



Full Length Article

Accelerated computation of lattice thermal conductivity using neural network interatomic potentials

Jeong Min Choi^{a,1}, Kyeongpung Lee^{a,1}, Sangtae Kim^a, Minseok Moon^a, Wonseok Jeong^{b,*},
Seungwu Han^{a,*}

^a Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Republic of Korea

^b Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94550, USA



ARTICLE INFO

Keywords:

Machine learning potential
Lattice thermal conductivity
Phonon Boltzmann transport equation
Accelerated computation

ABSTRACT

With the development of the density functional theory (DFT) and ever-increasing computational capacity, an accurate prediction of lattice thermal conductivity based on the Boltzmann transport theory becomes computationally feasible, contributing to a fundamental understanding of thermal conductivity as well as a choice of the optimal materials for specific applications. However, steep costs in evaluating third-order force constants limit the theoretical investigation to crystals with high symmetry and few atoms in the unit cell. Currently, machine learning potentials are garnering attention as a computationally efficient high-fidelity model of DFT, and several studies have demonstrated that the lattice thermal conductivity could be computed accurately via the machine learning potentials. However, test materials were mostly crystals with high symmetries, and the applicability of machine learning potentials to a wide range of materials has yet to be demonstrated. Furthermore, establishing a standard training set that provides consistent accuracy and computational efficiencies across a wide range of materials would be useful. To address these issues, herein we compute lattice thermal conductivities at 300 K using neural network interatomic potentials. As test materials, we select 25 materials with diverse symmetries and a wide range of lattice thermal conductivities between 10^{-1} and $10^3 \text{ Wm}^{-1}\text{K}^{-1}$. Among various choices of training sets, we find that molecular dynamics trajectories at 50–700 K consistently provide results at par with DFT for the test materials. In contrast to pure DFT approaches, the computational cost in the present approach is uniform over the test materials, yielding a speed gain of 2–10 folds. When a smaller reduced training set is used, the relative efficiency increases by up to ~ 50 folds without sacrificing accuracy significantly. The current work will broaden the application scope of machine learning potentials by establishing a robust framework for accurately computing lattice thermal conductivity with machine learning potentials.

1. Introduction

The prediction of thermal conductivity is crucial in selecting materials for various applications [1]. For example, materials with low thermal conductivities can be used for thermoelectrics [2–4] or thermal insulations [5], while materials with high thermal conductivities are suitable for the thermal management of electronic devices [6]. Over the last decades, the development of *ab initio* methods combined with increasing computational power enabled a reliable prediction of lattice thermal conductivity (κ_l) [4,7–12]. This is particularly impressive because the computed κ_l spans a wide range of scales from 10^{-1} to $10^3 \text{ Wm}^{-1}\text{K}^{-1}$. However, the *ab initio* evaluation of κ_l becomes expensive in

multicomponent or low-symmetry materials due to high computational costs in obtaining anharmonic terms in the interatomic force constant. For instance, the monoclinic $\beta\text{-Ga}_2\text{O}_3$ requires thousands of single-point density functional theory (DFT) calculations to obtain κ_l . Several approaches that exploit regression techniques have been proposed to save the computational cost, resulting in a significant increase in the computational efficiency [13–19]. One example is the compressive sensing lattice dynamics method, which uses sparsity in the force constants [13,14]. The potential energy surface is Taylor-expanded with the high-order force constants (fourth-order and beyond) in this method, which is fitted to DFT atomic forces for reference structures [14]. The temperature-dependent effective-potential method (TDEP) also expands

* Corresponding authors.

E-mail addresses: jeong3@llnl.gov (W. Jeong), hansw@snu.ac.kr (S. Han).

¹ Jeong Min Choi and Kyeongpung Lee contributed equally to this paper.

the potential energy surface but the series is usually truncated beyond the third order [15–18]. Through a fit to *ab initio* molecular dynamics (AIMD) trajectories, the finite-temperature effects of higher-order terms on the force constant and anharmonicity are incorporated in TDEP. Several high-throughput calculations of κ_l have been reported using semi-empirical approaches based on the Debye-Callaway or Leibfried-Schlömann models [20–25]. However, for materials outside of the dataset used to fit the model, the κ_l prediction accuracy would deteriorate.

Machine-learning interatomic potentials (MLPs) have recently been used as surrogate models of DFT in calculating force constants and κ_l [26–43]. To fit the potential energy surface produced by *ab initio* calculations, preferably DFT [44], MLPs use flexible regression models. For regression models, artificial neural networks [45], kernel-based methods [46], and linear fitting [47] are popular choices. Once trained, MLPs infer the energy, atomic force, and virial stress of the given structure with an accuracy comparable to DFT, but at a fraction of the cost. Until now, several crystals and alloys have been studied, and κ_l values obtained by MLPs have been close to the reference DFT data [26–38,42,43]. Notably, the κ_l of BAs was successfully predicted by MLP, which demands up to four-phonon scattering in the Boltzmann transport equation [38]. MLPs are also advantageous for investigating the κ_l of disordered phases such as amorphous [40] and liquid systems [39]. For materials dominated by nonperturbative phonon scattering, extensive molecular dynamics (MD) simulations are necessary to get an accurate κ_l , which can be handled efficiently by MLP [41].

Although previous works support that the MLP is a powerful tool in calculating κ_l , several issues need to be addressed. For example, most materials in the previous studies retain high symmetries, three or fewer elements, relatively simple configurations, and a limited range of κ_l ($>10 \text{ Wm}^{-1}\text{K}^{-1}$). Consequently, it is unclear whether the MLP's prediction accuracy can be maintained for materials in general. Furthermore, recipes for creating the training set differ significantly across the literature, making it difficult to establish a standard approach. Motivated by these observations, we herein investigate the effect of material complexity and different types of training sets on the accuracy and computational cost of κ_l prediction by MLPs. We adopt Behler-Parrinello-type neural network potential (NNP) as an MLP model [45] and consider three types of training sets: i) snapshots of crystals with randomly displaced atoms [42], ii) AIMD trajectory [26–30,33–39], and iii) snapshots of crystals with atoms displaced along phonon eigenmodes [40]. Our main goal is to develop a recipe for building a training set that predicts the room-temperature (300 K) κ_l of general bulk materials with high efficiency and reasonable accuracy. The following is how the remaining sections are built: the details of computational methods are described in Section 2. In Section 3.1, we conduct a preliminary test on the methods of constructing training sets. In Section 3.2, based on the method chosen in the previous section, we calculate κ_l of 25 materials with diverse cell symmetry and a wide range of κ_l values and analyze relative errors and computational efficiencies. We also check the effect of reducing the size of the training set. Finally, we summarize and conclude in Section 4.

2. Methods

2.1. Density functional theory calculations

All the DFT calculations in the present work are performed using Vienna *Ab initio* Simulation Package (VASP) [48–51]. The generalized gradient approximation by Perdew–Burke–Ernzerhof (PBE) [52] is used for the exchange–correlation functional. The initial structure for each material is obtained from the inorganic crystal structure database (ICSD) [53], which is relaxed further within DFT. The plane-wave cutoff energy and \mathbf{k} -points grids for the unit-cell optimization are selected such that the energy and atomic forces converge to within 1 meV/atom and 5 meV/Å, respectively. The selected parameters are summarized in the [Supplementary Information](#). In addition, the PREC tag is set to

“Accurate”, and the convergence criteria for the self-consistent cycle is set to 10^{-8} eV. Then, the unit cells are fully optimized including lattice vectors until remaining atomic forces become smaller than 1 meV/Å. The final structure becomes the reference one in generating the training set. The computational parameters used for constructing the training data will be discussed in Section 3.1.

2.2. Computation of lattice thermal conductivity

We calculate the lattice thermal conductivity by solving the phonon Boltzmann transport equation linearized under the relaxation-time approximation. In the following, we briefly summarize the equations involved in computing κ_l . Detailed discussions can be found elsewhere [8,9,54,55]. The potential energy surface U can be expanded as

$$U = U_0 + \frac{1}{2!} \sum_{ij\alpha\beta} \Phi_{ij}^{\alpha\beta} u_i^\alpha u_j^\beta + \frac{1}{3!} \sum_{ijk\alpha\beta\gamma} \Phi_{ijk}^{\alpha\beta\gamma} u_i^\alpha u_j^\beta u_k^\gamma + \dots, \quad (1)$$

where U_0 is a constant and $\Phi_{ij}^{\alpha\beta}$ and $\Phi_{ijk}^{\alpha\beta\gamma}$ are the second- and third-order interatomic force constants, respectively. u_i is the atomic displacement of the i th atom from the equilibrium position, and α , β , and γ are the Cartesian indices. At temperature T , the lattice thermal conductivity tensor is expressed as

$$\kappa_{\alpha\beta} = \frac{1}{\Omega} \sum_{\mathbf{q}s} C_{\mathbf{q}s} v_{\mathbf{q}s}^\alpha v_{\mathbf{q}s}^\beta \tau_{\mathbf{q}s}, \quad (2)$$

where Ω is the volume of the crystal and $C_{\mathbf{q}s}$, $v_{\mathbf{q}s}$, and $\tau_{\mathbf{q}s}$ correspond to the heat capacity, phonon velocity, and phonon lifetime with the wave vector \mathbf{q} and branch index s , respectively. In Eq. (2), the heat capacity is defined as

$$C_{\mathbf{q}s} = \hbar \omega_{\mathbf{q}s} \frac{\partial n_{\mathbf{q}s}}{\partial T}, \quad (3)$$

where \hbar is the reduced Planck constant, $\omega_{\mathbf{q}s}$ is the angular frequency of the phonon mode $\mathbf{q}s$, and $n_{\mathbf{q}s}$ is the Bose–Einstein occupation number at temperature T . The phonon lifetime is calculated as the inverse of the total scattering rate

$$\frac{1}{\tau_\lambda} = \sum_{\lambda'\lambda''}^+ \Gamma_{\lambda\lambda'\lambda''}^+ + \frac{1}{2} \sum_{\lambda'\lambda''}^- \Gamma_{\lambda\lambda'\lambda''}^- + \sum_{\lambda'} \Gamma_{\lambda\lambda'}, \quad (4)$$

where Γ^+ and Γ^- are three-phonon scattering rates, corresponding to absorption and emission processes, respectively. Here the phonon mode $\mathbf{q}s$ is abbreviated to a single index λ . The last term on the right-hand side of Eq. (4) corresponds to scattering by isotopes present in nature. The three-phonon scattering rates Γ^+ and Γ^- are expressed as

$$\Gamma_{\lambda\lambda'\lambda''}^+ = \frac{\hbar\pi}{4} \frac{n_{\lambda'} - n_{\lambda''}}{\omega_\lambda \omega_{\lambda'} \omega_{\lambda''}} |V_{\lambda\lambda'\lambda''}^+|^2 \delta(\omega_\lambda + \omega_{\lambda'} - \omega_{\lambda''}), \quad (5)$$

$$\Gamma_{\lambda\lambda'\lambda''}^- = \frac{\hbar\pi}{4} \frac{n_{\lambda'} + n_{\lambda''} + 1}{\omega_\lambda \omega_{\lambda'} \omega_{\lambda''}} |V_{\lambda\lambda'\lambda''}^-|^2 \delta(\omega_\lambda - \omega_{\lambda'} - \omega_{\lambda''}). \quad (6)$$

In Eqs. (5) and (6), the Dirac delta functions δ impose the energy conservation during scattering. The scattering matrix elements $V_{\lambda\lambda'\lambda''}^\pm$ are computed as

$$V_{\lambda\lambda'\lambda''}^\pm = \sum_{ijk} \sum_{\alpha\beta\gamma} \Phi_{ijk}^{\alpha\beta\gamma} \frac{\mathbf{e}_i^\alpha(\lambda) \mathbf{e}_j^\beta(\pm\lambda') \mathbf{e}_k^\gamma(-\lambda'')}{\sqrt{M_i M_j M_k}}, \quad (7)$$

where $\mathbf{e}_i(\lambda)$ and M_i are normalized eigenvectors of mode λ and the atomic mass of the i th atom, respectively. Here $-\lambda$ refers to the phonon mode in the wave vector $-\mathbf{q}$ and branch s . Finally, the summation index i and indices j and k in Eq. (7) indicate atoms in the unit cell and supercell, respectively.

The second-order interatomic force constants $\Phi_{ij}^{\alpha\beta}$ in Eq. (1) are

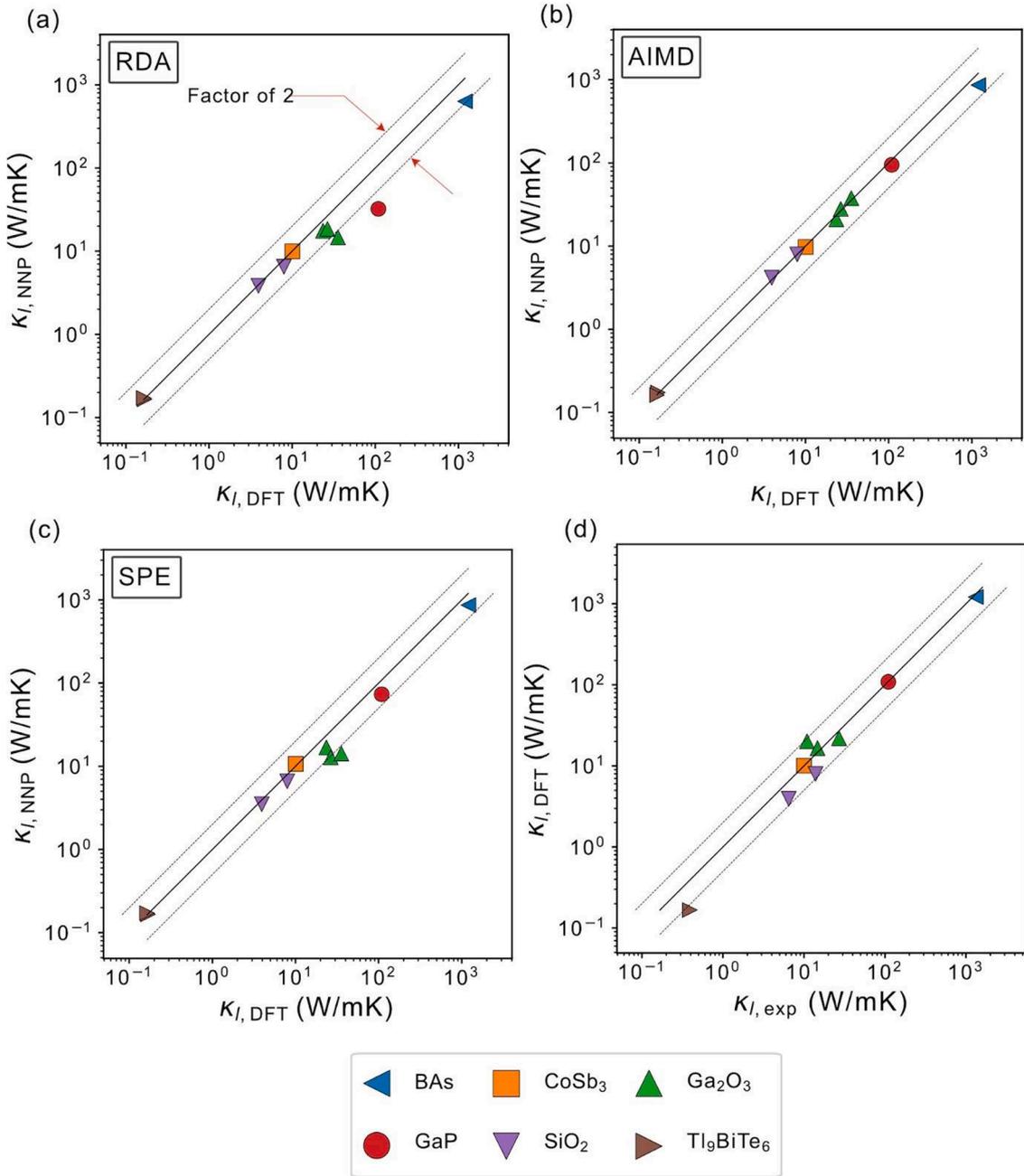


Fig. 1. κ_l of test materials calculated by density functional theory (DFT) and neural network potentials (NNPs) that are generated with various training sets. (a) Random displacements of atoms (RDA), (b) *ab initio* molecular dynamics (AIMD), (c) superposition of phonon eigenmodes (SPE). (d) Comparison between DFT and experimental values. Multiple values for each symbol denote the diagonal components of κ_l , and gray lines denote the error with a factor of 2. The data for experimental κ_l values are from Ref. [64] (BAs), Ref. [65] (CoSb₃), Ref. [66] (β -Ga₂O₃), Ref. [67] (GaP), Ref. [68] (α -SiO₂), and Ref. [69] (Tl₉BiTe₆).

calculated with a finite displacement method using the PHONOPY package [56]. The side lengths of supercells used in the computation are ~ 20 Å each. The third-order interatomic force constants ($\Phi_{ijk}^{\alpha\beta\gamma}$ in Eq. (1)), also calculated with a finite displacement method, and lattice thermal conductivities at 300 K are evaluated by the ShengBTE package [9]. Here the supercell dimension is typically ~ 10 Å with ~ 6 Å for the cutoff radii of the interatomic interaction. The number of single-point force calculations increases with the number of possible atomic pairs in the supercell, considering the cutoff radii of the interatomic interaction. For the \mathbf{q} -point sampling, a uniform mesh grid with a density of ~ 19 points Å⁻¹ is used. The Dirac delta functions in Eqs. (5) and (6) are approximated by the adaptive Gaussian with a proportionality factor of 0.1 [9,57]. Under the relaxation time approximation, the Boltzmann

transport equation is not solved iteratively. Because the current work focuses on comparing the results of DFT and NNP, we do not consider nonanalytic corrections for LO-TO splitting. The full details on the computational parameters used to evaluate force constants can be found in the [Supplementary Information](#).

2.3. Neural network potential

For training NNPs, we employ the SIMPLE-NN package [58]. Atom-centered symmetry functions G^2 and G^4 [59] are adopted to describe radial and angular distributions of neighboring atoms, respectively, defined as follows.

$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_c)^2} \cdot f_c(R_{ij}), \quad (8)$$

$$G_i^4 = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos \theta_{ijk})^\zeta \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}), \quad (9)$$

where i, j , and k are atomic indices, R_{ij} is the distance between atom i and j , and η, ζ , and λ , are hyperparameters. f_c is the cutoff function such that atoms outside the cutoff radius R_c do not contribute to the atomic energy.

$$f_c(R_{ij}) = \begin{cases} 0.5 \cdot \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1 \right] & (R_{ij} \leq R_c) \\ 0 & (R_{ij} > R_c). \end{cases} \quad (10)$$

The cutoff radius is set to 6.5 Å, and 26, 70, and 132 symmetry functions are employed for unary, binary, and ternary systems, respectively. The selected hyperparameters [60] are summarized in the [Supplementary Information](#). The network architecture comprises two hidden layers with 60 hidden nodes each and one output layer that provides atomic energy. The input vector is decorrelated by principal component analysis and then whitened to increase the learning speed [60]. The NNP is trained by minimizing the loss function (Γ) defined as follows.

$$\Gamma = \frac{1}{M} \sum_{i=1}^M \left(\frac{E_i^{\text{DFT}} - E_i^{\text{NNP}}}{N_i} \right)^2 + \frac{\mu_1}{3 \sum_{i=1}^M N_i} \sum_{i=1}^M \sum_{j=1}^{N_i} \left| \mathbf{F}_{ij}^{\text{DFT}} - \mathbf{F}_{ij}^{\text{NNP}} \right|^2 + \frac{\mu_2}{6M} \sum_{i=1}^M \times \sum_{k=1}^6 \left| S_{ik}^{\text{DFT}} - S_{ik}^{\text{NNP}} \right|^2, \quad (11)$$

where M is the total number of structures in the training set, N_i is the number of atoms in the i th structure, and $E_i^{\text{DFT(NNP)}}$, $\mathbf{F}_{ij}^{\text{DFT(NNP)}}$, and $S_{ik}^{\text{DFT(NNP)}}$ are its total energy, atomic force of the j th atom, and the k th component of the virial stress tensor, respectively. μ_1 and μ_2 are the parameters that scale the relative importance of atomic force and stress with respect to the total energy when minimizing the loss function. We set the parameters μ_1 and μ_2 as 10^2 and 10^{-6} , respectively. To avoid overfitting and obtain more regularized NNPs, we apply a dropout technique in which half of the nodes are randomly selected and fixed at each training iteration [61]. The learning rate starts at 0.01 and scales by 0.1 every 5000 iterations. We train two independent NNPs and use averaged forces to obtain interatomic force constants for given atomic displacements. For evaluating atomic forces, we use the LAMMPS package [62].

3. Results

3.1. Construction of training sets

To select an optimal approach to generate training sets, preliminary tests are conducted on six materials: BAs, CoSb₃, β -Ga₂O₃, GaP, α -SiO₂, and Tl₉BiTe₆ that span three orders of κ_l from 10^{-1} to $10^2 \text{ Wm}^{-1}\text{K}^{-1}$. We compare the three methods of constructing the training set: (i) random displacements of atoms (RDA), (ii) AIMD, and (iii) superposition of phonon eigenmodes (SPE). Each atom is displaced with random directions and amplitudes in the RDA method. The amplitudes are chosen from a normal distribution with a standard deviation of 0.1 Å. The AIMD method comprises two steps. First, the AIMD using a canonical ensemble is conducted over 1 ps with the time step of 1 fs, at temperatures of 50, 300, 500, and 700 K, following the choice of the simulation time and temperatures in Ref. [30]. (We also tested the AIMD method with NPT ensemble, but the resulting κ_l showed no significant difference.) The simulation cell, which contains ~ 100 atoms, is identical to the one used to calculate the third-order interatomic force constants for the most

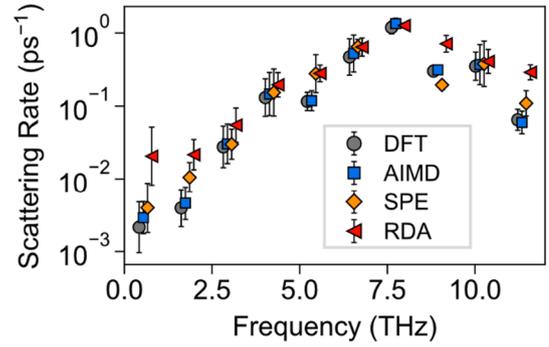


Fig. 2. The frequency-dependent phonon scattering rates of GaP obtained by DFT and NNPs. Scattering rates are averaged over every 1.2 THz intervals, and error bars indicate standard deviation within the interval.

cases of the test materials (see [Section 2.2](#), and [Supplementary Information](#)). During AIMD, the computational parameters are slightly loosened such that the self-consistency criteria is set to 10^{-4} eV with the default plane-wave energy cutoff and the Γ -point sampling. Next, we sample the 4-ps AIMD trajectories in 10- or 80-fs intervals depending on the target size of the training set, and more accurate DFT calculations are performed on the sampled structures. This is required because the interatomic force constants are sensitive to computational precision, so the training data must maintain high precision. In these calculations, the plane-wave energy cutoff, k-point grids, and convergence criteria of self-consistent calculation are set to the same as those of the unit cell optimization (See [Section 2.1](#)). Finally, the SPE method requires the second-order interatomic force constants that determine the phonon eigenmodes. To roughly calculate the second-order interatomic force constants with minimal computational costs, we use the Γ point and a small supercell size adopted in calculating third-order interatomic force constants (see [Section 2.2](#)). Each atom is displaced along the superposed phonon modes with random amplitudes and phase factors [19,63].

In each of the three approaches described above, the training set comprises 400 structures. To calculate third-order interatomic force constants, interactions up to ~ 6 Å or 6th, 7th, 19th, 5th, 11th, and 4th nearest neighbors are considered for BAs, CoSb₃, β -Ga₂O₃, GaP, α -SiO₂, and Tl₉BiTe₆, respectively. For the phonon dispersion, every NNP trained by the three methods reproduces phonon dispersions for all six materials (see [Fig. S1](#) in the [Supplementary Information](#)). [Fig. 1\(a\)–\(c\)](#) present the comparisons of the diagonal components of the κ_l tensor between NNP and DFT. We also compare the DFT results with experimental data in [Fig. 1\(d\)](#) to validate computational settings. It is seen that NNP-AIMD shows higher accuracy than other methods, with the error for all test materials being less than 30%. Conversely, NNP-RDA and NNP-SPE show error levels larger than a factor of 2 for GaP and β -Ga₂O₃, respectively. This means that RDA and SPE method may produce atomic displacements that have no bearing on third-order interatomic force constants. Since the RDA method does not make any assumptions about atomic correlations, it is more likely to produce unphysical atomic configurations with small interatomic distances and large repulsive atomic forces. This analysis is supported by the phonon scattering rates of GaP shown in [Fig. 2](#); while NNP-AIMD shows good accuracy in predicting frequency-scattering rates relations, the other two NNPs show large discrepancies, especially NNP-RDA.

The computational cost of the SPE method varies greatly depending on the cell symmetry due to the part on obtaining phonon dispersion. The costs of the AIMD and RDA methods, however, are less affected by crystal symmetry. Consequently, we conclude that the AIMD method is the best for generating training data. In passing, we note that for BAs, the present DFT result of κ_l (~ 1200) in [Fig. 1](#) is at variance with the previous literature (>2000) [30,70]. This is because we did not solve the Boltzmann transport equation iteratively (i.e., relaxation time approximation), which is known to affect the prediction accuracy for materials

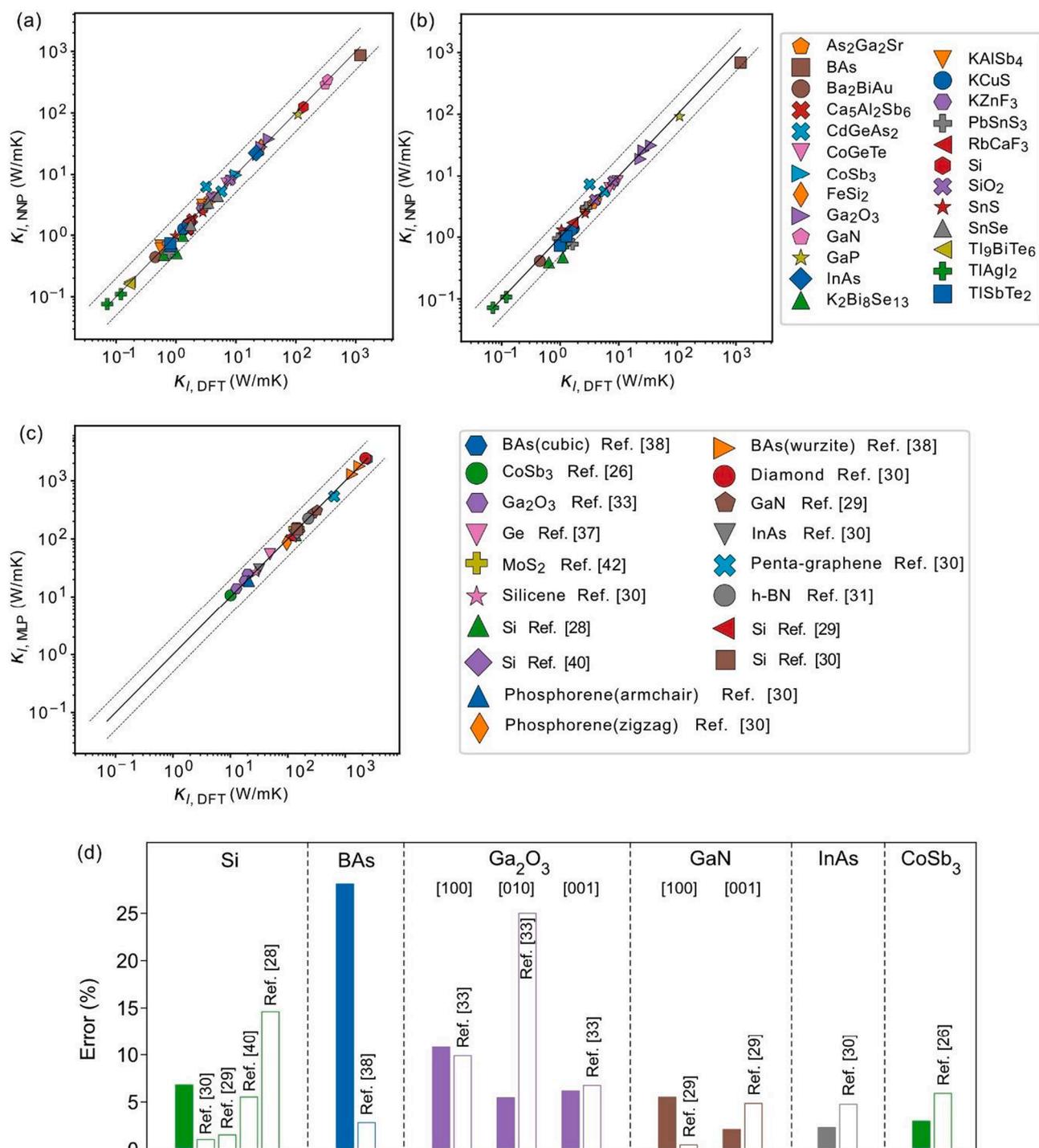


Fig. 3. κ_l computed by NNP-AIMD trained with (a) 400 and (b) 50 structures sampled from the same AIMD trajectories. (c) DFT and machine learning potentials (MLP) results quoted from the literature. (d) Comparison of errors in κ_l prediction by MLPs for the materials common in (a) and (c). The filled and empty bars are the present results and references, respectively.

with high κ_l [9]. When the Boltzmann transport equation is solved iteratively, we find that κ_l increases to ~ 1900 .

3.2. Computation of κ_l for diverse materials.

This section increases the benchmark cases to 25 materials with diverse cell symmetries and a wide range of κ_l values. The training set for each NNP comprises 400 structures generated by the AIMD method. Fig. 3(a) compares κ_l between NNP-AIMD and DFT. The root-mean-squared relative error (RMSRE) of test materials is 18.6%. Fig. 3(b)

shows the results when the training data is reduced, as discussed below. For comparison, we gather the κ_l values computed by MLP and DFT from the literature and plot them in Fig. 3(c) (For a fair comparison, only those obtained by solving the Boltzmann transport equation are presented). Except for a few materials, the error level in Fig. 3(a) and Fig. 3(c) is similar. Fig. 3(d) explicitly compares the error values for common materials in Fig. 3(a) and Fig. 3(c). It can be seen that the errors from the present approach are comparable to those of other references even though we use a consistent choice of the training set. However, BAs shows a much larger error than Ref. [38]. It is known that κ_l of BAs is

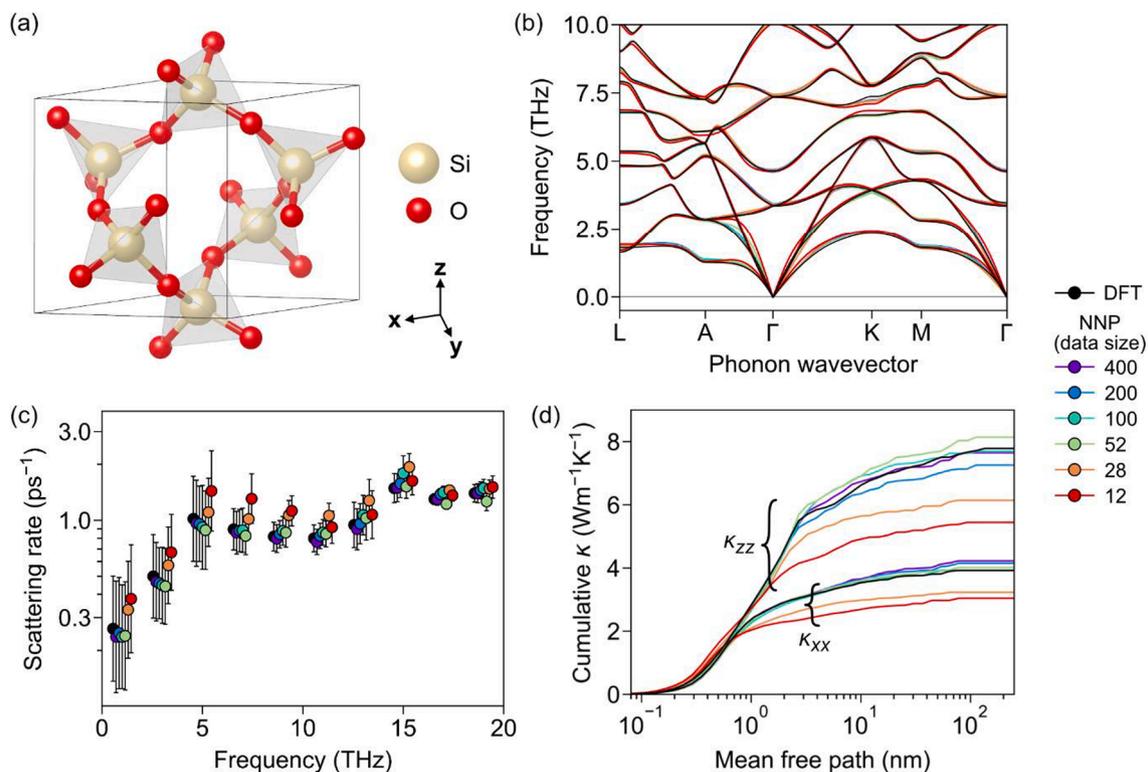


Fig. 4. (a) The unit cell of α -SiO₂ and data-size dependence of (b) phonon dispersion, (c) scattering rate, and (d) cumulative lattice thermal conductivity as a function of the phonon mean free path. The phonon dispersions are presented up to 10 THz for clarity. Scattering rates are averaged over every 2 THz intervals, and error bars indicate standard deviation within the interval. Atomic structures are visualized with VESTA [71].

mainly determined by the scattering of acoustic phonon modes. Ref. [38] employed NPT simulations and adopted DFT results for the harmonic part, which may have contributed to more accurate scatterings by acoustic phonons.

We note that materials with low symmetries often exhibit large errors in Fig. 3(a). To be specific, we classify the test materials into four groups based on the crystal system: (1) cubic, (2) tetragonal and hexagonal, (3) orthorhombic, and (4) monoclinic and triclinic. For each material, we first select the component of κ_l showing the largest error. For the selected components, we calculate the average error within each group. The results are 7.6%, 13.5%, 15.7%, and 24.2% for the group (1)-(4), respectively, which shows a trend of increasing errors with low crystal symmetries. The low symmetry may require longer AIMD simulations for sufficient sampling. To test this, we extend AIMD of K₂Bi₈Se₁₃, which has $\bar{P}1$ space group and thus has the largest error, up to 2 ps, increasing the training set by two folds. However, the resulting NNP for K₂Bi₈Se₁₃ produces almost identical κ_l . Therefore, the origin of the error and a systematic solution need further investigation in future.

The high-precision DFT calculations used to construct the training set account for about 80% of the computational costs in the present work. Consequently, it would be worthwhile to see if the training set could be shrunk further to save on computational costs. To this end, we investigate the effect of dataset size on κ_l using α -SiO₂ as an example (see Fig. 4(a)). First, we choose 12, 28, 52, 100, 200, and 400 structures from 108-atom AIMD data, which train six different NNPs. The simulation time and temperatures are the same as those in the previous section. In Fig. 4(b)-(d), we calculate phonon dispersions, scattering rates, and cumulative lattice thermal conductivities as a function of the phonon mean free path. Fig. 4(b) shows that harmonic properties are well described by every NNP, but a slight deviation of the transverse acoustic mode is observed along the Γ -A line with the data size of 12. The scattering rates and cumulative κ_l in Fig. 4(c) and 4(d), respectively, are well reproduced for data sizes exceeding 28. With data sizes of 12 or 28,

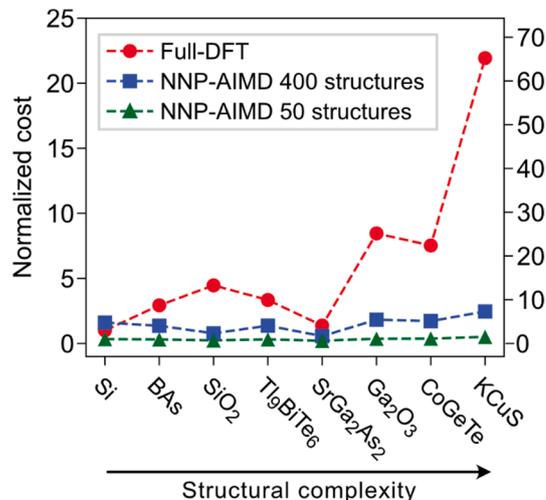


Fig. 5. The computational cost of calculating κ_l by DFT and NNPs trained by 400 or 50 structures. The computational cost is normalized with respect to Si computed within the fully DFT approach (left scale) or the NNP trained with 50 structures (right scale).

significant deviations in κ_l are observed mostly at mean free paths of 1–10 nm.

Following the above analysis, we reduce the training set to 50 structures for 18 materials selected from Fig. 3(a) by sampling more sparsely over the same MD trajectories. The κ_l values compared between the resulting NNPs and DFT are shown in Fig. 3(b). The RMSRE of the κ_l is 31%, increasing from 18% in Fig. 3(a), which is reasonable considering the drastic decrease in computational cost. Again, K₂Bi₈Se₁₃ shows the largest error as in Fig. 3(a), which can be understood by the low

symmetry in the crystal structure.

Additionally, we analyze the computational efficiency of the NNP-based calculation of κ_l with respect to the DFT-based calculation. Fig. 5 shows how the computational cost changes as the material complexity increases. The number of structures required to calculate the third-order force constants, which encompasses the number of atomic elements, the number of atoms in the unit cell, and the crystal symmetry, is used to rank the material complexity. The computational cost is normalized with respect to that of Si computed using the full DFT approach (left scale) or the NNP trained with 50 structures. (right scale). Here, the computational cost of the DFT-based κ_l calculation is the summation of those obtaining second- and third-order force constants. The majority of the computational costs in NNP-based calculations come from building the training set (AIMD and single-point calculations) and training NNPs, while the cost of calculating force constants using the finite displacement method is negligible. As the material complexity increases, the number of structures needed to generate the third-order force constants increases rapidly, which is confirmed in Fig. 5. For instance, the DFT computational cost for KCuS is more than 20 times higher than that of Si. Conversely, NNP-based κ_l calculations show nearly constant computational costs regardless of the material complexity. This is because the cost of AIMD and high-precision DFT calculations are similar among the materials and the atomic forces can be calculated with almost no cost once the NNP has been trained, even for a very large supercell.

To check the applicability for the metallic materials, we additionally calculate the κ_l of Al with the same procedure as in the above. The results calculated by DFT and NNP trained with 400 and 50 training sets are 10.2, 10.6, and 9.9 $\text{Wm}^{-1}\text{K}^{-1}$, respectively, and the corresponding errors are 3.9%, and 2.9%. In addition, the κ_l result of a narrow-gap material InAs, which PBE functional misclassifies as a metal, is given with the error of 2.3% in Fig. 3(a). Both results of metallic, and narrow-gap materials show that the present method can be applied for a wide range of materials with reasonable errors. However, for the metallic materials, electron-phonon scatterings may have significant effects on the κ_l [72], while our method only considers phonon-phonon scatterings. Lastly, special care is needed when calculating the κ_l of materials with high anharmonicity or κ_l at high temperatures, because we did not consider the phonon renormalization and higher-order phonon processes than third order.

4. Conclusion

In conclusion, we propose a standard protocol for building the training set of NNPs targeted for computing κ_l efficiently without fine-tuning for each material. The protocol requires 1-ps AIMD simulations at various temperatures and accurate single-point calculations for 400 structures sampled along the MD trajectory. Testing over 25 materials with diverse symmetries and wide range of κ_l , it is confirmed that NNP-AIMD provides consistent accuracies comparable to reported values in the literature. The uniform cost across material types for the proposed method makes it especially efficient for complex materials whose κ_l prediction would be costly in a full-DFT approach. Furthermore, the NNPs showed reasonable accuracy even when the training set was reduced to 50 structures. The current work will broaden the application scope of machine learning potentials by establishing a robust framework for accurately computing lattice thermal conductivity with machine learning potentials.

Data Availability

The entire data set of the lattice thermal conductivity calculated by DFT and NNP are available in Table S1 of the [Supplementary Information](#).

CRedit authorship contribution statement

Jeong Min Choi: Data curation, Formal analysis, Writing – original draft. **Kyeongpung Lee:** Data curation, Formal analysis, Writing – original draft. **Sangtae Kim:** Investigation. **Minseok Moon:** Validation. **Wonseok Jeong:** Supervision, Writing – review & editing. **Seungwu Han:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported Korea Institute of Ceramic Engineering and Technology (KICET) (N0002599), and Creative Materials Discovery Program (RIAM) through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (2017M3D1A1040689). The computations were carried out at Korea Institute of Science and Technology Information (KISTI) National Supercomputing Center (KSC-2020-CRE-0296). Part of the work is performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract Number DE-AC52-07NA27344.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.commatsci.2022.111472>.

References

- [1] N. Burger, A. Laachachi, M. Ferriol, M. Lutz, V. Toniazzo, D. Ruch, Review of thermal conductivity in composites: mechanisms, parameters and theory, *Prog. Polym. Sci.* 61 (2016) 1–28.
- [2] M.N. Hasan, H. Wahid, N. Nayan, M.S. Mohamed Ali, Inorganic thermoelectric materials: a review, *Int. J. Energy Res.* 44 (8) (2020) 6170–6222.
- [3] J. Wei, L. Yang, Z. Ma, P. Song, M. Zhang, J. Ma, F. Yang, X. Wang, Review of current high-ZT thermoelectric materials, *J. Mater. Sci.* 55 (27) (2020) 12642–12704.
- [4] P. Gorai, V. Stevanović, E.S. Toberer, Computationally guided discovery of thermoelectric materials, *Nat. Rev. Mater.* 2 (2017) 17053.
- [5] D.R. Clarke, S.R. Phillpot, Thermal barrier coating materials, *Mater. Today* 8 (6) (2005) 22–29.
- [6] X.C. Tong, *Advanced materials for thermal management of electronic packaging*, Springer Science & Business Media, 2011.
- [7] X. Qian, J. Zhou, G. Chen, Phonon-engineered extreme thermal conductivity materials, *Nat. Mater.* 20 (9) (2021) 1188–1202.
- [8] A. Togo, L. Chaput, I. Tanaka, Distributions of phonon lifetimes in Brillouin zones, *Phys. Rev. B* 91 (2015), 094306.
- [9] W. Li, J. Carrete, N.A. Katcho, N. Mingo, ShengBTE: a solver of the Boltzmann transport equation for phonons, *Comput. Phys. Commun.* 185 (2014) 1747–1758.
- [10] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, I. Tanaka, Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization, *Phys. Rev. Lett.* 115 (2015), 205901.
- [11] A. van Roekeghem, J. Carrete, C. Oses, S. Curtarolo, N. Mingo, High-throughput computation of thermal conductivity of high-temperature solid phases: the case of oxide and fluoride perovskites, *Phys. Rev. X* 6 (2016), 041061.
- [12] J. Carrete, W. Li, N. Mingo, S. Wang, S. Curtarolo, Finding unprecedentedly low-thermal conductivity half-Heusler semiconductors via high-throughput materials modeling, *Phys. Rev. X* 4 (2014), 011019.
- [13] F. Zhou, W. Nielson, Y. Xia, V. Ozoliņš, Lattice anharmonicity and thermal conductivity from compressive sensing of first-principles calculations, *Phys. Rev. Lett.* 113 (2014), 185501.
- [14] F. Zhou, W. Nielson, Y. Xia, V. Ozoliņš, Compressive sensing lattice dynamics. I. General formalism, *Phys. Rev. B* 100 (2019), 184308.
- [15] O. Hellman, I.A. Abrikosov, S.I. Simak, Lattice dynamics of anharmonic solids from first principles, *Phys. Rev. B* 84 (2011), 180301.
- [16] O. Hellman, P. Steneteg, I.A. Abrikosov, S.I. Simak, Temperature dependent effective potential method for accurate free energy calculations of solids, *Phys. Rev. B* 87 (2013), 104111.
- [17] O. Hellman, D.A. Broido, Phonon thermal transport in Bi_2Te_3 from first principles, *Phys. Rev. B* 90 (2014), 134309.

- [18] F. Bottin, J. Bieder, J. Bouchet, a-TDEP: temperature dependent effective potential for ABINIT – lattice dynamic properties including anharmonicity, *Comput. Phys. Commun.* 254 (2020), 107301.
- [19] F. Eriksson, E. Fransson, P. Erhart, The hiphive package for the extraction of high-order force constants by machine learning, *Adv. Theory Simul.* 2 (2019) 1800184.
- [20] J. Yan, P. Gorai, B. Ortiz, S. Miller, S.A. Barnett, T. Mason, V. Stevanović, E. S. Toberer, Material descriptors for predicting thermoelectric performance, *Energy Environ. Sci.* 8 (3) (2015) 983–994.
- [21] S.A. Miller, P. Gorai, B.R. Ortiz, A. Goyal, D. Gao, S.A. Barnett, T.O. Mason, G. J. Snyder, Q. Lv, V. Stevanović, E.S. Toberer, Capturing anharmonicity in a lattice thermal conductivity model for high-throughput predictions, *Chem. Mater.* 29 (6) (2017) 2494–2501.
- [22] R. McKinney, P. Gorai, E.S. Toberer, V. Stevanovic, Rapid prediction of anisotropic lattice thermal conductivity: application to layered materials, *Chem. Mater.* 31 (2019) 2048–2057.
- [23] C. Toher, J.J. Plata, O. Levy, M. de Jong, M. Asta, M.B. Nardelli, S. Curtarolo, High-throughput computational screening of thermal conductivity, Debye temperature, and Grüneisen parameter using a quasiharmonic Debye model, *Phys. Rev. B* 90 (2014), 174107.
- [24] C. Toher, C. Oses, J.J. Plata, D. Hicks, F. Rose, O. Levy, M. de Jong, M. Asta, M. Fornari, M.B. Nardelli, S. Curtarolo, Combining the AFLOW GIBBS and elastic libraries to efficiently and robustly screen thermomechanical properties of solids, *Phys. Rev. Mater.* 1 (2017), 015401.
- [25] H. Xie, J. Yan, X. Gu, H. Bao, A scattering rate model for accelerated evaluation of lattice thermal conductivity bypassing anharmonic force constants, *J. Appl. Phys.* 125 (2019), 205104.
- [26] P. Korotaev, I. Novoselov, A. Yanilkin, A. Shapeev, Accessing thermal conductivity of complex compounds by machine learning interatomic potentials, *Phys. Rev. B* 100 (2019), 144308.
- [27] C. Zhang, Q. Sun, Gaussian approximation potential for studying the thermal conductivity of silicene, *J. Appl. Phys.* 126 (2019), 105103.
- [28] H. Babaei, R. Guo, A. Hashemi, S. Lee, Machine-learning-based interatomic potential for phonon transport in perfect crystalline Si and crystalline Si with vacancies, *Phys. Rev. Mater.* 3 (2019), 074603.
- [29] E. Minamitani, M. Ogura, S. Watanabe, Simulating lattice thermal conductivity in semiconducting materials using high-dimensional neural network potential, *Appl. Phys. Express* 12 (2019), 095001.
- [30] B. Mortazavi, E.V. Podryabinkin, I.S. Novikov, T. Rabczuk, X. Zhuang, A. V. Shapeev, Accelerating first-principles estimation of thermal conductivity by machine-learning interatomic potentials: a MTP/ShengBTE solution, *Comput. Phys. Commun.* 258 (2021), 107583, <https://doi.org/10.1016/j.cpc.2020.107583>.
- [31] Y. Zhang, C. Shen, T. Long, H. Zhang, Thermal conductivity of h-BN monolayers using machine learning interatomic potential, *J. Phys. Condens. Matter.* 33 (2021), 105903.
- [32] R. Li, Z. Liu, A. Rohskopf, K. Gordiz, A. Henry, E. Lee, T. Luo, A deep neural network interatomic potential for studying thermal conductivity of β -Ga₂O₃, *Appl. Phys. Lett.* 117 (2020), 152102.
- [33] Y.B. Liu, J.Y. Yang, G.M. Xin, L.H. Liu, G. Csányi, B.Y. Cao, Machine learning interatomic potential developed for molecular simulations on thermal properties of β -Ga₂O₃, *J. Chem. Phys.* 153 (2020), 144501.
- [34] L. Han, X. Chen, Q. Wang, Y. Chen, M. Xu, L. Wu, C. Chen, P. Lu, P. Guan, Neural network potential for studying the thermal conductivity of Sn, *Comput. Mater. Sci.* 200 (2021), 110829.
- [35] H. Liu, X. Qian, H. Bao, C.Y. Zhao, X. Gu, High-temperature phonon transport properties of SnSe from machine-learning interatomic potential, *J. Phys. Condens. Matter.* 33 (2021), 405401.
- [36] C. Verdi, F. Karsai, P. Liu, R. Jinnouchi, G. Kresse, Thermal transport and phase transitions of zirconia by on-the-fly machine-learned interatomic potentials, *npj Comput. Mater.* 7 (2021) 156.
- [37] C. Mangold, S. Chen, G. Barbalinardo, J. Behler, P. Pochet, K. Termentzidis, Y. Han, L. Chaput, D. Lacroix, D. Donadio, Transferability of neural network potentials for varying stoichiometry: phonons and thermal conductivity of Mn_xGe_y compounds, *J. Appl. Phys.* 127 (2020), 244901.
- [38] Z. Liu, X. Yang, B.o. Zhang, W.u. Li, High thermal conductivity of wurtzite boron arsenide predicted by including four-phonon scattering with machine learning potential, *ACS Appl. Mater. Interfaces* 13 (45) (2021) 53409–53415.
- [39] J. Deng, L. Stixrude, Thermal conductivity of silicate liquid determined by machine learning potentials, *Geophys. Res. Lett.* 48 (2021), e2021GL093806, <https://doi.org/10.1029/2021GL093806>.
- [40] X. Qian, S. Peng, X. Li, Y. Wei, R. Yang, Thermal conductivity modeling using machine learning potentials: application to crystalline and amorphous silicon, *Mater. Today Phys.* 10 (2019), 100140.
- [41] Z. Zeng, C. Zhang, Y. Xia, Z. Fan, C. Wolverton, Y. Chen, Nonperturbative phonon scatterings and the two-channel thermal transport in Ti₃VSe₄, *Phys. Rev. B* 103 (2021), 224307.
- [42] X. Gu, C.Y. Zhao, Thermal conductivity of single-layer MoS_{2(1-x)Se_{2x}} alloys from molecular dynamics simulations with a machine-learning-based interatomic potential, *Comput. Mater. Sci.* 165 (2019) 74–81.
- [43] J. George, G. Hautier, A.P. Bartók, G. Csányi, V.L. Deringer, Combining phonon accuracy with high transferability in Gaussian approximation potential models, *J. Chem. Phys.* 153 (2020), 044104.
- [44] S. Watanabe, W. Li, W. Jeong, D. Lee, K. Shimizu, E. Mimanitani, Y. Ando, S. Han, High-dimensional neural network atomic potentials for examining energy materials: some recent simulations, *J. Phys. Energy* 3 (2020), 012003, <https://doi.org/10.1088/2515-7655/abc7f3>.
- [45] J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, *Phys. Rev. Lett.* 98 (2007), 146401.
- [46] A.P. Bartók, M.C. Payne, R. Kondor, G. Csányi, Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons, *Phys. Rev. Lett.* 104 (2010), 136403.
- [47] A.V. Shapeev, Moment tensor potentials: a class of systematically improvable interatomic potentials, *Multiscale Model. Simul.* 14 (3) (2016) 1153–1173.
- [48] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* 6 (1) (1996) 15–50.
- [49] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* 54 (16) (1996) 11169–11186.
- [50] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* 59 (3) (1999) 1758–1775.
- [51] G. Kresse, Ab initio molecular dynamics for liquid metals, *J. Non-Cryst. Solids* 192–193 (1995) 222–229.
- [52] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* 77 (18) (1996) 3865–3868.
- [53] M. Hellenbrandt, The Inorganic Crystal Structure Database (ICSD) - present and future, *Crystallogr. Rev.* 10 (2004) 17–22.
- [54] J.M. Ziman, P.W. Levy, Electrons and phonons; the theory of transport phenomena in solids, *Physics Today* 14 (11) (1961) 64–66.
- [55] G.P. Srivastava, The physics of phonons, CRC Press, 1990.
- [56] A. Togo, I. Tanaka, First principles phonon calculations in materials science, *Scr. Mater.* 108 (2015) 1–5.
- [57] W. Li, N. Mingo, L. Lindsay, D.A. Broido, D.A. Stewart, N.A. Katcho, Thermal conductivity of diamond nanowires from first principles, *Phys. Rev. B* 85 (2012), 195436.
- [58] K. Lee, D. Yoo, W. Jeong, S. Han, SIMPLE-NN: an efficient package for training and executing neural-network interatomic potentials, *Comput. Phys. Commun.* 242 (2019) 95–103.
- [59] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, *J. Chem. Phys.* 134 (2011), 074106.
- [60] D. Lee, K. Lee, D. Yoo, W. Jeong, S. Han, Crystallization of amorphous GeTe simulated by neural network potential addressing medium-range order, *Comput. Mater. Sci.* 181 (2020), 109725.
- [61] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* 15 (2014) 1929–1958.
- [62] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P. S. Crozier, P.J. in 't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trit, S.J. Plimpton, *Comput. Phys. Comm.* 271 (2022) 108177, <https://doi.org/10.1016/j.cpc.2021.108171>.
- [63] D. West, S.K. Estreicher, First-principles calculations of vibrational lifetimes and decay channels: hydrogen-related modes in Si, *Phys. Rev. Lett.* 96 (2006), 115504.
- [64] F. Tian, B. Song, X.i. Chen, N.K. Ravichandran, Y. Lv, K.e. Chen, S. Sullivan, J. Kim, Y. Zhou, T.-H. Liu, M. Goni, Z. Ding, J. Sun, G.A.G. Udalamatta Gamage, H. Sun, H. Ziyae, S. Huyan, L. Deng, J. Zhou, A.J. Schmidt, S. Chen, C.-W. Chu, P. Y. Huang, D. Broido, L.i. Shi, G. Chen, Z. Ren, Unusual high thermal conductivity in boron arsenide bulk crystals, *Science* 361 (6402) (2018) 582–585.
- [65] T. Caillat, A. Borschchevsky, J.-P. Fleurial, Properties of single crystalline semiconducting CoSb₃, *J. Appl. Phys.* 80 (8) (1996) 4442–4449.
- [66] Z. Guo, A. Verma, X. Wu, F. Sun, A. Hickman, T. Masui, A. Kuramata, M. Higashiwaki, D. Jena, T. Luo, Anisotropic thermal conductivity in single crystal β -gallium oxide, *Appl. Phys. Lett.* 106 (2015), 111909.
- [67] N. Dural, M.V. Romalis, Gallium phosphide as a new material for anodically bonded atomic sensors, *APL Mater.* 2 (2014), 086101.
- [68] H. Kanamori, N. Fujii, H. Mizutani, Thermal diffusivity measurement of rock-forming minerals from 300° to 1100°K, *J. Geophys. Res.* 73 (2) (1968) 595–605.
- [69] B. Wölfing, C. Kloc, J. Teubner, E. Bucher, High performance thermoelectric Ti₉BiTe₆ with an extremely low thermal conductivity, *Phys. Rev. Lett.* 86 (2001) 4350–4353.
- [70] L. Lindsay, D.A. Broido, T.L. Reinecke, First-principles determination of ultrahigh thermal conductivity of boron arsenide: a competitor for diamond? *Phys. Rev. Lett.* 111 (2013), 025901.
- [71] K. Momma, F. Izumi, VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data, *J. Appl. Crystallogr.* 44 (2011) 1272–1276.
- [72] Y. Wang, Z. Lu, X. Ruan, First principles calculation of lattice thermal conductivity of metals considering phonon-phonon and phonon-electron scattering, *J. Appl. Phys.* 119 (2016), 225109.