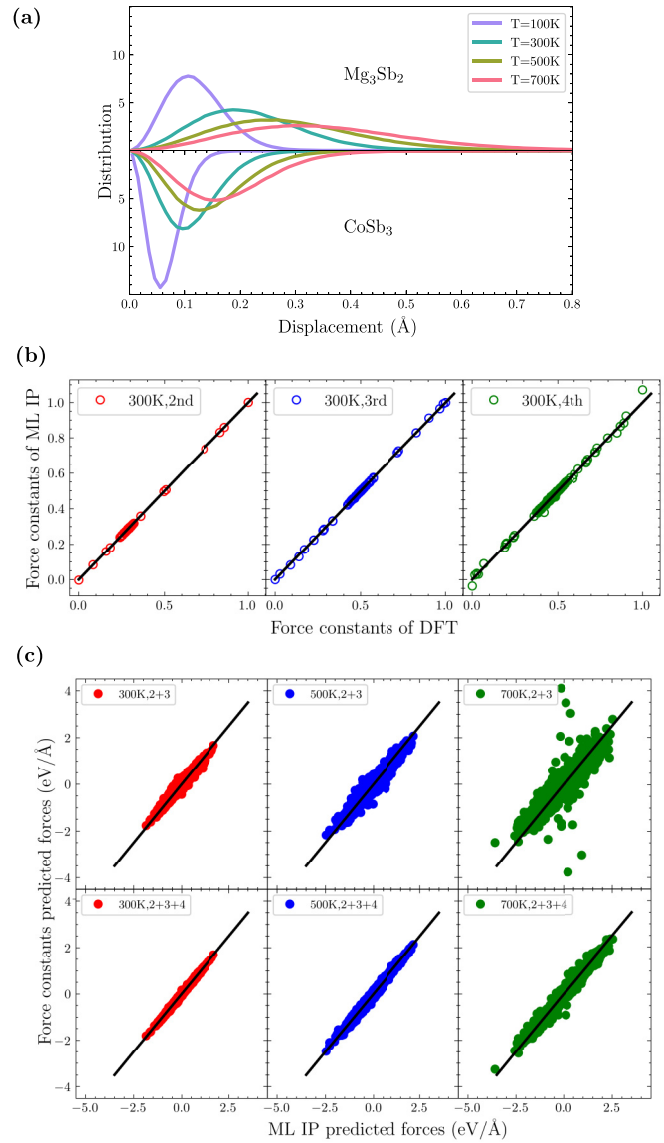

FIG. 4. (a) The distribution of displacements of the atoms away from their equilibrium positions for $CoSb_3$ (lower) and Mg_3Sb_2 (upper). (b) Comparison of the DFT results and the ML-IP results for the effective force constants at 300 K for Mg_3Sb_2. We normalized the force constants of ML-IP and DFT with the maximum and minimum force constants of DFT results. (c) The force contributed by high-order force constants, second- and third-order force constants only [(c) upper], and second-, third- and fourth-order force constants [(c) lower] for Mg_3Sb_2.

explored by the atoms due to the presence of weak chemical bonds poses a greater challenge for sampling, e.g., ML-IP trained with structures sampling at 300 K in Mg_3Sb_2 does not correctly describe the dynamics at 700 K. In DAS, the inner loop explores only a very narrow configuration space, which is not much influenced by the strength of the chemical bond. When the convergence criterion of the inner loop is satisfied, further additional structures in the corresponding sampling block are not likely to further improve the accuracy of the ML-IP. This is also the reason why a complementary

outer loop has to be adopted to cover the configuration spaces corresponding to a wide spectrum of temperatures.

The final on the fly threshold $\bar{a}^{\alpha}$ of each sampling block is also presented in Tables S1 and S2 of the Supplemental Material [50]. It is worth noting that the values of $\bar{a}^{\alpha}$ have a significant difference in different sampling blocks; for example, $\bar{a}^4$ is about 6 times larger than $\bar{a}^1$ for Mg_3Sb_2, indicating the importance of the on the fly thresholds for the convergence criterion of the inner loop. Furthermore, we also found that the $\bar{a}^{\alpha}$'s show material dependence by comparing Tables S1 and S2. Therefore, it is more efficient to use an adaptive $\bar{a}_i^{\alpha}$ rather than a fixed threshold. For example, a low threshold increases the number of selected structures as shown in Table S3 of the Supplemental Material [50]. The DAS together with the on the fly threshold are the key for the training dataset construction in this work.

In traditional perturbation-based thermal transport theory, the accurate prediction of lattice thermal conductivity requires that second- and third-order force constants are accurately predicted. Many previous works have shown that the effective force constants can describe high-order anharmonicity correctly [27,34,67]. However, in order to obtain reasonable high-order effective force constants for complex materials, structures with both small and large displacements need to be carefully considered [34]. In DAS, the improvement of the force constants' accuracy is achieved by systematically reducing the error in the energies and forces in an iterative manner. Different vibration or phonon modes are excited at high temperatures, which contributes to the displacements in MD simulations. The forces on atoms from MD simulations implicitly contain many force constants [67] and the larger displacements make the proportion of high-order force constants higher. If the error of some force constant components is large, the error in forces of corresponding structures will also be large; namely, there is high ambiguity. To validate whether the ML-IP with DAS method can describe the high-order force constants, we employed the least-square fitting method [68] to obtain the effective high-order force constants with forces from ML-IP and DFT calculations, respectively. The agreement between effective force constants that were extracted from the forces of DFT and ML-IP is excellent as shown in Fig. 4(b). The errors in second-, third-, and fourth-order maximum force constants of the ML-IP are 0.17%, 0.12%, and 13.3% for Mg_3Sb_2, respectively. Figure S10 of the Supplemental Material [50] shows the phonon dispersion along high-symmetry paths calculated from the effective second-order force constants of ML-IP and that from DFT for Mg_3Sb_2, respectively. The phonon dispersions are in good agreement except for a tiny difference in the optical branches. The agreement of effective force constants for $CoSb_3$ is also excellent [50].

In Fig. 4(c), we compare the ML-IP forces and corresponding forces contributed by effective force constants for Mg_3Sb_2 at 300, 500, and 700 K, respectively. Here, we show the forces contributed by force constants up to third order [Fig. 4(c), upper] and that with force constants up to fourth order [Fig. 4(c), lower]. The contribution of the fourth-order force constants to the forces increases with increasing temperature and, in particular, the fourth-order force constants are indispensable for reproducing the forces at 700 K. Previous study has shown that due to the influence of the fourth-order force constants to the renormalization of the harmonic phonons, the low-lying transverse acoustic branches become harder with increasing temperatures [27]. We then fitted the forces by considering only the second- and third-order force constants to reproduce the phonon renormalization. Figure S9 of the Supplemental Material [50] plots the phonon dispersions of the second-order force constants from the least-square fitting and frozen phonon method. The hardening of the low-lying transverse acoustic phonon modes at the *M*, *A*, and *L* points is successfully reproduced by our ML-IP method. This demonstrates that the present DAS approach could be used for evaluating the complex thermal transports for the materials with high-order phonon nonlinear scattering.

V. CONCLUSIONS

In summary, we developed a DAS method that significantly reduces the effort to design the training dataset for ML-IP by adaptively updating the ambiguity threshold for selecting structures at a wide range of thermodynamic conditions. The developed sampling method exhibits both good robustness and transferability as demonstrated in the thermal transport for rigid material such as $CoSb_3$ and material with chemical bond hierarchy such as Mg_3Sb_2. The lattice thermal conductivities of $CoSb_3$ and Mg_3Sb_2 calculated by the Green-Kubo method based on ML-IP agree very well with the experimental results. Our method is expected to be applicable in more complex materials such as part-crystalline, part-liquid materials [23,26] and these complex applications will be investigated in future work. Our further analysis shows that the sampling method ensures that high-order effective force constants such as third and fourth order are also highly accurate despite the fact that only energies and forces are used for fitting. Our work provides an avenue to effectively construct the training dataset for ML-IP in complex materials with chemical bond hierarchy and accurately predict thermal transport properties.

ACKNOWLEDGMENTS

This work is supported in part by the National Key Research and Development Program of China (Grants No. 2017YFB0701600, No. 2018YFB0703600, and No. 2019YFA0704901) and the National Natural Science Foundation of China (Grant No. 51632005). W.Z. also acknowledges the support from the Guangdong Innovation Research Team Project (Grant No. 2017ZT07C062), Guangdong Provincial Key-Lab program (Grant No. 2019B030301001), Shenzhen Municipal Key-Lab program (Grant No. ZDSYS20190902092905285), and the Centers for Mechanical Engineering Research and Education at MIT and SUSTech. We acknowledge the computational support provided by the Center for Computational Science and Engineering at Southern University of Science and Technology.

[1] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 2nd ed. (Oxford University Press, Oxford, 2017).
[2] J. Tersoff, Phys. Rev. B **39**, 5566 (1989).
[3] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984).
[4] J. Behler, J. Phys.: Condens. Matter **26**, 183001 (2014).
[5] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[6] A. V. Shapeev, Multiscale Model. Simul. **14**, 1153 (2016).
[7] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018).
[8] N. Artrith and J. Behler, Phys. Rev. B **85**, 045439 (2012).
[9] E. V. Podryabinkin and A. V. Shapeev, Comput. Mater. Sci. **140**, 171 (2017).
[10] L. Zhang, D.-Y. Lin, H. Wang, R. Car, and W. E, Phys. Rev. Mater. **3**, 023804 (2019).
[11] J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitberg, J. Chem. Phys. **148**, 241733 (2018).
[12] A. Singraber, T. Morawietz, J. Behler, and C. Dellago, J. Chem. Theory Comput. **15**, 3075 (2019).
[13] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, npj Comput. Mater. **6**, 20 (2020).
[14] N. Bernstein, G. Csányi, and V. L. Deringer, npj Comput. Mater. **5**, 99 (2019).
[15] K. Gubaev, E. V. Podryabinkin, G. L. W. Hart, and A. V. Shapeev, Comput. Mater. Sci. **156**, 148 (2019).
[16] P. Korotaev, I. Novoselov, A. Yanilkin, and A. Shapeev, Phys. Rev. B **100**, 144308 (2019).
[17] P. Korotaev and A. Shapeev, Phys. Rev. B **102**, 184305 (2020).
[18] R. Li, E. Lee, and T. Luo, Mater. Today Phys. **12**, 100181 (2020).
[19] J. Yang, H.-L. Yip, and A. K.-Y. Jen, Adv. Energy Mater. **3**, 549 (2013).
[20] J. Yang, L. Xi, W. Qiu, L. Wu, X. Shi, L. Chen, J. Yang, W. Zhang, C. Uher, and D. J. Singh, npj Comput. Mater. **2**, 15015 (2016).
[21] T. Zhu, Y. Liu, C. Fu, J. P. Heremans, J. G. Snyder, and X. Zhao, Adv. Mater. **29**, 1605884 (2017).
[22] H. Alam and S. Ramakrishna, Nano Energy **2**, 190 (2013).
[23] W. Qiu, L. Xi, P. Wei, X. Ke, J. Yang, and W. Zhang, Proc. Natl. Acad. Sci. USA **111**, 15031 (2014).
[24] H. Kim, S. Ballikaya, H. Chi, J.-P. Ahn, K. Ahn, C. Uher, and M. Kaviany, Acta Mater. **86**, 247 (2015).
[25] Y. Wang, H. Yang, W. Qiu, J. Yang, J. Yang, and W. Zhang, Phys. Rev. B **98**, 054304 (2018).
[26] J. Yang, Y. Wang, H. Yang, W. Tang, J. Yang, L. Chen, and W. Zhang, J. Phys.: Condens. Matter **31**, 183002 (2019).
[27] Y. Zhu, Y. Xia, Y. Wang, Y. Sheng, J. Yang, C. Fu, A. Li, T. Zhu, J. Luo, C. Wolverton, G. J. Snyder, J. Liu, and W. Zhang, Research **2020**, 4589786 (2020).
[28] T. Feng and X. Ruan, Phys. Rev. B **93**, 045202 (2016).
[29] T. Feng, L. Lindsay, and X. Ruan, Phys. Rev. B **96**, 161201(R) (2017).
[30] T. Feng and X. Ruan, Phys. Rev. B **97**, 045202 (2018).
[31] N. K. Ravichandran and D. Broido, Phys. Rev. X **10**, 021063 (2020).
[32] B. Mortazavi, E. V. Podryabinkin, I. S. Novikov, T. Rabczuk, X. Zhuang, and A. V. Shapeev, Comput. Phys. Commun. **258**, 107583 (2021).
[33] X. Qian, S. Peng, X. Li, Y. Wei, and R. Yang, Mater. Today Phys. **10**, 100140 (2019).
[34] T. Murakami, T. Shiga, T. Hori, K. Esfarjani, and J. Shiomi, EPL **102**, 46002 (2013).
[35] F. Ercolessi and J. B. Adams, EPL **26**, 583 (1994).
[36] A. Krogh and J. Vedelsby, in *Proceedings of the 7th International Conference on Neural Information Processing Systems* (The MIT Press, Cambridge, MA, 1994), pp. 231–238.
[37] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).
[38] C. van der Oord, G. Dusson, G. Csányi, and C. Ortner, Mach. Learn. Sci. Technol. **1**, 015004 (2020).
[39] I. S. Novikov, K. Gubaev, E. V. Podryabinkin, and A. V. Shapeev, Mach. Learn. Sci. Technol. **2**, 025002 (2021).
[40] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, and S. P. Ong, J. Phys. Chem. A **124**, 731 (2020).
[41] W. Li and N. Mingo, Phys. Rev. B **90**, 094302 (2014).
[42] R. Guo, X. Wang, and B. Huang, Sci. Rep. **5**, 7806 (2015).
[43] J. Zhang, L. Song, M. Sist, K. Tolborg, and B. B. Iversen, Nat. Commun. **9**, 4716 (2018).
[44] J. Xin, G. Li, G. Auffermann, H. Borrmann, W. Schnelle, J. Gooth, X. Zhao, T. Zhu, C. Felser, and C. Fu, Mater. Today Phys. **7**, 61 (2018).
[45] https://github.com/hlyang1992/das.
[46] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).
[47] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[48] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[49] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.104.094310 for details of the functional form of the ML-IP, parameter settings for the sampling process, the convergence test with respect to simulation cell size, the effect of the initial dataset and sampling temperature on the developed method, and the phonon dispersions from effective force constants.
[51] M. S. Green, J. Chem. Phys. **22**, 398 (1954).
[52] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
[53] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, Comput. Phys. Commun. **185**, 1747 (2014).
[54] A. J. McGaughey and J. M. Larkin, Annu. Rev. Heat Transfer **17**, 49 (2014).
[55] Z. Fan, L. F. C. Pereira, H.-Q. Wang, J.-C. Zheng, D. Donadio, and A. Harju, Phys. Rev. B **92**, 094301 (2015).
[56] D. A. McQuarrie, *Statistical Mechanics* (HarperCollins College Division, New York, 1976).
[57] C. Carbogno, R. Ramprasad, and M. Scheffler, Phys. Rev. Lett. **118**, 175901 (2017).
[58] Y. Huang, J. Kang, W. A. Goddard, and L.-W. Wang, Phys. Rev. B **99**, 064103 (2019).
[59] R. Hanus, X. Guo, Y. Tang, G. Li, G. J. Snyder, and W. G. Zeier, Chem. Mater. **29**, 1156 (2017).
[60] D. T. Morelli, T. Caillat, J.-P. Fleurial, A. Borshchevsky, J. Vandersande, B. Chen, and C. Uher, Phys. Rev. B **51**, 9622 (1995).
[61] T. Caillat, A. Borshchevsky, and J. -P. Fleurial, J. Appl. Phys. **80**, 4442 (1996).

[62] G. Li, Q. An, W. Li, W. A. Goddard, P. Zhai, Q. Zhang, and G. J. Snyder, Chem. Mater. **27**, 6329 (2015).
[63] M. de Jong, W. Chen, T. Angsten, A. Jain, R. Notestine, A. Gamst, M. Sluiter, C. Krishna Ande, S. van der Zwaag, J. J. Plata, C. Toher, S. Curtarolo, G. Ceder, K. A. Persson, and M. Asta, Sci. Data **2**, 150009 (2015).
[64] A. Bhardwaj and D. K. Misra, RSC Adv. **4**, 34552 (2014).
[65] J. Shuai, Y. Wang, H. S. Kim, Z. Liu, J. Sun, S. Chen, J. Sui, and Z. Ren, Acta Mater. **93**, 187 (2015).
[66] L. Song, J. Zhang, and B. B. Iversen, J. Mater. Chem. A **5**, 4932 (2017).
[67] K. Esfarjani and H. T. Stokes, Phys. Rev. B **77**, 144112 (2008).
[68] F. Eriksson, E. Fransson, and P. Erhart, Adv. Theory Simul. **2**, 1800184 (2019).