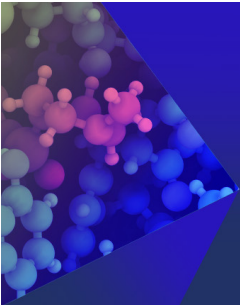# Correcting force error-induced underestimation of lattice thermal conductivity in machine learning molecular dynamics FREE

Xiguang Wu ; Wenjiang Zhou ; Haikuan Dong ; Penghua Ying ; Yanzhou Wang; Bai Song ;
Zheyong Fan ; Shiyun Xiong ✉

Check for updates

View Online

Export Citation

Correcting force error-induced underestimation of lattice thermal conductivity in machine learning molecular dynamics

View Online   Export Citation   CrossMark

Xiguang Wu,[1] (ID) Wenjiang Zhou,[2,3] (ID) Haikuan Dong,[4] (ID) Penghua Ying,[5] (ID) Yanzhou Wang,[6] Bai Song,[2,7,8,a] (ID) Zheyong Fan,[4,b] (ID) and Shiyun Xiong[1,c] (ID)

AFFILIATIONS

[1] Guangzhou Key Laboratory of Low-Dimensional Materials and Energy Storage Devices, School of Materials and Energy, Guangdong University of Technology, Guangzhou 510006, China
[2] Department of Energy and Resources Engineering, Peking University, Beijing 100871, China
[3] School of Advanced Engineering, Great Bay University, Dongguan 523000, China
[4] College of Physical Science and Technology, Bohai University, Jinzhou 121013, China
[5] Department of Physical Chemistry, School of Chemistry, Tel Aviv University, Tel Aviv 6997801, Israel
[6] MSP Group, QTF Centre of Excellence, Department of Applied Physics, Aalto University, FI-00076 Aalto Espoo, Finland
[7] Department of Advanced Manufacturing and Robotics, Peking University, Beijing 100871, China
[8] National Key Laboratory of Advanced MicroNanoManufacture Technology, Beijing 100871, China

[a] Electronic mail: songbai@pku.edu.cn
[b] Electronic mail: brucenju@gmail.com
[c] Author to whom correspondence should be addressed: syxiong@gdut.edu.cn

ABSTRACT

Machine learned potentials (MLPs) have been widely employed in molecular dynamics simulations to study thermal transport. However, the literature results indicate that MLPs generally underestimate the lattice thermal conductivity (LTC) of typical solids. Here, we quantitatively analyze this underestimation in the context of the neuroevolution potential (NEP), which is a representative MLP that balances efficiency and accuracy. Taking crystalline silicon, gallium arsenide, graphene, and lead telluride as examples, we reveal that the fitting errors in the machine-learned forces against the reference ones are responsible for the underestimated LTC as they constitute external perturbations to the interatomic forces. Since the force errors of a NEP model and the random forces in the Langevin thermostat both follow a Gaussian distribution, we propose an approach to correcting the LTC by intentionally introducing different levels of force noises via the Langevin thermostat and then extrapolating to the limit of zero force error. Excellent agreement with experiments is obtained by using this correction for all the prototypical materials over a wide range of temperatures. Based on spectral analyses, we find that the LTC underestimation mainly arises from increased phonon scatterings in the low-frequency region caused by the random force errors.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0213811

I. INTRODUCTION

Lattice thermal conductivity (LTC) of solids is a crucial physical property in many applications, including thermal management of electronics,[1,2] thermoelectric energy conversion,[3–5] and thermal barrier coatings.[6,7] Predicting and engineering LTC[8] is, therefore, of broad interest. Nevertheless, challenges abound owing to the presence of complex structures,[9] defects,[10] and disorders.[10] Among

various approaches to calculating LTC,[11] molecular dynamics (MD) simulation plays a unique role due to its versatility and its natural inclusion of the full lattice anharmonicity. MD simulations are widely applicable in crystals, glasses,[12] and also liquids.[13] Two basic categories are commonly used, including equilibrium molecular dynamics (EMD) based on the Green–Kubo formalism[14,15] and non-equilibrium molecular dynamics (NEMD) based on Fourier's law of heat conduction. Notably, the homogeneous non-equilibrium

01 July 2024 14:10:02

molecular dynamics (HNEMD) method, initially developed by Evans[16] for pairwise interactions and recently generalized to many-body interactions,[17] offers great efficiency for LTC calculations. However, the applicability and predictive power of MD simulations have long been limited by the availability and accuracy of empirical interatomic potentials.

A promising solution to this issue involves constructing machine learned potentials (MLPs) trained against reference energies, forces, and virial stresses of diverse atomic structures calculated at the quantum mechanical level. Many MLPs have been used for thermal conductivity modeling. Enabled by MLPs, the LTCs of many crystals with strong phonon anharmonicity or disorder have been successfully obtained through MD simulation driven by MLP (MLMD), including amorphous GeTe,[18] SnSe,[19] PbTe,[20] metal–organic frameworks,[21] and PH_4AlBr_4.[22] Moreover, with proper quantum corrections, quantitative agreement with experimental data has also been achieved for amorphous materials[23,24] and liquid water[25] over a wide range of temperatures. Despite these successes, previous studies have also shown that for materials with relatively high LTCs, such as $CoSb_3$[26] and cubic silicon (c-Si),[27] the predicted LTCs from MLMD calculations are generally lower than the experimental values. To the best of our knowledge, this discrepancy remains to be systematically understood and corrected, which constitutes the main focus of our present work.

In light of the critical impact of the interatomic forces on the accuracy of MD simulations, we first evaluate the effect of random forces on LTC using HNEMD simulations with a Langevin thermostat[28] based on empirical potentials. A decrease in LTC with an increasing level of random forces is consistently observed in six representative materials: amorphous silicon (a-Si), c-Si, cubic germanium (c-Ge), Si–Ge alloy, graphene, and (10, 10)-carbon nanotube (CNT). Subsequently, we focus on four benchmark materials, including c-Si, GaAs, graphene, and PbTe, and perform LTC calculations using MLMD. In particular, we employ the neuroevolution potential (NEP)[29–31] for its balanced efficiency and accuracy. Similar to literature results, we observe a consistent underestimation of LTC from the MLMD simulations, as compared to the experimental values. However, since the residual force errors of a NEP model and the random forces (white noises) in the Langevin thermostat both follow a Gaussian distribution, we propose an approach to correcting the LTC by intentionally introducing different levels of force noises via the Langevin thermostat and then extrapolating to the limit of zero force error. This extrapolation successfully corrects the LTCs, leading to excellent agreement with experimental data for all the materials considered in a wide range of temperatures. Spectral analyses reveal that the LTC underestimation before the correction mainly originates from increased phonon scatterings at low frequencies caused by the force errors.

II. METHODS

A. Neuroevolution potential

1. The NEP formalism

Here, we briefly review the NEP formalism.[29–31] NEP uses a feedforward neural network to correlate a local descriptor with the site energy U_i of atom i. In a single-hidden-layer neural network

comprising N_{neu} hidden neurons, U_i is expressed as

$$U_i = \sum_{\mu=1}^{N_{neu}} \omega_\mu^{(1)} \tanh\left(\sum_{\nu=1}^{N_{des}} \omega_{\mu\nu}^{(0)} q_\nu^i - b_\mu^{(0)}\right) - b^{(1)}, \quad (1)$$

where N_{des} is the number of descriptor components; q_ν^i is the νth descriptor component of atom i; $\omega_{\mu\nu}^{(0)}$, $\omega_\mu^{(1)}$, $b_\mu^{(0)}$, and $b^{(1)}$ are the trainable parameters; and $\tanh(x)$ is the nonlinear activation function in the hidden layer.

The descriptor vector in NEP includes radial and angular components. The radial components q_n^i $(0 \le n \le n_{max}^R)$ are defined as

$$q_n^i = \sum_{j \neq i} g_n(r_{ij}), \quad (2)$$

where r_{ij} is the distance between atoms i and j and $g_n(r_{ij})$ are a set of radial functions, of which each is formed by a linear combination of Chebyshev polynomials. The angular components include the so-called n-body $(n \ge 3)$ correlations. For example, the three-body ones q_{nl}^i $(0 \le n \le n_{max}^A, 1 \le l \le l_{max})$ are defined as

$$q_{nl}^i = \frac{2l+1}{4\pi} \sum_{j \neq i} \sum_{k \neq i} g_n(r_{ij}) g_n(r_{ik}) P_l(\cos\theta_{ijk}). \quad (3)$$

Here, P_l is the Legendre polynomial and θ_{ijk} is the angle formed by the ij and ik bonds. Note that the radial functions $g_n(r_{ij})$ for the radial and angular descriptor components can have different cutoff radii, which are denoted as r_c^R and r_c^A, respectively. The free parameters are optimized using the separable natural evolutionary strategy[32] by minimizing a loss function that is a weighted sum of the root-mean-square errors (RMSEs) of energy, force, and virial stress, for N_{gen} generations with a population size of N_{pop}. The hyperparameters used for all the materials considered in this work are listed in Table S1.

2. Training datasets

For c-Si, GaAs, graphene, and PbTe, we generate datasets through density functional theory (DFT) calculations using the Vienna *Ab initio* simulation package[33] with the ion–electron interactions described by the projector-augmented wave method.[33,34] For GaAs, the Perdew–Zunger functional with the local density approximation[35] is used to describe the exchange–correlation of electrons, while the Perdew–Burke–Ernzerhof functional with the generalized gradient approximation[36] is used for the other materials. The cutoff energy is 400 eV for PbTe and 600 eV for the other materials. The k-point mesh is $4 \times 4 \times 4$ for c-Si, $2 \times 2 \times 2$ for GaAs and PbTe, and $6 \times 6 \times 1$ for graphene. The energy convergence threshold is 10^{-6} eV for c-Si and graphene and 10^{-8} eV for GaAs and PbTe.

The dataset for each material consists of structures from *ab initio* molecular dynamics (AIMD) simulations (called AIMD structures below) possibly supplemented by those from random cell deformations and atom displacements (called perturbation structures below). For c-Si, there are 900 AIMD structures sampled at various temperatures (100–1000 K) and strain states (unstrained, uniaxial strains of $\pm1\%$ and $\pm2\%$, biaxial strains of $\pm0.5\%$ and $\pm1\%$) and 70 perturbation structures. Each c-Si structure has 64 atoms. For GaAs, there are 197 AIMD structures sampled at various temperatures (100–900 K) in the *NPT* ensemble and 99 perturbation

structures up to ±4% strains. Each GaAs structure has 250 atoms. For graphene, there are 700 AIMD structures sampled at various temperatures (100–1000 K) and strain states (unstrained, biaxial strains of ±0.5%, ±1%, and 2%). Each graphene structure has 72 atoms. For PbTe, there are 60 AIMD structures sampled from 100 to 1100 K with a fixed cell and 64 perturbation structures up to ±4% strains. Each PbTe structure has 216 atoms.

After obtaining the datasets, we used the GPUMD package[37] (the nep executable) to train the NEP models. The parity plots and accuracy metrics are detailed in Figs. S1–S4. Force test errors will be further discussed and used in Sec. III B.

B. Thermal conductivity calculation using MD

1. The HNEMD method

We use the efficient HNEMD method[17] for many-body potentials to calculate the LTCs. In HNEMD, an external driving force on each atom i

$$\mathbf{F}_i^{\text{ext}} = \mathbf{F}_e \cdot \mathbf{W}_i \qquad (4)$$

is applied during the simulation. Here, $\mathbf{F}_e$ is the driving force parameter (of the dimension of inverse length) and[29,31]

$$\mathbf{W}_i = \sum_{j \neq i} \mathbf{r}_{ij} \otimes \frac{\partial U_j}{\partial \mathbf{r}_{ji}} \qquad (5)$$

is the virial tensor of atom i, where U_j is the site energy of atom j, and $\mathbf{r}_{ij} \equiv \mathbf{r}_j - \mathbf{r}_i$, with $\mathbf{r}_i$ being the position of atom i. The driving force parameter should be large enough to ensure a large signal-to-noise ratio and be small enough to maintain the system in the linear-response regime. In the linear-response regime, the LTC tensor $\kappa_{\mu\nu}$ can be calculated from the following relation:[17]

$$\frac{\langle J_\mu(t) \rangle_{\text{ne}}}{TV} = \sum_\nu \kappa_{\mu\nu} F_e^{\ \nu}, \qquad (6)$$

where $\langle J_\mu(t) \rangle_{\text{ne}}$ represents a non-equilibrium ensemble average of the heat current, T is the system temperature, and V is the system volume. The heat current for the NEP model has been derived to be[29,31]

$$\mathbf{J} = \sum_i \mathbf{W}_i \cdot \mathbf{v}_i, \qquad (7)$$

where $\mathbf{v}_i$ is the velocity of atom i.

The HNEMD formalism also allows for an efficient calculation of the frequency-resolved LTC $\kappa(\omega)$ via the following relation:[17]

$$\frac{2}{VT} \int_{-\infty}^{+\infty} e^{i\omega t} K^\mu(t) dt = \sum_\nu \kappa_{\mu\nu}(\omega) F_e^{\ \nu}, \qquad (8)$$

where

$$K^\mu(t) = \sum_i \sum_\nu \langle W_i^{\mu\nu}(0) v_i^\nu(t) \rangle_{\text{ne}} \qquad (9)$$

is the virial-velocity correlation function.

2. Thermostats in HNEMD simulations

HNEMD simulations are normally performed in the *NVT* ensemble realized by using a *global* thermostat, such as the Nosé–Hoover chain (NHC)[38] or the Bussi–Donadio–Parrinello[39] thermostat. In contrast, a *local* thermostat such as the Langevin thermostat[28] is avoided because it can introduce (white) noises through random forces, leading to the following equations of motion:

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i}, \qquad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i - \frac{\mathbf{p}_i}{\tau_T} + \mathbf{f}_i. \qquad (10)$$

Here, τ_T is a time parameter; $\mathbf{r}_i$, $\mathbf{p}_i$, and m_i are, respectively, the position, momentum, and mass of atom i; $\mathbf{F}_i$ is the force on atom i resulting from the interatomic potential; and $\mathbf{f}_i$ is the random force on atom i. Each component of the random force forms a Gaussian distribution with zero mean and a variance of

$$\sigma_{\text{L}}^2 = \frac{2k_{\text{B}} T m}{\tau_T \Delta t}, \qquad (11)$$

where m is the average atom mass in the system, k_{B} is the Boltzmann constant, and Δt is the integration time step. The random forces can affect the dynamics of the system and thus time-correlation properties such as the heat current autocorrelation function, leading to reduced LTC as compared to the case of using a global thermostat. Clearly, a smaller τ_T gives a larger random force variance and a stronger reduction in the LTC. We will demonstrate this effect using examples.
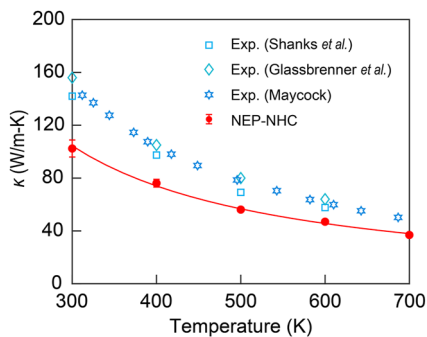
3. MD simulation details

All the MD simulations are performed using the GPUMD package[37] (the gpumd executable), with a time step of 1 fs. For all the materials, we use a sufficiently large simulation cell to eliminate finite-size effects. In MD simulations with empirical potentials, the simulation cells contain 32 768 atoms for Si–Ge alloy, c-Ge, c-Si, and a-Si, 15 416 atoms for graphene, and 16 280 atoms for the (10,10)-CNT. The a-Si samples are prepared by employing a melt–quench–anneal process, first equilibrating at 2000 K for 10 ns, then quenching down to 300 K during 30 ns, and finally annealing at 300 K for 10 ns. In MD simulations with NEP models, the simulation cells contain 13 824, 8000, 16 000, and 36 000 atoms for c-Si, GaAs, graphene, and PbTe, respectively. These cells have been tested to be large enough to eliminate the finite-size effects in HNEMD simulations (see Fig. S7). For each material, we first equilibrate the system in the *NPT* ensemble (with a target pressure of zero) for 2 ns and *NVT* ensemble for another 2 ns and then calculate the LTC in the *NVT* ensemble during a production time of 10–20 ns. For each material at each temperature, three to five independent runs are performed to improve the statistical accuracy and obtain an error estimate. The error bars are calculated from the statistical standard error of independent simulations. An example of c-Si at 300 K is shown in Fig. S6.
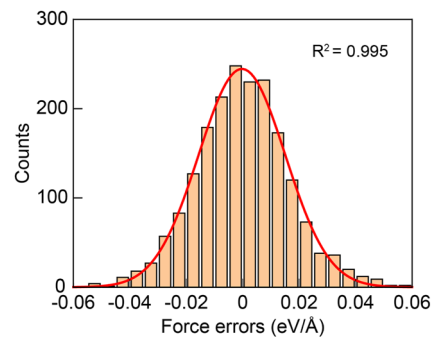
III. RESULTS AND DISCUSSION

A. Thermal conductivity underestimation in MLMD

To begin with, we take c-Si as an example to demonstrate the thermal conductivity underestimation in MLMD simulations, using NEP as a representative MLP. As illustrated in Fig. 1, the calculated

FIG. 1. Comparison of κ for c-Si from NEP-MD simulations and experimental measurements.[40–42] Here, the NHC thermostat is used. Error bars are smaller than the symbol sizes for the calculated values.

FIG. 2. Force error distribution for c-Si at T = 300 K. Fitting to the Gaussian distribution yields a coefficient of determination R^2 = 0.995.

TABLE I. RMSEs σ_{mlp} of force prediction for the four NEP models at various temperatures.

T (K)	σ_{mlp} (meV/Å)			
	c-Si	GaAs	Graphene	PbTe
300	16.7	16.6	29.2	27.0
400	21.3	19.8	30.1	29.9
500	28.3	23.5	32.5	34.4
600	30.1	26.8	36.6	37.1
700	41.6	30.2	42.4	42.1

LTC values from 300 to 700 K are consistently lower than experimental measurements, especially at low temperatures. For instance, at 300 K, MLMD simulations yield a LTC of 102 ± 6 W/m K. While this is more accurate than the value of 240 W/m K as obtained from a Stillinger–Weber potential,[43] it is still ~32% lower than the experimental value of about 150 W/m K.[40] An EMD simulation based on the Gaussian approximation potential (GAP) also reported a lower-than-experiment value of 121 W/m K.[27] A similar trend of underestimation is observed for c-Si at other temperatures by GAP,[27] for GeTe by NEP,[44] and for CoSb$_3$[26] by moment tensor potential.

B. Role of force noises in reducing LTC

To understand the underestimation of the LTC from MLMD simulations, we notice that a MLP usually has a certain level of error for force prediction compared to the reference data. The RMSEs σ_{mlp} of force prediction for the four materials we considered at different temperatures are presented in Table I.

A crucial observation is that the force errors follow a Gaussian distribution, as shown in Fig. 2 for the example of c-Si at 300 K. This distribution is the same as that for the random forces in the Langevin thermostat, i.e., a Gaussian distribution with zero mean and a certain variance. Based on this similarity, an understanding of the underestimation of the LTC by MLMD simulations can thus be obtained by studying the effect of the Langevin thermostat on the LTC. When the system is coupled to the Langevin thermostat, a random frictional force will be added to all atoms, affecting the dynamics of the system.

According to Newton's equation of motion, the effect of random forces on the atoms is similar to that from randomly varied atomic masses. Thus, the coupling to Langevin thermostat introduces an additional phonon scattering term, of which the strength can be tuned by varying the coupling constant. One could directly use a NEP model for this test, but due to the lower computational cost of empirical potentials, we first use the Tersoff empirical potential[45] to study this effect.

In Fig. 3, we show the inverse LTC ($1/\kappa$) at 300 K as a function of σ_L for six representative materials, including a-Si, Si–Ge alloy, c-Ge, c-Si, graphene, and (10,10)-CNT. As expected, $1/\kappa$ increases with increasing σ_L, which indicates a stronger effect of the random forces in the Langevin thermostat in reducing the calculated LTC. Notably, for all the six materials, $1/\kappa$ exhibits a linear relationship with σ_L. This suggests that the intrinsic LTC without the influence of the random forces in the Langevin thermostat can be obtained by extrapolating to $\sigma_L = 0$. Indeed, the extrapolated values align well with the results from HNEMD simulations based on the NHC thermostat that does not involve random forces, with the largest relative error being <1.5% (see Table S2).

The linear relation between $1/\kappa$ and σ_L can be justified based on the kinetic theory of phonons and Matthiessen's rule. Taking the random forces in the Langevin thermostat as an extra source of phonon scattering, we have
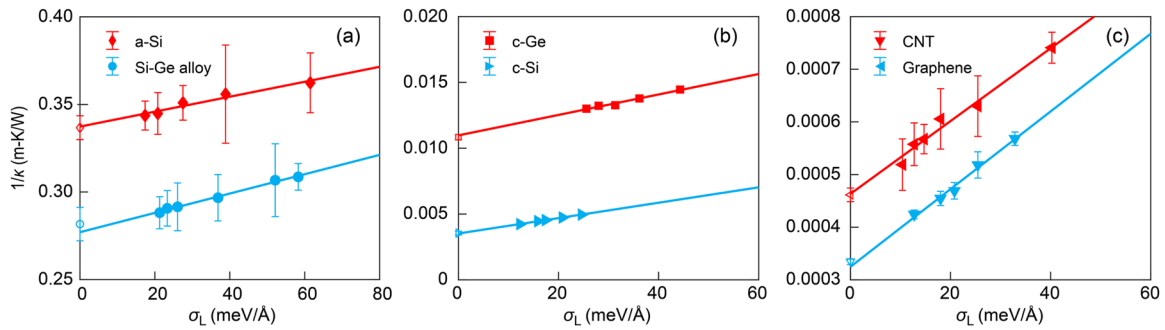
$$\frac{1}{\kappa} = \frac{1}{\kappa_0} + \frac{1}{1/3Cv_g\Lambda_L}, \tag{12}$$

where κ and κ_0 are the LTCs with and without the influence of the random forces, respectively, C is the heat capacity, v_g is the phonon group velocity, and Λ_L is the phonon mean free path resulting from the random forces in the Langevin thermostat. Under first-order approximation with sufficiently small σ_L, $1/\Lambda_L$ should be proportional to σ_L, which brings Eq. (12) to
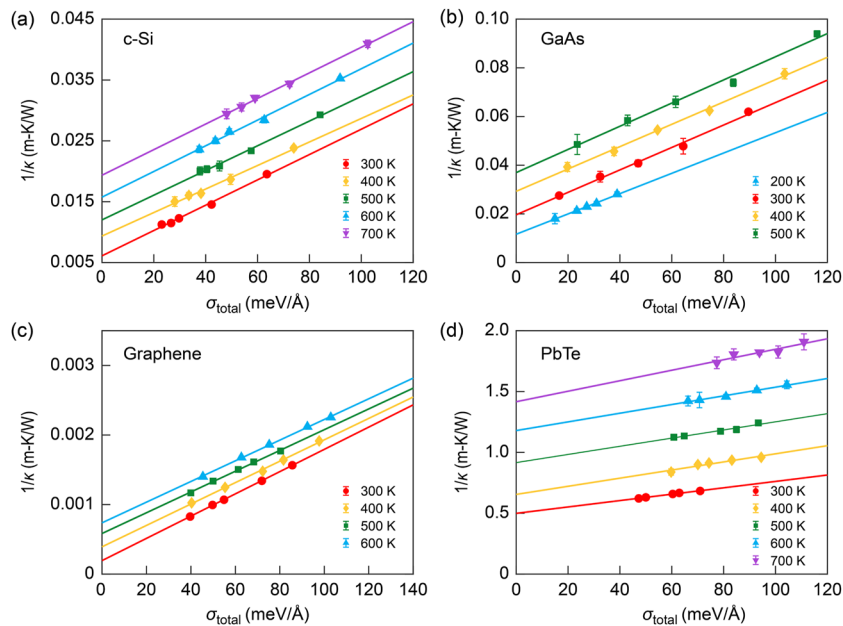
$$\frac{1}{\kappa} = \frac{1}{\kappa_0} + \beta\sigma_L, \tag{13}$$

which gives the observed linear relation between $1/\kappa$ and σ_L with β being a slope parameter.

FIG. 3. Inverse LTC ($1/\kappa$) as a function of the random force variance σ_L of the Langevin thermostat [see Eq. (11)] for (a) a-Si and Si–Ge alloy, (b) c-Ge and c-Si, and (c) graphene and (10,10)-CNT at 300 K. The hollow and filled symbols are the results from the NHC and Langevin thermostats, respectively. The solid lines represent the linear fits to the Langevin data only.

FIG. 4. Inverse LTC ($1/\kappa$) from NEP-MD simulations as a function of the total force error σ_{total} at different temperatures for (a) c-Si, (b) GaAs, (c) graphene, and (d) PbTe. The solid lines indicate linear fits, and the points of intersection at $\sigma_{total} = 0$ correspond to the corrected LTC values.
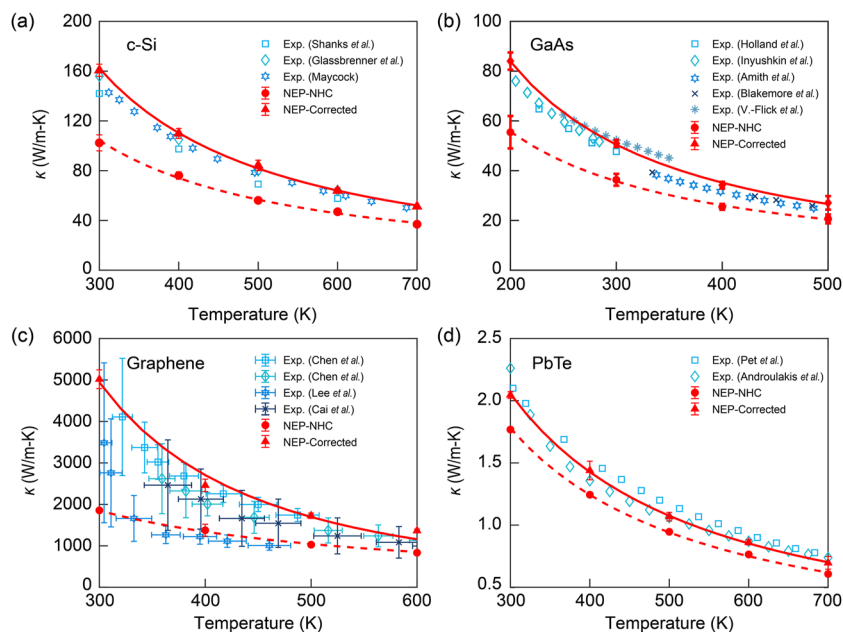
C. Correction of LTC in MLMD

Based on the results above, we can understand why MLMD usually underestimates the LTC, particularly for high-κ materials. According to the linear relation between the inverse LTC and the random force variance, we can devise a method to correct the underestimation of LTC due to the force errors in MLMD. To this end, we note that both the force errors in MLMD and the random forces in the Langevin thermostat follow a Gaussian distribution, and when they are present simultaneously, a new set of force errors are created with a larger variance given by

$$\sigma_{total}^2 = \sigma_L^2 + \sigma_{mlp}^2, \tag{14}$$

according to the properties of Gaussian distribution. Therefore, we can intentionally introduce extra force errors by using MLP-based HNEMD simulations with the Langevin thermostat. The LTC κ_0 without any force errors (including the force errors of the MLP) can be obtained by an extrapolation based on the following relation:

$$\frac{1}{\kappa} = \frac{1}{\kappa_0} + \beta\sigma_{total}, \tag{15}$$

where κ is the LTC of a material calculated by using MLMD with a certain force error variance σ_{mlp} and the Langevin thermostat with a certain random force variance σ_L. The linear relation between $1/\kappa$ and σ_{total} is unfailingly confirmed in Fig. 4 for the four representative
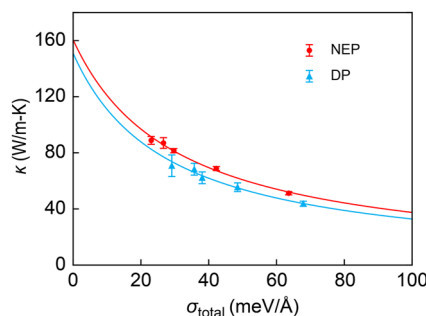
FIG. 5. Corrected (using the Langevin thermostat and extrapolation) and uncorrected (using the NHC thermostat) κ as a function of temperature for (a) c-Si, (b) GaAs, (c) graphene, and (d) PbTe. Experimental values are from Refs. 40–42 (c-Si), Refs. 46–50 (GaAs), Refs. 51–54 (graphene), and Refs. 55 and 56 (PbTe). It is worth noting that the experimentally synthesized samples may contain defects such as vacancies and dislocations. In addition, the synthesized samples usually have limited sizes. Thus, the experimentally measured samples may involve weak defect and boundary scatterings, leading to slight deviations between measured and predicted thermal conductivities.

materials in a wide range of temperatures, whose LTCs span three orders of magnitude.

In Fig. 5, we compare the uncorrected and corrected LTCs from MLMD simulations with experimental results for c-Si, GaAs, graphene, and PbTe. In all the systems, the uncorrected LTCs are consistently lower than the experimental results in the entire temperature range due to the presence of force errors in the MLPs. Remarkably, once the force errors in the MLPs are eliminated via our extrapolation scheme, the LTCs closely approach the experimental data at all the temperatures studied. For graphene, the corrected LTCs slightly exceed the measured values but remain within the experimental uncertainties. This minor discrepancy could arise from factors such as isotope scattering and finite-size effects in the experimental setups,[51–54] which generally lead to reduced LTCs.

To demonstrate that the thermal conductivity underestimation is not specific to NEP, we consider the deep potential (DP)[57,58] as an additional example. We train a DP model for silicon using the same training dataset as used for NEP. The force RMSE at 300 K is determined to be 29.0 meV/Å. Because HNEMD is not available for the DP model via the LAMMPS MD engine,[59] we perform EMD simulations instead. Similar to HNEMD simulations, we use the Langevin thermostat with coupling times of 350, 250, 100, and 40 ps to introduce additional force errors, giving rise to total force errors of 35.7, 38.0, 48.4, and 67.9 meV/Å, respectively. We also use the NHC thermostat corresponding to the total force error of 29.0 meV/Å. The results are shown in Fig. 6. Clearly, the thermal conductivity predicted from the DP model with no additional force errors (corresponds to the case of using the NHC thermostat) is

FIG. 6. Calculated LTC for c-Si from NEP-HNEMD and DP-EMD simulations as a function of the total force error σ_{total}. The DP-EMD results are obtained from 20 independent runs, each with a production time of 2 ns.

also underestimated compared to the experimental value. With the decrease in coupling time in the Langevin thermostat, the thermal conductivity reduces gradually. Based on our proposed extrapolation formula [Eq. (15)], the corrected thermal conductivity from DP is 151 W/m K, which is very close to the one obtained by NEP (160 W/m K) with a relative difference of ~5%. Therefore, we conclude that the underestimation of LTC is a common issue in MLPs and can be corrected by our proposed method.

The need for LTC correction is more pronounced in materials with higher LTCs and at lower temperatures. This is attributed to the weaker anharmonic phonon–phonon interactions, which leads to a
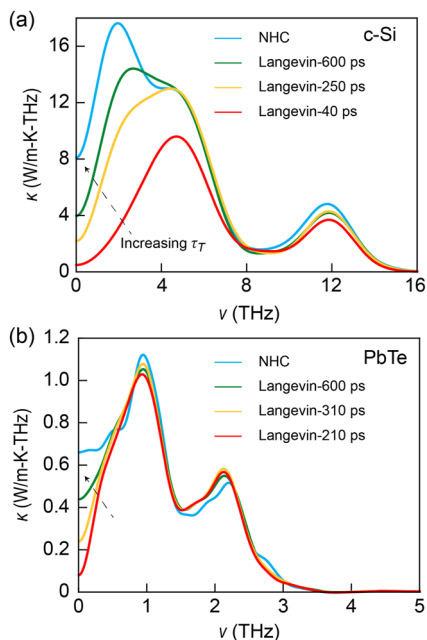
FIG. 7. Spectral LTC $\kappa(\omega)$ from NEP-MD simulations using the NHC and Langevin thermostats for (a) c-Si and (b) PbTe, both at 300 K.

relatively stronger contribution of the phonon scattering by the force errors. This also explains why the amount of correction is large for graphene, which is one of the most thermally conductive materials, intermediate for c-Si and GaAs that have intermediate LTCs, and small for PbTe that has a low LTC. Furthermore, the spectral LTC results in Fig. 7 show that the force errors mainly reduce $\kappa(\omega)$ in the low-frequency region. With increasing force errors, $\kappa(\omega)$ in the low-frequency region is more and more reduced. This further supports the large effect of the force errors in high-LTC materials, which usually have large $\kappa(\omega)$ in the low-frequency region. Therefore, MLMD simulations remain largely accurate for low-LTC materials, such as PbTe,[20] a-Si,[23] amorphous SiO_2,[24] and liquid water.[25]

IV. CONCLUSIONS

In summary, our systematic investigation revealed that the underestimation of lattice thermal conductivity commonly observed in the literature is primarily due to force fitting errors in machine learned potentials. Using empirical potentials and Langevin thermostat we demonstrated that introducing random forces on atoms can significantly reduce the lattice thermal conductivity, supporting our hypothesis. These random forces act as an additional source of phonon scattering, thereby reducing the lattice thermal conductivity. Employing the kinetic theory of phonons and Matthiessen's rule, we established a linear extrapolation formula to estimate the thermal conductivity in the absence of random forces. The validity of the extrapolation scheme was tested using empirical potentials on various materials, including a-Si, Si–Ge alloys, c-Si, c-Ge, graphene, and CNT.

We established that the force errors in machine-learned potentials follow a Gaussian distribution, akin to the distribution of

random forces in the Langevin thermostat. This similarity inspired us to intentionally introduce extra force noises via the Langevin thermostat and then extrapolate to the limit of zero force error. The extrapolated results show excellent agreement with experimental data over a broad temperature range for all the materials studied. Spectral thermal conductivity analyses further indicate that the underestimation of the lattice thermal conductivity is mainly due to increased acoustic phonon scatterings caused by the force errors. Our findings provide a clear explanation for the underestimated thermal conductivity often observed in molecular dynamics simulations based on machine learned potentials. The method of correcting this underestimation we developed will significantly enhance the applicability of machine learned potentials in the prediction of lattice thermal conductivity.

SUPPLEMENTARY MATERIAL

See the supplementary material for the hyperparameters used in NEP, the calculated thermal conductivity using empirical potentials, phonon dispersion relations from NEP models as compared to DFT calculations, size-convergence tests for thermal conductivity calculations, and the parity plots of trained NEPs.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

X.W. and W.Z. contributed equally to this work.

Xiguang Wu: Data curation (equal); Investigation (equal); Visualization (equal); Writing – original draft (equal). **Wenjiang Zhou**: Data curation (equal); Investigation (equal); Visualization (equal); Writing – original draft (equal). **Haikuan Dong**: Data curation (equal); Formal analysis (equal); Investigation (equal); Writing – review & editing (equal). **Penghua Ying**: Data curation (equal); Formal analysis (equal); Writing – review & editing (equal). **Yanzhou**

Wang: Software (equal); Writing – review & editing (equal). **Bai Song**: Funding acquisition (equal); Resources (equal); Supervision (supporting); Writing – review & editing (equal). **Zheyong Fan**: Formal analysis (equal); Methodology (equal); Software (lead); Writing – review & editing (equal). **Shiyun Xiong**: Conceptualization (lead); Formal analysis (equal); Funding acquisition (equal); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY

All the training datasets and the trained NEP models are freely available at https://gitlab.com/brucefan1983/nep-data. The DP training and potential files as well as MD input files are freely available at https://github.com/hityingph/supporting-info.

REFERENCES

[1] A. L. Moore and L. Shi, "Emerging challenges and materials for thermal management of electronics," Mater. Today **17**, 163–174 (2014).

[2] J. S. Kang, M. Li, H. Wu, H. Nguyen, T. Aoki, and Y. Hu, "Integration of boron arsenide cooling substrates into gallium nitride devices," Nat. Electron. **4**, 416–423 (2021).

[3] G. J. Snyder and E. S. Toberer, "Complex thermoelectric materials," Nat. Mater. **7**, 105–114 (2008).

[4] L.-D. Zhao, S.-H. Lo, Y. Zhang, H. Sun, G. Tan, C. Uher, C. Wolverton, V. P. Dravid, and M. G. Kanatzidis, "Ultralow thermal conductivity and high thermoelectric figure of merit in SnSe crystals," Nature **508**, 373–377 (2014).

[5] W. Zhou, Y. Dai, J. Zhang, B. Song, T.-H. Liu, and R. Yang, "Effect of four-phonon interaction on phonon thermal conductivity and mean-free-path spectrum of high-temperature phase SnSe," Appl. Phys. Lett. **121**, 112202 (2022).

[6] J. H. Perepezko, "The hotter the engine, the better," Science **326**, 1068–1069 (2009).

[7] R. Vaßen, M. O. Jarligo, T. Steinke, D. E. Mack, and D. Stöver, "Overview on advanced thermal barrier coatings," Surf. Coat. Technol. **205**, 938–942 (2010).

[8] X. Qian, J. Zhou, and G. Chen, "Phonon-engineered extreme thermal conductivity materials," Nat. Mater. **20**, 1188–1202 (2021).

[9] T. Tadano and S. Tsuneyuki, "Quartic anharmonicity of rattlers and its effect on lattice thermal conductivity of clathrates from first principles," Phys. Rev. Lett. **120**, 105901 (2018).

[10] R. Hanus, R. Gurunathan, L. Lindsay, M. T. Agne, J. Shi, S. Graham, and G. Jeffrey Snyder, "Thermal transport in defective and disordered materials," Appl. Phys. Rev. **8**, 031311 (2021).

[11] X. Gu, Z. Fan, and H. Bao, "Thermal conductivity prediction by atomistic simulation methods: Recent advances and detailed comparison," J. Appl. Phys. **130**, 210902 (2021).

[12] Y. H. Lee, R. Biswas, C. M. Soukoulis, C. Z. Wang, C. T. Chan, and K. M. Ho, "Molecular-dynamics simulation of thermal conductivity in amorphous silicon," Phys. Rev. B **43**, 6573–6580 (1991).

[13] R. Vogelsang, C. Hoheisel, and G. Ciccotti, "Thermal conductivity of the Lennard-Jones liquid by molecular dynamics calculations," J. Chem. Phys. **86**, 6371–6375 (1987).

[14] M. S. Green, "Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids," J. Chem. Phys. **22**, 398–413 (1954).

[15] R. Kubo, "Statistical–mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems," J. Phys. Soc. Jpn. **12**, 570–586 (1957).

[16] D. J. Evans, "Homogeneous NEMD algorithm for thermal conductivity—Application of non-canonical linear response theory," Phys. Lett. A **91**, 457–460 (1982).

[17] Z. Fan, H. Dong, A. Harju, and T. Ala-Nissila, "Homogeneous nonequilibrium molecular dynamics method for heat transport and spectral decomposition with many-body potentials," Phys. Rev. B **99**, 064308 (2019).

[18] G. C. Sosso, D. Donadio, S. Caravati, J. Behler, and M. Bernasconi, "Thermal transport in phase-change materials from atomistic simulations," Phys. Rev. B **86**, 104301 (2012).

[19] H. Liu, X. Qian, H. Bao, C. Y. Zhao, and X. Gu, "High-temperature phonon transport properties of SnSe from machine-learning interatomic potential," J. Phys.: Condens. Matter **33**, 405401 (2021).

[20] R. Cheng, X. Shen, S. Klotz, Z. Zeng, Z. Li, A. Ivanov, Y. Xiao, L.-D. Zhao, F. Weber, and Y. Chen, "Lattice dynamics and thermal transport of PbTe under high pressure," Phys. Rev. B **108**, 104306 (2023).

[21] P. Ying, T. Liang, K. Xu, J. Zhang, J. Xu, Z. Zhong, and Z. Fan, "Sub-micrometer phonon mean free paths in metal–organic frameworks revealed by machine learning molecular dynamics simulations," ACS Appl. Mater. Interfaces **15**, 36412–36422 (2023).

[22] P.-H. Du, C. Zhang, T. Li, and Q. Sun, "Low lattice thermal conductivity with two-channel thermal transport in the superatomic crystal Ph_4AlBr_4," Phys. Rev. B **107**, 155204 (2023).

[23] Y. Wang, Z. Fan, P. Qian, M. A. Caro, and T. Ala-Nissila, "Quantum-corrected thickness-dependent thermal conductivity in amorphous silicon predicted by machine learning molecular dynamics simulations," Phys. Rev. B **107**, 054303 (2023).

[24] T. Liang, P. Ying, K. Xu, Z. Ye, C. Ling, Z. Fan, and J. Xu, "Mechanisms of temperature-dependent thermal transport in amorphous silica from machine-learning molecular dynamics," Phys. Rev. B **108**, 184203 (2023).

[25] K. Xu, Y. Hao, T. Liang, P. Ying, J. Xu, J. Wu, and Z. Fan, "Accurate prediction of heat conductivity of water by a neuroevolution potential," J. Chem. Phys. **158**, 204114 (2023).

[26] P. Korotaev, I. Novoselov, A. Yanilkin, and A. Shapeev, "Accessing thermal conductivity of complex compounds by machine learning interatomic potentials," Phys. Rev. B **100**, 144308 (2019).

[27] X. Qian, S. Peng, X. Li, Y. Wei, and R. Yang, "Thermal conductivity modeling using machine learning potentials: Application to crystalline and amorphous silicon," Mater. Today Phys. **10**, 100140 (2019).

[28] G. Bussi and M. Parrinello, "Accurate sampling using Langevin dynamics," Phys. Rev. E **75**, 056707 (2007).

[29] Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen, and T. Ala-Nissila, "Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport," Phys. Rev. B **104**, 104309 (2021).

[30] Z. Fan, "Improving the accuracy of the neuroevolution machine learning potential for multi-component systems," J. Phys.: Condens. Matter **34**, 125902 (2022).

[31] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J. Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun, P. Erhart, Y. Su, and T. Ala-Nissila, "GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations," J. Chem. Phys. **157**, 114801 (2022).

[32] T. Schaul, T. Glasmachers, and J. Schmidhuber, "High dimensions and heavy tails for natural evolution strategies," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation, GECCO '11* (Association for Computing Machinery, New York, NY, USA, 2011), pp. 845–852.

[33] P. E. Blöchl, "Projector augmented-wave method," Phys. Rev. B **50**, 17953–17979 (1994).

[34] G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method," Phys. Rev. B **59**, 1758–1775 (1999).

[35] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," Phys. Rev. B **23**, 5048–5079 (1981).

[36] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**, 3865–3868 (1996).

[37] Z. Fan, W. Chen, V. Vierimaa, and A. Harju, "Efficient molecular dynamics simulations with many-body potentials on graphics processing units," Comput. Phys. Commun. **218**, 10–16 (2017).

01 July 2024 14:10:02

[38] G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé–Hoover chains: The canonical ensemble via continuous dynamics," J. Chem. Phys. **97**, 2635–2643 (1992).

[39] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," J. Chem. Phys. **126**, 014101 (2007).

[40] C. J. Glassbrenner and G. A. Slack, "Thermal conductivity of silicon and germanium from 3 K to the melting point," Phys. Rev. **134**, A1058–A1069 (1964).

[41] H. R. Shanks, P. D. Maycock, P. H. Sidles, and G. C. Danielson, "Thermal conductivity of silicon from 300 to 1400 K," Phys. Rev. **130**, 1743–1748 (1963).

[42] P. Maycock, "Thermal conductivity of silicon, germanium, III–V compounds and III–V alloys," Solid-State Electron. **10**, 161–168 (1967).

[43] S. G. Volz and G. Chen, "Molecular-dynamics simulation of thermal conductivity of silicon crystals," Phys. Rev. B **61**, 2651–2656 (2000).

[44] J. Zhang, H. C. Zhang, W. Li, and G. Zhang, "Thermal conductivity of GeTe crystals based on machine learning potentials," Chin. Phys. B **33**, 047402 (2024).

[45] J. Tersoff, "Empirical interatomic potential for silicon with improved elastic properties," Phys. Rev. B **38**, 9902–9905 (1988).

[46] A. Amith, I. Kudman, and E. F. Steigmeier, "Electron and phonon scattering in GaAs at high temperatures," Phys. Rev. **138**, A1270–A1276 (1965).

[47] J. S. Blakemore, "Semiconducting and other major properties of gallium arsenide," J. Appl. Phys. **53**, R123–R181 (1982).

[48] A. Vega-Flick, D. Jung, S. Yue, J. E. Bowers, and B. Liao, "Reduced thermal conductivity of epitaxial GaAs on Si due to symmetry-breaking biaxial strain," Phys. Rev. Mater. **3**, 034603 (2019).

[49] A. V. Inyushkin, A. N. Taldenkov, A. Y. Yakubovsky, A. V. Markov, L. Moreno-Garsia, and B. N. Sharonov, "Thermal conductivity of isotopically enriched ^{71}GaAs crystal," Semicond. Sci. Technol. **18**, 685 (2003).

[50] M. G. Holland, "Phonon scattering in semiconductors from thermal conductivity studies," Phys. Rev. **134**, A471–A480 (1964).

[51] S. Chen, A. L. Moore, W. Cai, J. W. Suk, J. An, C. Mishra, C. Amos, C. W. Magnuson, J. Kang, L. Shi, and R. S. Ruoff, "Raman measurements of thermal transport in suspended monolayer graphene of variable sizes in vacuum and gaseous environments," ACS Nano **5**, 321–328 (2011).

[52] S. Chen, Q. Wu, C. Mishra, J. Kang, H. Zhang, K. Cho, W. Cai, A. A. Balandin, and R. S. Ruoff, "Thermal conductivity of isotopically modified graphene," Nat. Mater. **11**, 203–207 (2012).

[53] J.-U. Lee, D. Yoon, H. Kim, S. W. Lee, and H. Cheong, "Thermal conductivity of suspended pristine graphene measured by Raman spectroscopy," Phys. Rev. B **83**, 081419 (2011).

[54] W. Cai, A. L. Moore, Y. Zhu, X. Li, S. Chen, L. Shi, and R. S. Ruoff, "Thermal transport in suspended and supported monolayer graphene grown by chemical vapor deposition," Nano Lett. **10**, 1645–1651 (2010).

[55] J. Androulakis, I. Todorov, D.-Y. Chung, S. Ballikaya, G. Wang, C. Uher, and M. Kanatzidis, "Thermoelectric enhancement in PbTe with K or Na codoping from tuning the interaction of the light- and heavy-hole valence bands," Phys. Rev. B **82**, 115209 (2010).

[56] Y. Pei, X. Shi, A. LaLonde, H. Wang, L. Chen, and G. J. Snyder, "Convergence of electronic bands for high performance bulk thermoelectrics," Nature **473**, 66–69 (2011).

[57] L. Zhang, J. Han, H. Wang, R. Car, and W. E, "Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics," Phys. Rev. Lett. **120**, 143001 (2018).

[58] H. Wang, L. Zhang, J. Han, and E. Weinan, "DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics," Comput. Phys. Commun. **228**, 178–184 (2018).

[59] S. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," J. Comput. Phys. **117**, 1–19 (1995).