Letter

Heat flux for semilocal machine-learning potentials

Marcel F. Langer [1,2,3,*] Florian Knoop [4,3,†] Christian Carbogno [3] Matthias Scheffler [3] and Matthias Rupp [3,5,6]

[1]*Machine Learning Group, Technische Universität Berlin, 10587 Berlin, Germany*
[2]*Berlin Institute for the Foundations of Learning and Data, 10623 Berlin, Germany*
[3]*The NOMAD Laboratory at the FHI of the Max-Planck-Gesellschaft and IRIS Adlershof of the Humboldt Universität zu Berlin, 14195 Berlin, Germany*
[4]*Theoretical Physics Division, Department of Physics, Chemistry and Biology (IFM), Linköping University, 581 83 Linköping, Sweden*
[5]*Department of Computer and Information Science, University of Konstanz, 78464 Konstanz, Germany*
[6]*Materials Research and Technology Department, Luxembourg Institute of Science and Technology, Belvaux, Luxembourg*

The Green-Kubo (GK) method is a rigorous framework for heat transport simulations in materials. However, it requires an accurate description of the potential-energy surface and carefully converged statistics. Machine-learning potentials can achieve the accuracy of first-principles simulations while allowing to reach well beyond their simulation time and length scales at a fraction of the cost. In this Letter, we explain how to apply the GK approach to the recent class of message-passing machine-learning potentials, which iteratively consider semilocal interactions beyond the initial interaction cutoff. We derive an adapted heat flux formulation that can be implemented using automatic differentiation without compromising computational efficiency. The approach is demonstrated and validated by calculating the thermal conductivity of zirconium dioxide across temperatures.

The thermal conductivity tensor $\boldsymbol{\kappa}$ describes the ability of a material to conduct heat when exposed to a temperature gradient. Its computational prediction is of great interest for the design of novel high-performance materials which are needed, for example, as thermal barrier coatings in engines [1] or thermoelectrics for waste heat recovery [2]. Such materials often feature complex structure and strongly anharmonic potential-energy surfaces (PESs) [3,4]. This implies the need to evaluate $\boldsymbol{\kappa}$ with a non-perturbative method such as the Green-Kubo (GK) method [5–9].

In the GK approach, $\boldsymbol{\kappa}$ is expressed in terms of the integral of the autocorrelation function of the instantaneous heat flux $\mathbf{J}(t)$ as observed in equilibrium molecular dynamics (MD) simulations,

$$\boldsymbol{\kappa}(T, p) = \frac{1}{k_\mathrm{B} T^2 V} \lim_{t \to \infty} \int_0^t d\tau \, \langle \mathbf{J}(\tau) \otimes \mathbf{J}(0) \rangle_{T,p}, \quad (1)$$

where k_B is the Boltzmann constant, T is the temperature, V is the simulation cell volume, and $\langle \cdot \rangle_{T,p}$ denotes an ensemble average at temperature T and pressure p.

High-accuracy MD simulations can be performed using density-functional theory (DFT) when the exchange-correlation approximation is reliable [10]. For the evaluation of Eq. (1), this approach [11,12] suffers from its numerical cost, which limits the system sizes and timescales that can be treated, and therefore requires additional denoising and extrapolation approaches [12–14]. The alternative, so far, has been the use of semiempirical force fields (FFs) [15]. Here, the interatomic interactions are described by a physically motivated analytical equation that includes free parameters which are fitted to experimental or *ab initio* results. This classical FF approach has been very successful, as it enables proper consideration of the ensemble averages needed in Eq. (1). However, the restricted flexibility of FFs may limit their generality and ability to model novel materials.

A new, more general class of FFs is the family of machine-learning potentials (MLPs) which leverage techniques like neural networks [16–21]. MLPs offer, in principle, unrestricted flexibility, but are limited to the mechanisms and information that are provided by their training data. In local MLPs, linear scaling with system size is achieved by using the short-range nature of chemical bonding [22] to decompose the total energy into contributions that depend only on local atomic environments. However, a strict locality assumption limits the flexibility and therefore accuracy of such MLPs. Some FFs therefore include explicit long-range electrostatic and van der Waals interactions [23–27]. Semilocal MLPs [28–39] build up longer-range correlations iteratively from local ones through message-passing mechanisms, thereby preserving linear scaling with system size. They have recently emerged as an alternative to strictly local MLPs and have shown promising performance in benchmark settings and first applications [35,36,40–42].

While local MLPs have been used to investigate thermal transport via GK [43–50], more recent semilocal MLPs have not yet been applied, partially because a heat flux formulation

---

*Corresponding author: mail@marcel.science
†Corresponding author: florian.knoop@liu.se

that incorporates message-passing mechanisms was lacking. In this work, we fill that gap and extend the GK approach to semilocal potentials. To this end, we derive a formulation of the heat flux that explicitly accounts for semilocal interactions, finding that the resulting thermal conductivity significantly differs from a purely local form. While the computation of this heat flux scales quadratically with system size, we show that an alternative, yet equivalent, form based on an extended auxiliary system can be introduced, leading to overall linear scaling and straightforward practical implementation of the approach via automatic differentiation (AD) [51,52]. Using the SchNet message-passing neural network (MPNN) [29,30], we demonstrate the accuracy and feasibility of large-scale semilocal MLP thermal conductivity calculations for zirconia (ZrO_2), an oxide known for its strongly anharmonic PES [53,54].

For any potential function $U(\{\mathbf{R}_J\})$ of the atomic positions $\{\mathbf{R}_J\}$ that can be decomposed into atomic contributions $U_I(\{\mathbf{R}_J\})$ such that $U = \sum_I U_I(\{\mathbf{R}_J\})$, the heat flux is given by the classical equivalent of a formula by Hardy [55], which we rederive in the Supplemental Material [56] to explicitly account for periodic boundary conditions. This full "Hardy" heat flux reads

$$\mathbf{J} = \sum_{\substack{I \in \mathcal{R}_{sc} \\ J \in \mathcal{R}_{all}}} \left[\mathbf{R}_{JI} \left(\frac{\partial U_I}{\partial \mathbf{R}_J} \cdot \dot{\mathbf{R}}_J \right) \right] + \sum_{I \in \mathcal{R}_{sc}} E_I \dot{\mathbf{R}}_I \qquad (2)$$

$$=: \mathbf{J}_{pot} + \mathbf{J}_{conv}, \qquad (3)$$

where $\mathbf{R}_{JI} = \mathbf{R}_I - \mathbf{R}_J$ is the atom-pair vector connecting atoms I and J, $\partial U_I/\partial \mathbf{R}_J$ is the change in the energy contribution U_I with respect to moving atom J, $\dot{\mathbf{R}}_J$ is the velocity of atom J, and $E_I = U_I + \frac{1}{2}m_I\dot{\mathbf{R}}_I^2$ is the total energy of atom I. The double sum extends over all atoms $\mathcal{R}_{sc}$ in the simulation cell, while $\mathcal{R}_{all}$ enumerates the full, infinite, bulk system. As this work considers FFs and MLPs that explicitly define atomic potential energies U_I, total atomic energies E_I and, consequently, $\mathbf{J}_{conv}$ can be computed in a straightforward manner. We therefore discuss only the more involved computation of $\mathbf{J}_{pot}$ in the following, whereas the heat flux used for calculating κ in this work is always equivalent to the full flux given by Eq. (2). The nomenclature for heat flux contributions and the relation of Eq. (2) to DFT formulations are further discussed in the Supplemental Material [56].

Evaluating $\mathbf{J}_{pot}$ requires disentangling the contributions of every atom, including those in the bulk, to every atomic potential energy U_I. This can be challenging for nonpairwise, many-body potentials, leading to the development of specialized expressions for different FFs [57–62], many of which were recently unified and shown to be equivalent to Eq. (2) by Fan *et al.* [60].

In the next step, we introduce an interaction cutoff radius r_c which defines an atomic neighborhood $\mathcal{N}(I) = \{\mathbf{R}_J : |\mathbf{R}_{IJ}^{MIC}| \leqslant r_c, \mathbf{R}_J \in \mathcal{R}_{sc}\}$ for each atom I using the minimum image convention (MIC). The cutoff needs to be chosen such that MIC wrapping happens outside the local neighborhood. This leads to the notion of a *local potential* $U_I = U_I(\{\mathbf{R}_{IJ}^{MIC} : J \in \mathcal{N}(I)\})$. In an interatomic potential of this type, each atomic potential-energy contribution is a function *only* of outgoing atom-pair vectors $\mathbf{R}_{IJ}^{MIC}$ with neighbors in $\mathcal{N}(I)$,

and the corresponding incoming vectors $\mathbf{R}_{JI}^{MIC}$ are treated as independent variables. Then,

$$\frac{\partial U_I}{\partial \mathbf{R}_J} = \sum_{K \in \mathcal{N}(I)} \frac{\partial U_I}{\partial \mathbf{R}_{IK}^{MIC}} \frac{\partial \mathbf{R}_{IK}^{MIC}}{\partial \mathbf{R}_J} = \frac{\partial U_I}{\partial \mathbf{R}_{IJ}^{MIC}}. \qquad (4)$$

The corresponding local formulation of $\mathbf{J}_{pot}$ is then

$$\mathbf{J}_{pot}^{local} = \sum_{I \in \mathcal{R}_{sc}} \sum_{J \in \mathcal{N}(I)} \mathbf{R}_{JI}^{MIC} \left(\frac{\partial U_I}{\partial \mathbf{R}_{IJ}^{MIC}} \cdot \dot{\mathbf{R}}_J \right), \qquad (5)$$

where we note that the partial derivative of U_I is now performed with respect to the atom-pair vector $\mathbf{R}_{IJ}$, similar to the heat flux formulation defined in Ref. [60]. The resulting expression can be implemented efficiently with AD, as discussed in the Supplemental Material [56] and Ref. [63].

While being exact for *local* potentials, this formulation of the heat flux does not apply to the *semilocal* case of MLPs based on MPNNs. In such potentials, longer-range interactions are introduced *without* explicitly increasing the cutoff r_c by building them up iteratively: Neighboring atoms are allowed to exchange information for a fixed number of iterations M [28]. Neighboring environments up to an effective cutoff radius $r_c^{eff} = Mr_c$ therefore become correlated; atomic potential energies U_I acquire a dependence on atom-pair vectors *outside* of their immediate neighborhoods $\mathcal{N}(I)$, rendering Eq. (4) and, consequently, Eq. (5) inapplicable. In principle, U_I can depend on atom-pair vectors in *all* atomic neighborhoods, $U_I = U_I(\{\mathbf{R}_{LK}^{MIC} : L \in \mathcal{R}_{sc}, K \in \mathcal{N}(L)\})$, so

$$\frac{\partial U_I}{\partial \mathbf{R}_J} = \sum_{\substack{L \in \mathcal{R}_{sc} \\ K \in \mathcal{N}(L)}} \frac{\partial U_I}{\partial \mathbf{R}_{LK}^{MIC}} \frac{\partial \mathbf{R}_{LK}^{MIC}}{\partial \mathbf{R}_J}$$

$$= \sum_{K \in \mathcal{N}(J)} \left(\frac{\partial U_I}{\partial \mathbf{R}_{KJ}^{MIC}} - \frac{\partial U_I}{\partial \mathbf{R}_{JK}^{MIC}} \right), \qquad (6)$$

where we keep both terms since ingoing and outgoing atom-pair vectors are treated independently. Substitution into Eq. (2) yields the heat flux for the semilocal case

$$\mathbf{J}_{pot}^{semilocal} = \sum_{\substack{I \in \mathcal{R}_{sc} \\ J \in \mathcal{R}_{sc} \\ K \in \mathcal{N}(J)}} \mathbf{R}_{JI}^{MIC} \left[\left(\frac{\partial U_I}{\partial \mathbf{R}_{KJ}} - \frac{\partial U_I}{\partial \mathbf{R}_{JK}} \right) \cdot \dot{\mathbf{R}}_J \right]. \qquad (7)$$

This semilocal form of the heat flux generalizes $\mathbf{J}_{pot}^{local}$ defined in Eq. (5) to semilocal MLPs . In the case of $M = 1$, the two forms become identical.

Equation (7) reflects the standard construction of semilocal MLPs; a double sum over all atoms is required, and its evaluation formally scales quadratically with system size. As shown in Fig. 1, a direct implementation of this form is therefore impractical. While force predictions for a semilocal MLP based on the SchNet architecture [29,30] remain below 100 ms for all system sizes investigated, the unoptimized calculation of the heat flux dominates the computational cost by several orders of magnitude at the system sizes required for the GK method.

If the analytical form of U_I were known, a lower-scaling evaluation of the heat flux might be accessible by deriving
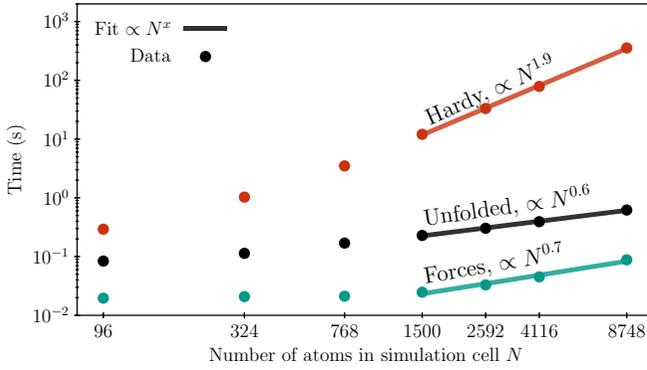
FIG. 1. Computation time per time step for different system sizes N for zirconia, evaluating a SchNet MPNN, for different heat flux formulations on a single Tesla Volta V100 32 GB GPU. To estimate the asymptotic scaling, a function proportional to N^x has been fitted to the results for large N. Note that in this setup with limited memory, the truly asymptotic limit cannot be reached.

and implementing analytical derivatives. Modern MLPs, however, typically rely on AD [51,52] for efficiently computing derivatives without requiring detailed information about the functional form of the MLP .

To take advantage of this, we now derive an adapted form of the heat flux that preserves the implicit treatment of interactions beyond local environments to retain the computational efficiency of semilocal MLPs while explicitly attributing all contributions to U_I to bulk positions for $\mathbf{J}_{\text{pot}}$ in Eq. (2). This is achieved by constructing an extended simulation cell that includes all replicas that interact with atoms in the simulation cell [64,65]. This yields a set of "unfolded" positions $\mathcal{R}_{\text{unf}}$, consisting of the original simulation cell and a shell of replica positions up to the effective cutoff radius $r_c^{\text{eff}} = M r_c$.

Atomic neighborhoods are then constructed *without* periodic boundary conditions and the MIC. Interactions across simulation cell boundaries are therefore now treated explicitly. The potential energy obtained by summing over the original simulation cell U remains unchanged in this construction. This allows retaining the small cutoffs needed for efficiency while enabling AD to distinguish derivatives with respect to positions in the simulation cell and replica positions.

With this construction, Eq. (2) can be rewritten as

$$\mathbf{J}_{\text{pot}}^{\text{unfolded}} = \sum_{\substack{I \in \mathcal{R}_{\text{sc}} \\ J \in \mathcal{R}_{\text{unf}}}} \left(\mathbf{R}_I \frac{\partial U_I}{\partial \mathbf{R}_J} \cdot \dot{\mathbf{R}}_J - \mathbf{R}_J \frac{\partial U_I}{\partial \mathbf{R}_J} \cdot \dot{\mathbf{R}}_J \right). \quad (8)$$

Both terms can be implemented efficiently with AD by executing the sum over I before differentiation. For the first term, this can be achieved by excluding $\mathbf{R}_I$ from the computational graph. For the second term, the sum can be simply moved inside the derivative. Further details are given in the Supplemental Material [56]. The overall computational cost is then proportional to $\mathcal{R}_{\text{unf}}$. Since the number of additional positions is proportional only to the surface area of the simulation cell and the number of interactions M, the overall asymptotic linear scaling is restored, with $|\mathcal{R}_{\text{unf}}| \propto N + N^{2/3}$ (see Fig. 1).

To validate the approach, we benchmark the performance of a semilocal MLP, in particular the SchNet [29,30] MPNN

architecture, for GK calculations on zirconia (ZrO_2) and compare our results to results obtained with size-extrapolated *ab initio* GK [12], as well as GK with a local MLP [50], and experimental measurements [66–68].

Training and validation data were generated using *ab initio* MD in the NpT ensemble, with four different trajectories heating up an initially tetragonal simulation cell with 96 atoms to target temperatures 750, 1500, 2250, and 3000 K. In total, 10 000 single-point calculations were performed using FHI-AIMS [69] and FHI-VIBES [70], using the PBEsol [71] functional and otherwise following the computational approach of Ref. [12].

On these data, we train a SchNet MPNN, implemented in SCHNETPACK [72], with cutoff radius $r_c = 5$ Å. We choose an interaction depth $M = 2$, leading to an effective cutoff of 10 Å. In line with recent findings by others [36], we find this to be sufficient, as test set error does not significantly decrease for higher values of M or r_c. Further details on the training procedure, choice of hyperparameters, and testing of the MLP can be found in the Supplemental Material [56].

We find that this simple approach yields an MLP capable of describing the dynamics in monoclinic and tetragonal zirconia up to temperatures of approximately 2000 K. In this temperature range, the anharmonic vibrational density of states matches that obtained from DFT. Beyond 2000 K, the oxygen atoms become more mobile, and different types of dynamical events are observed, in particular exchange-type oxygen diffusion. This behavior is also present in the training data, in line with recent literature [73], although slightly different diffusion events are observed given the smaller simulation cells and trajectory lengths. When diffusion increases at higher temperatures, the MLP becomes unstable. This might be due to the limited amount of training data for these processes, especially for thermodynamic conditions close to the tetragonal-to-cubic phase transition. These observations suggest that an accurate description of defect formation is necessary to investigate zirconia above 2000 K, which is beyond the scope of the current work.

Figure 2 compares our efficient implementation with the full semilocal heat flux, as well as the purely local heat flux formulation. Due to the high computational cost of the unoptimized implementation, we use a small simulation cell with $N = 768$ atoms and rely on the noise reduction scheme introduced in Ref. [14]. The results confirm that our implementation $\mathbf{J}_{\text{pot}}^{\text{unfolded}}$ is equivalent to the semilocal heat flux $\mathbf{J}_{\text{pot}}^{\text{semilocal}}$, while the local flux $\mathbf{J}_{\text{pot}}^{\text{local}}$ is not, underestimating the thermal conductivity by approximately 40% due to missing interactions beyond $M = 1$. A similar effect was observed when formulations applicable to pairwise additive potentials were used for many-body force fields [61,62].

Enabled by the computationally efficient access to $\mathbf{J}$ for semilocal MLPs, we then predict the thermal conductivity of zirconia across temperatures. Since the focus of the present work is the heat flux, we do not treat the thermodynamics of zirconia with the MLP but use experimentally determined lattice parameters [74,75] to account for lattice expansion. At 1400 K, both phases are investigated, as the monoclinic phase is sufficiently stable during the course of the simulations, which consist of 11 trajectories of 1 ns each, with an
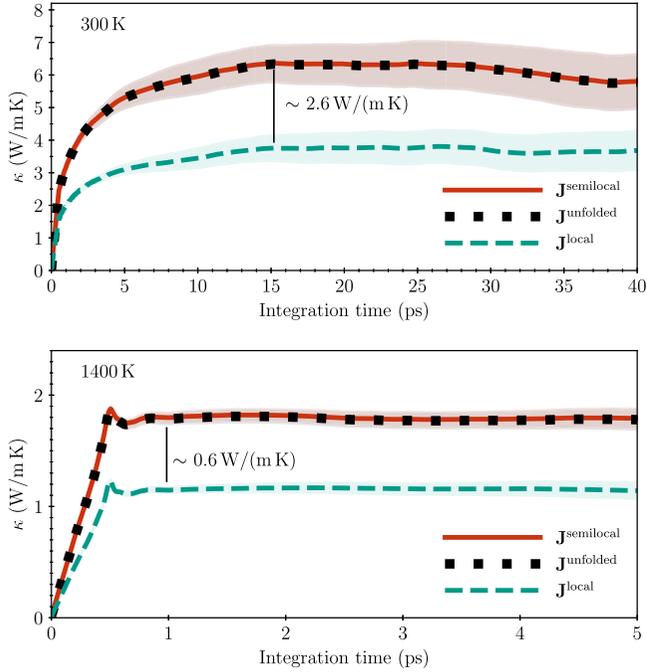
FIG. 2. Comparison of the integrals of the heat flux autocorrelation function for different formulations of the heat flux. The efficient reformulation of the heat flux $\mathbf{J}^{\mathrm{unfolded}}$ is equivalent to the full heat flux $\mathbf{J}^{\mathrm{semilocal}}$ but not to $\mathbf{J}^{\mathrm{local}}$, which neglects semilocal interactions. Results are given for an MPNN with $M = 2$ and zirconia at 300 K in the monoclinic phase (top) and 1400 K in the tetragonal phase (bottom) for a simulation cell with 768 atoms. Shaded regions indicate standard error across 11 trajectories.

FIG. 3. Thermal conductivity across temperatures computed with an MPNN using $M = 2$ message-passing steps and experimentally determined lattice parameters [74,75], compared with another MLP without extrapolation [50], size-extrapolated *ab initio* GK [12], and experimental measurements [66–68]. Error bars are shown as given in the respective publications; the ones in the present work reflect the standard error across 11 trajectories. Letters "t" and "m" indicate results for the tetragonal and monoclinic phases, respectively.

$N = 1500$ simulation cell. These settings yield fully size- and time-converged results (see the Supplemental Material [56] for details).

The results presented in Fig. 3 are in good agreement with both experimental measurements in the monoclinic phase and theoretical MLP predictions in the monoclinic and tetragonal phases. As this work uses similar lattice parameters and the same exchange-correlation functional as the work by Verdi *et al.* [50], the observed close agreement is to be expected. Remaining differences between the MLP results may be due to the larger simulation cells used in the present work, enabled by the favorable scaling of computational cost due to the efficient heat flux implementation, and the semilocal nature of the employed MPNN. Compared to experiment, both MLPs are found to systematically underestimate κ by approximately 10% to 20%, which may be related to the intrinsic approximation of a finite-range MLP or the underlying density functional approximation. Larger differences are observed with the size-extrapolated *ab initio* GK results reported by Carbogno *et al.* [12], which, however, were computed for the tetragonal phase at all temperatures. Additionally, due to the high computational cost of first-principles calculations, only three trajectories of 60 ps each were used, which is reflected in the larger statistical error.

We conclude that the adapted GK approach for semilocal MLPs introduced in this work can successfully and efficiently predict the thermal conductivity of zirconia across temperatures, using 10 000 first-principles calculations in total. Despite a moderate system size of 96 atoms for training, fully size-converged results were obtained without requiring additional extrapolation schemes.

In summary, we have demonstrated the feasibility of applying AD-based semilocal MLPs to the prediction of thermal conductivities with the GK method. For this, we investigated the impact of semilocal interactions on the heat flux and derived an adapted heat flux $\mathbf{J}^{\mathrm{unfolded}}_{\mathrm{pot}}$ that can be efficiently implemented via AD. This heat flux has an asymptotically linear runtime and requires no further restrictions on the form of the potential. Its formulation is independent of the body-order of the potential-energy function, making no distinction between pair, angle-dependent, and many-body potentials. As it relies on explicitly constructing an extended simulation cell, it is applicable to semilocal MLPs with moderate effective interaction ranges.

Data and code required to reproduce all figures can be found at Zenodo [76]. First-principles calculations for the training data are additionally available from the NOMAD repository [77]. Further information and software can be found in Refs. [56,78].

[1] A. Evans, D. Clarke, and C. Levi, The influence of oxides on the performance of advanced gas turbines, J. Eur. Ceram. Soc. **28**, 1405 (2008).

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

[3] 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 (London) **508**, 373 (2014).

[4] F. Knoop, T. A. R. Purcell, M. Scheffler, and C. Carbogno, Anharmonicity measure for materials, Phys. Rev. Mater. **4**, 083809 (2020).

[5] L. Onsager, Reciprocal relations in irreversible processes. I, Phys. Rev. **37**, 405 (1931).

[6] L. Onsager, Reciprocal relations in irreversible processes. II, Phys. Rev. **38**, 2265 (1931).

[7] M. S. Green, Markoff random processes and the statistical mechanics of time-dependent phenomena, J. Chem. Phys. **20**, 1281 (1952).

[8] 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 (1957).

[9] R. Kubo, M. Yokota, and S. Nakajima, Statistical-mechanical theory of irreversible processes. II. Response to thermal disturbance, J. Phys. Soc. Jpn. **12**, 1203 (1957).

[10] A. M. Teale *et al.*, DFT exchange: Sharing perspectives on the workhorse of quantum chemistry and materials science, Phys. Chem. Chem. Phys. **24**, 28700 (2022).

[11] A. Marcolongo, P. Umari, and S. Baroni, Microscopic theory and quantum simulation of atomic heat transport, Nat. Phys. **12**, 80 (2016).

[12] C. Carbogno, R. Ramprasad, and M. Scheffler, *Ab Initio* Green-Kubo Approach for the Thermal Conductivity of Solids, Phys. Rev. Lett. **118**, 175901 (2017).

[13] A. Marcolongo, L. Ercole, and S. Baroni, Gauge fixing for heat-transport simulations, J. Chem. Theory Comput. **16**, 3352 (2020).

[14] F. Knoop, M. Scheffler, and C. Carbogno, *Ab initio* Green-Kubo simulations of heat transport in solids: Method and implementation, Phys. Rev. B **107**, 224304 (2023).

[15] M. González, Force fields and molecular dynamics simulations, J. Neutronique **12**, 169 (2011).

[16] J. B. Witkoskie and D. J. Doren, Neural network models of potential energy surfaces: Prototypical examples, J. Chem. Theory Comput. **1**, 14 (2005).

[17] S. Lorenz, A. Groß, and M. Scheffler, Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks, Chem. Phys. Lett. **395**, 210 (2004).

[18] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, Phys. Rev. Lett. **98**, 146401 (2007).

[19] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, Phys. Rev. Lett. **104**, 136403 (2010).

[20] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Machine learning force fields, Chem. Rev. **121**, 10142 (2021).

[21] I. Poltavsky and A. Tkatchenko, Machine learning force fields: Recent advances and remaining challenges, J. Phys. Chem. Lett. **12**, 6551 (2021).

[22] E. Prodan and W. Kohn, Nearsightedness of electronic matter, Proc. Natl. Acad. Sci. USA **102**, 11635 (2005).

[23] A. K. Rappe and W. A. Goddard, Charge equilibration for molecular dynamics simulations, J. Phys. Chem. **95**, 3358 (1991).

[24] T. A. Halgren and W. Damm, Polarizable force fields, Curr. Opin. Struct. Biol. **11**, 236 (2001).

[25] P. Cieplak, F.-Y. Dupradeau, Y. Duan, and J. Wang, Polarization effects in molecular mechanical force fields, J. Phys.: Condens. Matter **21**, 333102 (2009).

[26] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin, The ReaxFF reactive force-field: Development, applications and future directions, npj Comput. Mater. **2**, 15011 (2016).

[27] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, General-purpose machine learning potentials capturing nonlocal charge transfer, Acc. Chem. Res. **54**, 808 (2021).

[28] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, Neural message passing for quantum chemistry, in *Proceedings of the 34th International Conference on Machine Learning (ICML 2017)* (PMLR, 2017), pp. 1263–1272.

[29] K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, SchNet: A continuous-filter convolutional neural network for modeling quantum interactions, in *Advances in Neural Information Processing Systems 30 (NIPS 2017)* (Curran Associates, Red Hook, NY, 2017), pp. 992–1002.

[30] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, SchNet—A deep learning architecture for molecules and materials, J. Chem. Phys. **148**, 241722 (2018).

[31] O. T. Unke and M. Meuwly, PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges, J. Chem. Theory Comput. **15**, 3678 (2019).

[32] J. Klicpera, J. Groß, and S. Günnemann, Directional message passing for molecular graphs, in *Proceedings of the Eighth International Conference on Learning Representations (ICLR 2020)* (OpenReview, 2020).

[33] J. Klicpera, F. Becker, and S. Günnemann, GemNet: Universal directional graph neural networks for molecules, in *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)* (Curran Associates, Red Hook, NY, 2021), pp. 6790–6802.

[34] O. T. Unke, S. Chmiela, M. Gastegger, K. T. Schütt, H. E. Sauceda, and K.-R. Müller, SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects, Nat. Commun. **12**, 7273 (2021).

[35] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials, Nat. Commun. **13**, 2453 (2022).

[36] I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner, and G. Csányi, MACE: Higher order equivariant message passing neural networks for fast and accurate force fields, in *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)* (Curran Associates, Red Hook, NY, 2022), pp. 11423–11436.

[37] I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, The design space of E(3)-equivariant atom-centered interatomic potentials, arXiv:2205.06643.

[38] A. Bochkarev, Y. Lysogorskiy, C. Ortner, G. Csányi, and R. Drautz, Multilayer atomic cluster expansion for semilocal interactions, Phys. Rev. Res. **4**, L042019 (2022).

[39] T. Frank, O. Unke, and K.-R. Müller, So3krates: Equivariant attention for interactions on arbitrary length-scales in molecular systems, in *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)* (Curran Associates, Red Hook, NY, 2022), pp. 29400–29413.

[40] S. Stocker, J. Gasteiger, F. Becker, S. Günnemann, and J. T. Margraf, How robust are modern graph neural network potentials in long and hot molecular dynamics simulations? Mach. Learn.: Sci. Technol. **3**, 045010 (2022).

[41] O. T. Unke, M. Stöhr, S. Ganscha, T. Unterthiner, H. Maennel, S. Kashubin, D. Ahlin, M. Gastegger, L. Medrano Sandonas, A. Tkatchenko, and K.-R. Müller, Accurate machine learned quantum-mechanical force fields for biomolecular simulations, arXiv:2205.08306.

[42] C. Chen and S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, Nat. Comput. Sci. **2**, 718 (2022).

[43] 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).

[44] C. Mangold, S. Chen, G. Barbalinardo, J. Behler, P. Pochet, K. Termentzidis, Y. Han, L. Chaput, D. Lacroix, and D. Donadio, Transferability of neural network potentials for varying stoichiometry: Phonons and thermal conductivity of Mn_xGe_y compounds, J. Appl. Phys. **127**, 244901 (2020).

[45] 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).

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

[47] R. Li, E. Lee, and T. Luo, A unified deep neural network potential capable of predicting thermal conductivity of silicon in different phases, Mater. Today Phys. **12**, 100181 (2020).

[48] R. Li, Z. Liu, A. Rohskopf, K. Gordiz, A. Henry, E. Lee, and T. Luo, A deep neural network interatomic potential for studying thermal conductivity of β-Ga_2O_3, Appl. Phys. Lett. **117**, 152102 (2020).

[49] 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).

[50] C. Verdi, F. Karsai, P. Liu, R. Jinnouchi, and G. Kresse, Thermal transport and phase transitions of zirconia by on-the-fly machine-learned interatomic potentials, npj Comput. Mater. **7**, 156 (2021).

[51] A. Griewank and A. Walther, *Evaluating Derivatives* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008).

[52] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: A survey, J. Mach. Learn. Res. **18**, 1 (2018).

[53] S. Fabris, A. T. Paxton, and M. W. Finnis, Free energy and molecular dynamics calculations for the cubic-tetragonal phase transition in zirconia, Phys. Rev. B **63**, 094101 (2001).

[54] C. Carbogno, C. G. Levi, C. G. Van de Walle, and M. Scheffler, Ferroelastic switching of doped zirconia: Modeling and understanding from first principles, Phys. Rev. B **90**, 144109 (2014).

[55] R. J. Hardy, Energy-flux operator for a lattice, Phys. Rev. **132**, 168 (1963).

[56] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.108.L100302 for a brief introduction to AD and MPNNs; the derivation of 2; discussion of heat flux in solids; additional details on the derivation of Eqs. (5), (7), and (8); details of the implementation of 8; details of hyperparameter selection, training, and testing of the potential, including test set errors, phonon band structures, and the vibrational density of states; convergence of κ with simulation cell size, simulation duration, and noise removal settings; and details of the Green-Kubo workflow. It contains Refs. [79–98].

[57] Y. Chen, Local stress and heat flux in atomistic systems involving three-body forces, J. Chem. Phys. **124**, 054113 (2006).

[58] N. C. Admal and E. B. Tadmor, Stress and heat flux for arbitrary multibody potentials: A unified framework, J. Chem. Phys. **134**, 184106 (2011).

[59] D. Torii, T. Nakano, and T. Ohara, Contribution of inter- and intramolecular energy transfers to heat conduction in liquids, J. Chem. Phys. **128**, 044504 (2008).

[60] Z. Fan, L. F. C. Pereira, H.-Q. Wang, J.-C. Zheng, D. Donadio, and A. Harju, Force and heat current formulas for many-body potentials in molecular dynamics simulations with applications to thermal conductivity calculations, Phys. Rev. B **92**, 094301 (2015).

[61] D. Surblys, H. Matsubara, G. Kikugawa, and T. Ohara, Application of atomic stress to compute heat flux via molecular dynamics for systems with many-body interactions, Phys. Rev. E **99**, 051301 (2019).

[62] P. Boone, H. Babaei, and C. E. Wilmer, Heat flux for many-body interactions: Corrections to LAMMPS, J. Chem. Theory Comput. **15**, 5579 (2019).

[63] M. F. Langer, J. T. Frank, and F. Knoop, Stress and heat flux via automatic differentiation, arXiv:2305.01401.

[64] A. P. Thompson, S. J. Plimpton, and W. Mattson, General formulation of pressure and stress tensor for arbitrary many-body interaction potentials under periodic boundary conditions, J. Chem. Phys. **131**, 154107 (2009).

[65] A. Kinaci, J. B. Haskins, and T. Çağın, On calculation of thermal conductivity from Einstein relation in equilibrium molecular dynamics, J. Chem. Phys. **137**, 014106 (2012).

[66] S. Raghavan, H. Wang, R. B. Dinwiddie, W. D. Porter, and M. J. Mayo, The effect of grain size, porosity and yttria content on the thermal conductivity of nanocrystalline zirconia, Scr. Mater. **39**, 1119 (1998).

[67] J.-F. Bisson, D. Fournier, M. Poulain, O. Lavigne, and R. Mévrel, Thermal conductivity of yttria-zirconia single crystals, determined with spatially resolved infrared thermography, J. Am. Ceram. Soc. **83**, 1993 (2000).

[68] R. Mévrel, J.-C. Laizet, A. Azzopardi, B. Leclercq, M. Poulain, O. Lavigne, and D. Demange, Thermal diffusivity and conductivity of $Zr_{1-x}Y_xO_{2-x/2}$ ($x = 0, 0.084$ and 0.179) single crystals, J. Eur. Ceram. Soc. **24**, 3081 (2004).

[69] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Ab initio* molecular simulations with numeric atom-centered orbitals, Comput. Phys. Commun. **180**, 2175 (2009).

[70] F. Knoop, T.A.R. Purcell, M. Scheffler, and C. Carbogno, FHI-vibes: *Ab initio* vibrational simulations, J. Open Source Software **5**, 2671 (2020).

[71] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces, Phys. Rev. Lett. **100**, 136406 (2008).

[72] K. T. Schütt, P. Kessel, M. Gastegger, K. A. Nicoli, A. Tkatchenko, and K.-R. Müller, SchNetPack: A deep learning toolbox for atomistic systems, J. Chem. Theory Comput. **15**, 448 (2019).

[73] K. Tolborg and A. Walsh, Exploring the high-temperature stabilization of cubic zirconia from anharmonic lattice dynamics, Cryst. Growth Des. **23**, 3314 (2023).

[74] R. Patil and E. Subbarao, Axial thermal expansion of ZrO_2 and HfO_2 in the range room temperature to $1400\,^\circ C$, J. Appl. Crystallogr. **2**, 281 (1969).

[75] E. H. Kisi and C. Howard, Crystal structures of zirconia phases and their inter-relation, Key Eng. Mater. **153–154**, 1 (1998).

[76] M. F. Langer, F. Knoop, C. Carbogno, M. Scheffler, and M. Rupp, Data and software: Heat flux for semi-local machine-learning potentials, Zenodo (2023), https://doi.org/10.5281/zenodo.7767432.

[77] https://doi.org/10.17172/NOMAD/2023.03.24-2.

[78] https://marcel.science/gknet.

[79] F. Knuth, C. Carbogno, V. Atalla, V. Blum, and M. Scheffler, All-electron formalism for total energy strain derivatives and stress tensor components for numeric atom-centered orbitals, Comput. Phys. Commun. **190**, 33 (2015).

[80] N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, and P. Riley, Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds, arXiv:1802.08219.

[81] B. Anderson, T.-S. Hy, and R. Kondor, Cormorant: Covariant molecular neural networks, in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)* (Curran Associates, Red Hook, NY, 2019), pp. 14537–14546.

[82] K. T. Schütt, O. T. Unke, and M. Gastegger, Equivariant message passing for the prediction of tensorial properties and molecular spectra, in *Proceedings of the 38th International Conference on Machine Learning* (ICML 2021), pp. 9377–9388.

[83] J. J. Erpenbeck, Einstein-Kubo-Helfand and McQuarrie relations for transport coefficients, Phys. Rev. E **51**, 4296 (1995).

[84] R. J. Hardy, Formulas for determining local properties in molecular-dynamics simulations: Shock waves, J. Chem. Phys. **76**, 622 (1982).

[85] W. Noll, Die Herleitung der Grundgleichungen der Thermomechanik der Kontinua aus der statistischen Mechanik, J. Ration. Mech. Anal. **4**, 627 (1955).

[86] J. H. Irving and J. G. Kirkwood, The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics, J. Chem. Phys. **18**, 817 (1950).

[87] A. Ito and H. Nakamura, Energy current on multi-body potential with Dirac delta function, Prog. Theor. Phys. Suppl. **178**, 107 (2009).

[88] N. Chetty and R. M. Martin, First-principles energy density and its applications to selected polar surfaces, Phys. Rev. B **45**, 6074 (1992).

[89] A. J. Ladd, B. Moran, and W. G. Hoover, Lattice thermal conductivity: A comparison of molecular dynamics and anharmonic lattice dynamics, Phys. Rev. B **34**, 5058 (1986).

[90] L. Isaeva, G. Barbalinardo, D. Donadio, and S. Baroni, Modeling heat transport in crystals and glasses from a unified lattice-dynamical approach, Nat. Commun. **10**, 3853 (2019).

[91] A. McGaughey and M. Kaviany, Phonon transport in molecular dynamics simulations: Formulation and thermal conductivity prediction, Adv. Heat Transfer **39**, 169 (2006).

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

[93] P. C. Howell, Comparison of molecular dynamics methods and interatomic potentials for calculating the thermal conductivity of silicon, J. Chem. Phys. **137**, 224111 (2012).

[94] Data set PRL_2017_aiGK in the NOMAD repository, https://doi.org/10.17172/NOMAD/2017.04.13-1.

[95] C. Draxl and M. Scheffler, The NOMAD laboratory: From data sharing to artificial intelligence, J. Phys. Mater. **2**, 036001 (2019).

[96] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.

[97] G. J. Martyna, D. J. Tobias, and M. L. Klein, Constant pressure molecular dynamics algorithms, J. Chem. Phys. **101**, 4177 (1994).

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