

ARTICLE OPEN

Empirical interatomic potentials optimized for phonon properties

Andrew Rohskopf¹, Hamid R. Seyf¹, Kiarash Gordiz¹, Terumasa Tadano² and Asegun Henry^{1,3,4}

Molecular dynamics simulations have been extensively used to study phonons and gain insight, but direct comparisons to experimental data are often difficult, due to a lack of accurate empirical interatomic potentials for different systems. As a result, this issue has become a major barrier to realizing the promise associated with advanced atomistic-level modeling techniques. Here, we present a general method for specifically optimizing empirical interatomic potentials from ab initio inputs for the study of phonon transport properties, thereby resulting in phonon optimized potentials. The method uses a genetic algorithm to directly fit the empirical parameters of the potential to the key properties that determine whether or not the atomic level dynamics and most notably the phonon transport are described properly.

npj Computational Materials (2017)3:27; doi:10.1038/s41524-017-0026-y

INTRODUCTION

Over the last 25 years, the usage of molecular dynamics (MD) simulations to study phonons has grown markedly. The reason MD is a useful tool for studying phonons is because of three primary advantages/features over other methods: (1) it naturally includes anharmonicity to full order, (2) it naturally includes the full atomic level details of the structure (i.e., composition, defects, boundaries, etc.), and (3) classical MD can access the necessary time (10^{-6} – 10^2 ns) and length scales (10^{-2} to 10^3 nm) using today's high-performance computing hardware.^{1, 2} The last point is a great advantage of classical MD over ab initio MD. However, given the growing interest and usage of classical MD to study phonons, there has not been a correspondingly large increase in the number and fidelity of direct comparisons to experimental data. This has largely been because of the lack of suitable and accurate empirical interatomic potentials (EIPs) that can describe the various interatomic interactions involved in the actual systems being measured.^{3–6} Here it should be emphasized that the EIP is the most important aspect of a classical MD simulation, because it contains all the physics and in essence, a classical MD simulation is merely a way of sampling the EIP when the atoms reside in the phase space that the EIP was designed for. Another reason is that it is difficult to compare MD results directly to experiments because there is often insufficient detail known about the atomic level structure of the samples being measured to facilitate construction of an accurately representative MD supercell. Nonetheless, it is because of these challenges that fair and rigorous comparisons between MD simulation data and experiments are lacking, and this has stifled the ability for theorists to explain and/or predict material properties and anomalous behaviors observed in experiments, which would further our understanding.^{6–10}

Given the reality of these barriers to scientific advancement, the development of a means by which one can quickly, easily and accurately create EIPs for the purposes of studying phonon

transport has become a grand challenge for the field. Here, the notion that one should be able to quickly and easily parameterize EIPs is important to emphasize, as there have been major advances over the last 25 years that enable the creation of accurate EIPs.^{11–15} Most notably, these advances include the proliferation of first principles methods, such as density functional theory (DFT), which can be used to generate data that interatomic potentials can be fit to reproduce. In the past, many efforts to produce such EIPs involved modification of the functional form itself and the properties used for fitting often included experimental measurements of the lattice parameters, elastic constants, phonon frequencies, and other measurable quantities.^{16, 17} More recently, efforts to fit EIP parameters have shifted to ab initio data, which offers a more direct connection to atomistic level quantities, such as forces, energies, and stresses on supercells, which cannot be easily determined experimentally.^{14, 18} However, the N-dimensional optimization problem to find the best EIP parameters is still daunting, and in the past has been partially guided by chemical/physical intuition into the system of interest. For this reason, many popular functional forms for EIPs have been modified for different systems to achieve improved accuracy.^{19, 20} Consequently, many MD investigations resort to using whatever parameters can be found somewhere in the literature for an EIP that has already been coded, and have been applied to the specific system of interest. Furthermore, this usage of standard EIPs and parameters from literature often happens regardless of whether the EIP accurately describes the phonon transport or not, simply due to a lack of options. Thus, the grand challenge has been to develop a quick and easy method for creating EIPs that accurately describe phonons. The emphasis here on the process being quick and easy, is so that the EIPs can be created and employed with minimal effort, and the major intellectual investment can remain focused on the MD and phonon transport instead of the prerequisite issue of finding a suitable EIP. It should

¹George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA; ²Center for Materials Research by Information Integration, National Institute for Materials Science, Tsukuba 305-0047, Japan; ³School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA and ⁴Heat Lab, Georgia Institute of Technology, Atlanta, GA 30332, USA
Correspondence: Asegun Henry (ase@gatech.edu)

Received: 30 December 2016 Revised: 18 May 2017 Accepted: 8 June 2017

Published online: 12 July 2017

be noted that a Taylor expansion of the potential energy surface offers excellent agreement with phonon dispersion, but has been shown to be unstable in MD simulations at high temperatures.^{21, 22} The addition of higher order terms in a Taylor expansion potential does still not ensure stability at temperatures much higher than room temperature,²¹ and it is therefore of interest to parameterize other functional forms that yield stable MD simulations at high temperatures. In this article, we describe a methodology and a set of freely available codes that implement the methodology that can overcome this barrier of generating what we have termed phonon optimized potentials (POPs).²³

Before explaining the methodology itself it is useful to first highlight its goals—the basic tenets that underpin it, and several questions we seek to answer with the first example usage of it, which is described later:

Tenet 1: Many EIP functional forms are typically overdesigned for the study of phonons by possessing features and flexibility that allow for the study of regions of phase space beyond thermal vibrations around equilibrium. In optimizing the parameters of such functional forms, we therefore hypothesize that many solutions may exist to describe the properties of interest, which are associated with a reduced portion of phase space. Popular EIPs are designed to describe various configurations of the atoms, and most notably different atomic coordinations.^{24–26} However, it is most often the case that when one seeks to study phonons, all atoms by definition vibrate around their equilibrium sites²⁷ and the atomic coordination and configuration are fixed for all atoms throughout the entire MD simulation.²⁸ We therefore say that such EIPs are often overdesigned for the study of phonons, as their design goes beyond simple vibrations about equilibrium. Given this great overdesign of standard EIPs to describe unnecessary regions of phase space for the study of phonons, we hypothesize that many parameter sets exist for such EIPs that describe phonons, but may not describe other properties. Nonetheless, since the focus herein is to optimize for phonons, such EIPs would still be considered accurate for the purposes herein, despite their limited transferability.

Tenet 2: In order for an EIP to be optimized for describing phonons, the key quantities that must be well described are the total energy and its derivatives. Taking results from the fluctuation-dissipation theorem^{29, 30} as a basis for describing phonon transport properties, such as thermal conductivity^{28, 30} and interface conductance,³¹ this tenet is based on the idea that one will obtain the correct transport properties if all of the forces (including correct individual force components³²) and velocities of the atoms are correct. Also, from the formalism developed for crystalline thermal conductivity it is known that if one can correctly compute all the derivatives of the energy with respect to atomic displacements, one should theoretically properly describe phonon-phonon interactions.³³ Therefore an EIP is optimized for phonons when it accurately reproduces the derivatives of the potential energy with respect to the atomic displacements. Although, in concept one would need an infinite number of derivatives to be exact, we note that in practice only the energy and its first three or four derivatives ($E, \frac{dE}{dr_i}, \frac{d^2E}{dr_i dr_j}, \frac{d^3E}{dr_i dr_j dr_k}$) are actually needed for most systems/temperatures, but higher order terms can be included as deemed necessary.³⁴ Here, it is also important to clarify that the goal of POPs is to make EIPs that replicate the results of *ab initio* calculations and not necessarily experiments. In this sense the goal is to make POPs based purely on first principles data, thereby enabling them with predictive power.

Tenet 3: Assuming the preceding tenet is correct, then one can set as components of an objective function, the relative error in energy and its derivatives to assess the viability of a potential to describe phonon properties. This then provides a universal scale upon which any EIP can be assessed. For example, one can assess that a given EIP reproduces the energy of the *ab initio* model within 3%, the forces within 5%, the second derivatives within 10% on average, and so on. In this respect one can also invoke

statistical techniques as well (i.e., standard deviation, root mean squared etc.) to better assess the EIP accuracy. It may also happen that a certain functional form is unable to reduce the error in energy and its derivatives to say below ~40%. Such a functional form may therefore not be suitable for the system of interest and it is important that the methodology be able to determine this. It is therefore a goal of the methodology to enable assessment of the suitability of the functional form itself.

Tenet 4: A major goal of the POPs approach is to make the creation of EIPs for a given system easy and quick. Here, the term “easy” is to imply that minimal input, chemical insight and management on the part of the user is required and instead the procedure itself is capable of handling most of the effort. Furthermore, it is highly desirable to have the approach be built in such a way, that it does not require new coding on the part of the user, which inherently requires time for debugging and is a strong deterrent.

Considering the aforementioned goals/tenets and hypotheses, several questions arise that we seek to answer herein, namely: (1) Considering the aforementioned hypothesis in Tenet 1 regarding the possibility of finding multiple/many sets of parameters that nearly degenerately minimize the target objective function, it is not clear if all of such solutions will exhibit commonalities, e.g., they can somehow be reduced to a single unique best solution, or if some are uniquely different, exhibiting drastically different parameters that somehow still yield similar objective function values. (2) It is not clear a priori according to Tenet 3 whether common and standard EIP functional forms will be able to accurately reproduce *ab initio* results at all.

Herein we seek to determine if the proposed approach can actually yield useful POPs that can at least reproduce thermal conductivity within ~10%. Considering the four Tenets, we seek to create a generalized user-friendly code and method that can create POPs with ease. Due to the drastic nonlinearity in the search space of EIP parameters, we choose to use a genetic algorithm (GA). More details on the optimization code and procedure are found in the Methods section and the SI.

RESULTS

With the preceding framework implemented we then tested the code on the two most popular example systems that are of very high technological/applications interest, namely crystalline silicon (c-Si) and crystalline germanium (c-Ge). For many semiconductors it is well known that long range interactions are important³⁵ yet the most popular EIPs such as Tersoff,³⁶ Stillinger–Weber (SW),³⁷ the environment-dependent interatomic potential²⁴ and others are restricted to first nearest neighbors. Thus, to illustrate the power of the POPs methodology and to describe c-Si and c-Ge more accurately we fit the *ab initio* data with a combination of short ranged and long ranged functional forms and procedures described in the Methods section and the SI. One of the most interesting and important outputs of the fitting procedure, which was executed in less than 1 day using four processors per trial, was that the various fits yielded >50 uniquely different parameter sets that all had less than 10% error in forces, energies, and stresses (Tenet 1). Figure 1 shows the objective function convergence with generations for all 50 trials in three different EIPs. Three EIPs considered here were: (1) the SW potential, added to the Born potential and a long-ranged Coulomb potential, abbreviated SWBC; (2) the Tersoff potential, added to the Born potential and a long-ranged Coulomb potential, abbreviated TBC; (3) the Morse potential added to an explicit three-body angular term, as well as the Born potential and long-ranged coulomb potential, abbreviated M3BC. The results in Fig. 1, highlight that the GA-based approach was able to quickly exhaust the options associated with the different EIPs (Tenet 3), and although the TBC and M3BC potentials were able to minimize the objective function

substantially, the SWBC potential was not. This illustrated that the SWBC functional form itself is simply incapable of properly describing the system of interest.

After being optimized, these parameter sets were then tested against 50 configurations that were randomly displaced by a maximum of 0.05 Å, for which they had not been fit to, and their error in forces, energies and stresses was <10%. Thus, they can be substituted for direct DFT calculations to yield the same forces energies and stresses for an arbitrary configuration to within 10%. Here, the importance of the nominal value of ~10% error is illustrated in Fig. 2, as one can visually see the difference between predicted and fitted forces when the error is 50%, vs. when it is below 10%. As a point of comparison, consider the differences in forces for the same configurations using different pseudopotentials, which subsequently result in ~15% standard deviation in thermal conductivity.³⁸ Thus, errors much greater than 10% are quite substantial and are larger than the differences one would expect to observe from improving the ab initio calculations. Consequently, even though these popular EIPs have been extensively used to model c-Si and c-Ge, the dynamics are not

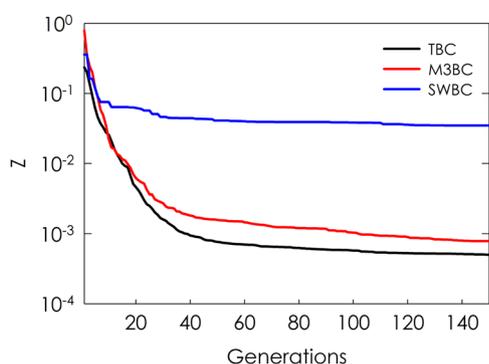


Fig. 1 Average Z value convergence with generations. Each line in this figure represents the average Z value (Eq. 1) per generation, across all 100 trials performed on the TBC, M3BC, and SWBC potentials. The following weights were used for each Z component: $w_f = 0.15$, $w_e = 0.25$, $w_s = 0.2$, $w_{ifc-2} = 0.2$, and $w_{ifc-3} = 0.2$. This type of analysis shows that some functional forms are able to simultaneously reproduce certain properties better than others. The SWBC, which on average only experienced an order of magnitude decrease from a random guess in Z, could not reproduce all quantities at once. The TBC and M3BC potentials were able to decrease Z by 3 orders of magnitude from a random guess. The small final discrepancy between TBC and M3BC is due to the fact that TBC better reproduced the 3rd order IFCs, which also led to better thermal conductivity agreement as seen with Figs. 2 and 3

representative of the real materials. Instead, what has been modeled is likely a fictitious material that happens to have similar thermal conductivity, but not the same phonon trajectories as Si and Ge. For perspective, it is important to realize that we found that the most common parameter sets used for the Tersoff³⁶ and SW³⁷ potentials to describe c-Si result in 35 and 210% error in forces, respectively.

DISCUSSION

Clearly the 50% level of error in Fig. 2 would lead to incorrect dynamics and would result in unacceptably large errors in phonon transport properties. However, conceptually a < 10% error in forces, might translate to ~10% error in the heat flux operator³² and therefore errors in thermal conductivity²⁸ and interface conductance³¹ on the order of 10% might not be surprising. Nonetheless, one would also expect that the relationships between force error and thermal conductivity could be quite nuanced and complicated. Other authors have experienced situations where a potential can obtain proper phonon dispersion but incorrect thermal conductivity.³⁹ We note that this stresses the importance of fitting to both the 2nd and higher order interatomic force constants (IFCs), since the 2nd order IFCs determine the phonon dispersion, while the higher order IFCs are also needed to determine the thermal conductivity. The validity of an approximately 10% error in force, energy and stress metric, along with <10% errors in IFCs is consistent with estimations of the thermal conductivity for c-Si using the Boltzmann transport equation within the relaxation time approximation as implemented in Alamode.⁴⁰ Figures 3 and 4 show the excellent agreement between the DFT derived thermal conductivity that was obtained when fitting to energies, forces, stresses, and 2nd and 3rd order IFCs employing the TBC (<5% thermal conductivity error) and M3BC (<15% thermal conductivity error) potentials for c-Si and c-Ge. It is interesting to note here that the SWBC potential was unable to simultaneously minimize all parts of the objective function to within 10% error of DFT, as seen in the poor fit displayed in Fig. 1. The SWBC resulted in unstable dynamics and/or much larger (>50% errors in IFCs) disagreement with phonon transport properties, such as dispersion or forces. These results answer question (1), namely whether or not multiple parameter sets can be determined that can accurately reproduce the trajectories that would have been obtained if a DFT MD calculation could be evaluated at the requisite length and time scales. We have also answered the question (2) as to whether or not some common functional forms (e.g., SWBC) simply fail to simultaneously describe all properties in the objective function, as shown in Fig. 1.

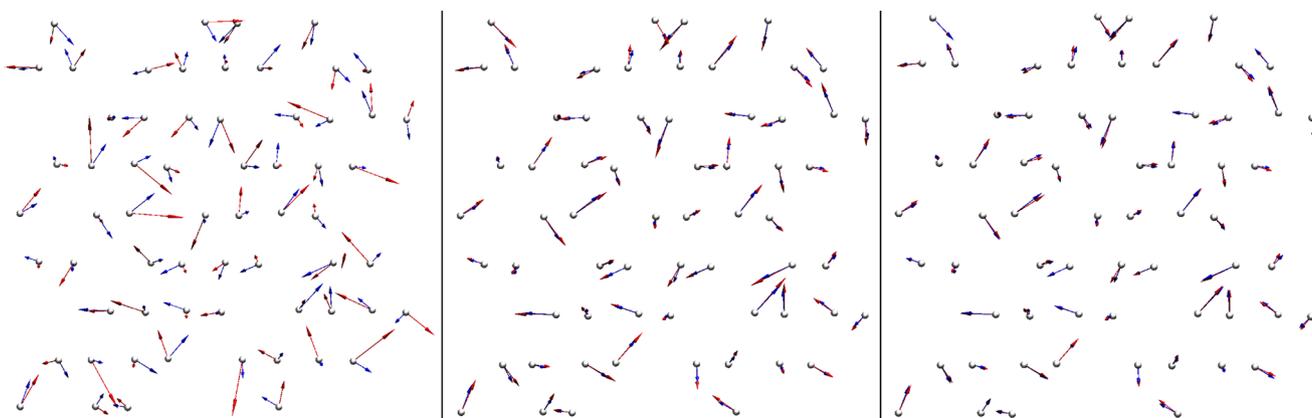


Fig. 2 Visual representation of different errors in forces. DFT forces (blue vectors) are shown on the atoms and are compared to EIP forces (red vectors) for 50%, 10% and 3% errors in forces from left to right

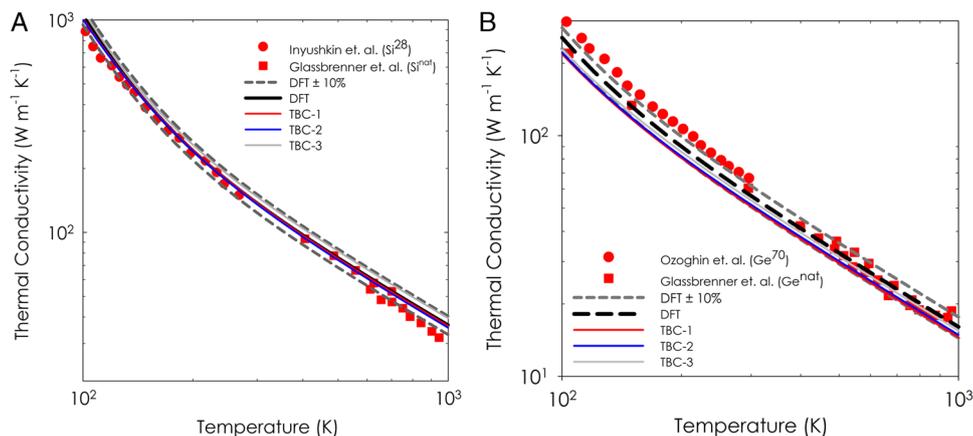


Fig. 3 TBC, DFT, and experimental^{59, 60} thermal conductivity vs. temperature. **a** Thermal conductivity as a function of temperature for c-Si with DFT, experimental,^{59, 60} and TBC POPs. **b** Thermal conductivity as a function of temperature for c-Ge with DFT, experimental,^{59, 61} and TBC POPs. The TBC POPs were able to consistently obtain thermal conductivities within 10% of DFT values

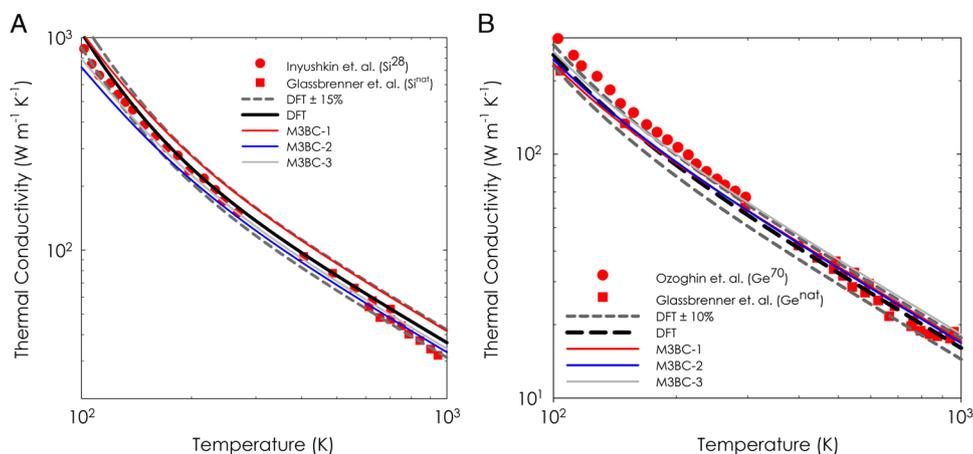


Fig. 4 M3BC, DFT, and experimental^{59, 60} thermal conductivity vs. temperature. **a** Thermal conductivity as a function of temperature for c-Si with DFT, experimental,^{59, 60} and M3BC POPs. **b** Thermal conductivity as a function of temperature for c-Ge with DFT, experimental,^{59, 61} and M3BC POPs. While M3BC-3 is within 10% of the DFT thermal conductivity for c-Si, other M3BC potentials were not able to reach this mark. DFT thermal conductivity $\pm 15\%$ is therefore displayed to show the performance of the M3BC potential for c-Si

For an EIP to be optimized for phonons it is important to fit to the IFCs. The 2nd order IFCs are more important to include since the harmonic components of forces usually comprise the overwhelming majority (>90%) of the forces and energy⁴¹. This information is not explicitly contained in the total forces and therefore it is important to separately include the 2nd order derivatives (e.g., the Hessian matrix) as part of the objective function, since we have observed cases where it is possible to have different EIPs reproduce the total forces correctly with drastically different force components. Phonon dispersion relations for various POPs are shown in Figs. 5 and 6 for c-Si and c-Ge. Figure 5a is an interesting case since it shows a qualitative error in phonon dispersion associated with increasing error in 2nd order IFCs, thus showing the importance of including the 2nd order IFCs in fitting to ensure the interactions between atoms are properly scaled and yield the correct dispersion. Figure 6 shows the fitted POPs in comparison to the most two most popular potentials used for silicon—Tersoff¹⁷ and SW.³³ It is seen that these standard potentials for silicon greatly overestimate the optical phonon frequencies as compared to the POPs. The main error in phonon frequencies with the TBC and M3BC potentials is seen with their failure to reproduce the flattening of frequencies near the zone boundaries, due to the inability of the pair potential to reproduce long range IFCs. This issue, however, has been overcome with

other analytical functional forms, such as the bond charge model⁴² and valence force field model,^{43, 44} which assume different interaction parameters for every pair. The issue of reproducing long range IFCs is also overcome by interatomic potentials which model the potential energy surface as dependent upon a smooth and flexible atomic neighborhood for every atom, such as the Gaussian approximated potential⁴⁵ and the spectral neighbor analysis potential.⁴⁶ Lastly, the issue of dispersion is also overcome by a Taylor expansion of the potential energy around the equilibrium positions. Such methods, could in the future be used in combination with more standard EIPs, to produce even more accurate POPs, that not only replicate forces, energies, stresses and thermal conductivities, but also more accurately reproduce dispersions.

We have developed a framework that consistently produces reliable POPs out of any suitable EIP or combination of EIPs from first principles inputs, by employing a GA to find parameters. In our first demonstration of the methodology we have answered two important rather fundamental questions regarding EIP fitting: (1) It was confirmed that common EIPs are overdesigned for the purposes of exclusively modeling phonon transport and thus many nearly degenerately performing solutions exist that have drastically different parameters. This finding is particularly important, because different solutions could be more transferrable

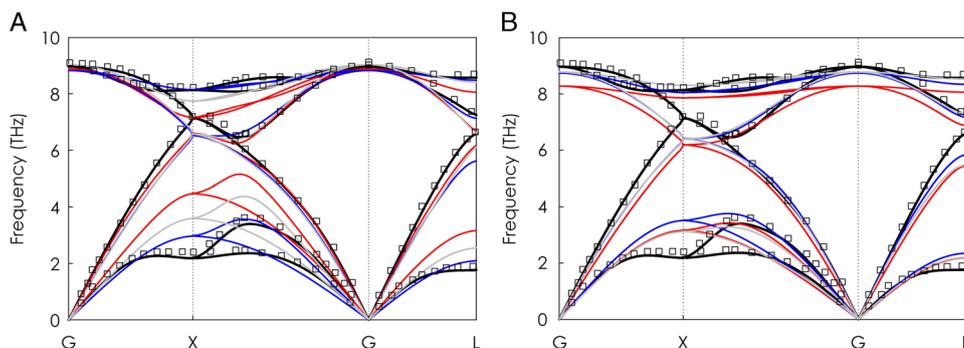


Fig. 5 DFT, and experimental⁶² phonon dispersion for Ge, along with different POPs. **a** DFT (black), M3BC-1 (red), M3BC-2 (blue), M3BC-3 (gray) and experiments (squares) are shown. These EIPs are shown to demonstrate the importance of 2nd order force constants in reproducing correct phonon dispersion. From best to worst agreement in 2nd order force constant error, M3BC-2 (5.1×10^{-4} Ry/Bohr² mean absolute error in 2nd order IFCs), M3BC-3 (9.7×10^{-4} Ry/Bohr² mean absolute error in 2nd order IFCs), and M3BC-1 (7.4×10^{-3} Ry/Bohr² mean absolute error in 2nd order IFCs) exhibit expected qualitative trends in phonon frequency error, especially at the zone boundaries where flattening is unable to occur. **b** DFT (black), TBC-1 (red), TBC-2 (blue), TBC-3 (gray) and experiments (squares) are shown

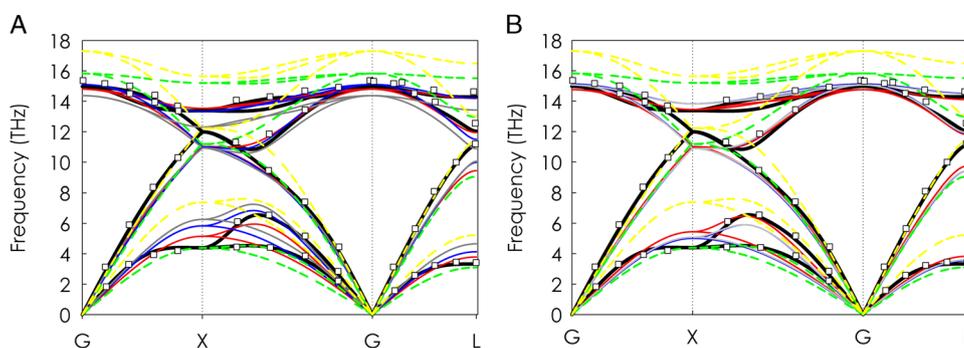


Fig. 6 DFT and experimental⁴² phonon dispersion for Si, along with various POPs. The original Tersoff¹⁷ and Stillinger–Weber³⁷ potentials are shown as yellow and green dotted lines, respectively. **a** DFT (black), TBC-1 (red), TBC-2 (blue), TBC-3 (gray) and experiments (squares) are shown. **b** DFT (black), M3BC-1 (red), M3BC-2 (blue), M3BC-3 (gray) and experiments (squares) are shown

or better/worse at describing other properties/phenomena of interest and thus it is useful to report all of such solutions, so that other users can determine the extent to which POPs parameterizations can be used for other properties. (2) We have confirmed that common functional forms for c-Si and c-Ge can in fact reproduce DFT results within $\sim 10\%$, and therefore serve as useful substitutes to enable probing of larger length and time scales accessible to DFT directly. Furthermore, the GA approach proved useful at performing sufficiently exhaustive searches through parameters, such that one can evaluate the suitability of the functional form itself for a given system.

Future work will involve using POPs for alloys and disordered systems in order to study thermal transport, since the asymmetry of these systems render them computationally intractable to be studied directly by current ab initio methods.³³ The creation of POPs will also allow for the study of thermal transport using MD, with much greater fidelity, enabling direct comparisons with experiments, when some of the most general methodologies are employed.^{28, 31} Also of critical importance is the fact that the POPs methodology can be used for fields outside of thermal transport. For example, functional forms such as REAXFF⁴⁷ can be used in the context of the study of chemical reaction kinetics, and common functional forms used herein can be optimized using the same algorithm to study defects, grain boundaries, interfaces and surfaces by first fitting to the most closely accessible DFT configurations. In this way, it is anticipated that the POPs methodology can serve as a significant advancement in many other areas of science/engineering, beyond that of thermal

transport, since phonons are important for many other non-heat transfer centered phenomena as well.⁴⁸

METHODS

The specific case of creating EIPs for c-Si and c-Ge is discussed here in detail, because it highlights important features and answers to the questions/hypotheses outlined in the Tenets. The POPs methodology described in the following has been implemented in a freely available C++ code⁴⁹ and can be run massively in parallel. The approach uses a GA to search for parameters that minimize an objective function representing some form of error between candidate EIPs and ab initio results. The GA is a metaheuristic that mimics natural selection to guide the algorithm towards a solution of a multi-dimensional problem.^{50–52} Gradient based methods work well for problems with less dimensions and much less non-linearity. However, for more complex problems with large numbers of dimensions and strongly non-linear behavior, gradient based methods easily fail and alternative schemes such as a GA are needed.^{12, 53} Since the GA approach itself inherently searches for random perturbations to potential solutions (i.e., via crossover and mutation), one is in practice guaranteed to exhaustively exploit many local minima if many trials are run in parallel. Many parallel trials also exhaust all possibilities for a given functional form, which allows one to determine its suitability for a given system (Tenet 3). For example, if many GA trials yield a high percentage of undesirable solutions, it can be said that the functional form may not be an appropriate candidate for the system of interest. Additional details associated with the POPs GA are described in the SI.

The code couples with the open source MD software LAMMPS⁵⁴ as a calculator for EIPs, and the open source code Alamode⁴⁰ as a calculator for IFCs. LAMMPS was selected because it has many standard EIPs already coded within it as well as many common variants. This prevents users from having to write new code to try different EIP functional forms, which

contributes strongly towards making the process “easy” (Tenet 4). Alamode was selected because it can take as inputs a series of arbitrary atomic displacements and forces to then compute 2nd, 3rd and 4th order derivatives of the energy with respect to the atomic displacements, as required by Tenets 2 and 3 to create a POP.⁴⁰ Using LAMMPS and Alamode as calculators for EIP properties, Tenet 3 requires that an objective function is needed to determine the viability of an EIP compared to reference values. In general, the objective function consists of different quantities that are weighted according to their relative importance on a normalized scale via Eq. 1:

$$Z = \sum_i w_i z_i, \quad (1)$$

where each z_i generically represents an error between the values produced by a candidate EIP as compared to the ab initio values, weighted by a factor w_i . In this work, Eq. 1 is a sum of the weighted errors of Hellman–Feynman forces, energies, stresses, and 2nd and 3rd order IFCs with weights w_{f_i} , w_{e_i} , w_{s_i} , w_{ifc-2} , and w_{ifc-3} , respectively. These quantities are inspired from the discussion in Tenets 2 and 3, and the stresses were found to be necessary to ensure crystal stability in MD simulations. The format of the error z_i for forces, energies, stresses, and IFCs is described in the SI.

Lastly the interface with ab initio results has been generalized so that users can employ any code of interest. The inputs are simply a series of supercell snapshots containing atomic coordinates, the total energy, the individual atom total forces, and supercell stresses for configurations with atoms randomly displaced from equilibrium positions as well as different volumes. Several previous works have indicated that using random displacements, or even displacements from ab initio MD trajectories, is highly effective in capturing anharmonicity.^{40, 55, 56} Each randomly displaced configuration contains information about many interatomic interactions and not just a subset of atoms as is the case for the direct displacement method.³³ As a result, it has been found that using random displacements reduces the search space by providing more valuable information on anharmonicity and other interactions that result in stability during a MD simulation.

We selected three popular short ranged EIPs: the SW, Tersoff, and the Morse⁵⁷ potential⁵⁸ as implemented in LAMMPS and we also added the Buckingham potential and damped-shifted force (DSF) Coulomb potential⁵⁷ to Tersoff and SW, and we added the Born–Mayer–Huggins⁵⁸ and DSF Coulomb EIPs to Morse. Since the Morse potential is two-body in nature, we added a harmonic three body term of the form $E = K(\theta - \theta_0)^2$ to account for the covalent three body interactions in c-Si and c-Ge. This then yielded three candidate functional forms: Tersoff + Buckingham + Coulomb (TBC), Stillinger–Weber + Buckingham + Coulomb (SWBC), and Morse + 3 Body + Born + Coulomb (M3BC)—all of which are already coded in LAMMPS. The ranges used for each parameter search and a general rationale for selecting the range is described in the SI.

These potentials were fit to snapshots of DFT calculated configurations of 64 atoms in a supercell using the POPs code, whereby it calls LAMMPS as a library to evaluate the energy and forces of candidate parameter sets, along with Alamode to determine the 2nd and 3rd order IFCs. The objective function in Eq. 1 was then minimized with the following relative weightings: $w_f = 0.15$, $w_e = 0.25$, $w_s = 0.2$, $w_{ifc-2} = 0.2$, and $w_{ifc-3} = 0.2$.

Data availability

The open-source POPS code and manual are located at www.pops.gatech.edu. The parameters for potentials used in this study are attached as separate files as supplementary material.

ACKNOWLEDGEMENTS

All DFT calculations and fitting procedures were run on the Partnership for Advanced Computing Environment (PACE) computer cluster at the Georgia Institute of Technology. We acknowledge support from the National Science Foundation through a Career Award (1554050) for A. Henry. T.T. was partly supported by JSPS KAKENHI Grant Number 16K17724 and “Materials research by Information Integration” Initiative (MI2I) project of the Support Program for Starting Up Innovation Hub from Japan Science and Technology Agency (JST).

AUTHOR CONTRIBUTIONS

A.R. created the generalized GA code that creates POPs, along with the method. H.R. S. and K.G. aided in the testing and refinement of the code and method, along with debugging. T.T., the author of Alamode, helped with the calculation of force

constants for use in fitting. A.H. provided insight into interatomic potential modeling, DFT calculations, and served as an advisor on this project.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Computational Materials* website (doi:10.1038/s41524-017-0026-y).

Competing interests: The authors declare that they have no competing financial interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Frenkel, D. & Smit, B. *Understanding Molecular Simulations: From Algorithms to Applications* (Academic, 1996).
- Chen, G. *Nanoscale Energy Transport and Conversion*. (Oxford University Press, 2005).
- Berber, S., Kwon, Y.-K. & Tománek, D. Unusually high thermal conductivity of carbon nanotubes. *Phys. Rev. Lett.* **84**, 4613–4616 (2000).
- Schelling, P. K., Phillpot, S. R. & Keblinski, P. Comparison of atomic-level simulation methods for computing thermal conductivity. *Phys. Rev. B* **65**, 144306 (2002).
- Volz, S. G. & Chen, G. Molecular-dynamics simulation of thermal conductivity of silicon crystals. *Phys. Rev. B* **61**, 2651–2656 (2000).
- He, Y., Savic, I., Donadio, D. & Galli, G. Lattice thermal conductivity of semi-conducting bulk materials: atomistic simulations. *Phys. Chem. Chem. Phys.* **14**, 16209–16222 (2012).
- Goicochea, J. V., Madrid, M. & Amon, C. Thermal properties for bulk silicon based on the determination of relaxation times using molecular dynamics. *J. Heat Transf.* **132**, 012401 (2010).
- Broido, D. A., Ward, A. & Mingo, N. Lattice thermal conductivity of silicon from empirical interatomic potentials. *Phys. Rev. B* **72**, 014308 (2005).
- Lee, J.-H. & Grossman, J. C. Thermoelectric properties of nanoporous Ge. *Appl. Phys. Lett.* **95**, 013106 (2009).
- Lee, J.-H., Galli, G. A. & Grossman, J. C. Nanoporous Si as an Efficient Thermoelectric Material. *Nano Lett.* **8**, 3750–3754 (2008).
- Malshe, M. et al. Parametrization of analytic interatomic potential functions using neural networks. *J. Chem. Phys.* **129**, 044111 (2008).
- Solomon, J., Chung, P., Srivastava, D. & Darve, E. Method and advantages of genetic algorithms in parameterization of interatomic potentials: metal oxides. *Comput. Mater. Sci.* **81**, 453–465 (2014).
- Shimada, T., Wakahara, K., Umeno, Y. & Kitamura, T. Shell model potential for PbTiO₃ and its applicability to surfaces and domain walls. *J. Phys. Condens. Matter* **20**, 325225 (2008).
- Brommer, P. & Gähler, F. Effective potentials for quasicrystals from ab-initio data. *Philos. Mag.* **86**, 753–758 (2006).
- Barrett, C. & Wang, L.-W. A systematic fitting procedure for accurate force field models to reproduce ab initio phonon spectra of nanostructures. *Comput. Phys. Commun.* **200**, 27–36 (2016).
- Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids*. (Oxford University Press, 1989).
- Tersoff, J. Empirical interatomic potential for silicon with improved elastic properties. *Phys. Rev. B* **38**, 9902–9905 (1988).
- Ercolessi, F. & Adams, J. B. Interatomic potentials from first-principles calculations: the force-matching method. *Europhys. Lett.* **26**, 583–588 (1994).
- El-Mellouhi, F., Sekkal, W. & Zaoui, A. A modified Tersoff potential for the study of finite temperature properties of BP. *Physica A* **311**, 130–136 (2002).
- Brenner, D. W. in *Computer Simulation of Materials at Atomic Level* 23–40 (Wiley-VCH Verlag GmbH & Co. KGaA, 2005).
- Murakami, T., Shiga, T., Hori, T., Esfarjani, K. & Shiomi, J. Importance of local force fields on lattice thermal conductivity reduction in PbTe_{1-x}Sex alloys. *Europhys. Lett.* **102**, 46002 (2013).
- Hellman, O., Steneteg, P., Abrikosov, I. A. & Simak, S. I. Temperature dependent effective potential method for accurate free energy calculations of solids. *Phys. Rev. B* **87**, 104111 (2013).
- Rohskopf, A., Seyf, H. R., Gordiz, K. & Henry, A. Phonon optimized potentials. Preprint at <https://arxiv.org/abs/1610.02353> (2016).
- Justo, J. F., Bazant, M. Z., Kaxiras, E., Bulatov, V. V. & Yip, S. Interatomic potential for silicon defects and disordered phases. *Phys. Rev. B* **58**, 2539–2550 (1998).
- Bazant, M. Z., Kaxiras, E. & Justo, J. F. Environment-dependent interatomic potential for bulk silicon. *Phys. Rev. B* **56**, 8542–8552 (1997).

26. Pettifor, D. G. & Oleinik, I. I. Analytic bond-order potentials beyond Tersoff-Brenner. I. Theory. *Phys. Rev. B* **59**, 8487–8499 (1999).
27. Dove, M. T. *Introduction to Lattice Dynamics*, Vol. 4 (Cambridge University Press, 1993).
28. Wei, L. & Asegun, H. Direct calculation of modal contributions to thermal conductivity via Green–Kubo modal analysis. *New J. Phys.* **18**, 013028 (2016).
29. Kubo, R. The fluctuation-dissipation theorem. *Rep. Prog. Phys.* **29**, 255–284 (1966).
30. Allen, P. B. & Feldman, J. L. Thermal conductivity of disordered harmonic solids. *Phys. Rev. B* **48**, 12581–12588 (1993).
31. Kiarash, G. & Asegun, H. A formalism for calculating the modal contributions to thermal interface conductance. *New J. Phys.* **17**, 103002 (2015).
32. Hardy, R. J. Energy-flux operator for a lattice. *Phys. Rev.* **132**, 168–177 (1963).
33. Broido, D. A., Malorny, M., Birner, G., Mingo, N. & Stewart, D. A. Intrinsic lattice thermal conductivity of semiconductors from first principles. *Appl. Phys. Lett.* **91**, 231922 (2007).
34. Feng, T. & Ruan, X. Quantum mechanical prediction of four-phonon scattering rates and reduced thermal conductivity of solids. *Phys. Rev. B* **93**, 045202 (2016).
35. Esfarjani, K. & Stokes, H. T. Method to extract anharmonic force constants from first principles calculations. *Phys. Rev. B* **77**, 144112 (2008).
36. Tersoff, J. Modeling solid-state chemistry: interatomic potentials for multi-component systems. *Phys. Rev. B* **39**, 5566–5568 (1989).
37. Stillinger, F. H. & Weber, T. A. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* **31**, 5262–5271 (1985).
38. Jain, A. & McGaughey, A. J. Effect of exchange–correlation on first-principles-driven lattice thermal conductivity predictions of crystalline silicon. *Comput. Mater. Sci.* **110**, 115–120 (2015).
39. Chernatynskiy, A., Flint, C., Sinnott, S. B. & Phillpot, S. R. Critical assessment of UO2 classical potentials for thermal conductivity calculations. *J. Mater. Sci.* **47**, 7693–7702 (2012).
40. Tadano, T., Gohda, Y. & Tsuneyuki, S. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *J. Phys. Condens. Matter* **26**, 225402 (2014).
41. Gordiz, K. & Henry, A. Phonon transport at crystalline Si/Ge Interfaces: the role of interfacial modes of vibration. *Sci. Rep.* **6**, 23139 (2016).
42. Weber, W. Adiabatic bond charge model for the phonons in diamond, Si, Ge, and alpha-Sn. *Phys. Rev. B* **15**, 4789–4803 (1977).
43. Steiger, S. et al. Enhanced valence force field model for the lattice properties of gallium arsenide. *Phys. Rev. B* **84**, 155204 (2011).
44. Musgrave, M. J. P. & Pople, J. A. A general valence force field for diamond. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **268**, 474–484 (1962).
45. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
46. Thompson, A. P., Swiler, L. P., Trott, C. R., Foiles, S. M. & Tucker, G. J. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330 (2015).
47. van Duin, A. C. T., Dasgupta, S., Lorant, F. & Goddard, W. A. ReaxFF: a reactive force field for hydrocarbons. *J. Phys. Chem. A* **105**, 9396–9409 (2001).
48. Kim, K. & Kaviani, M. Phonocatalysis. An ab initio simulation experiment. *AIP Adv.* **6**, 065124 (2016).
49. Rohskopf, A. *Phonon Optimized Potentials* (2017) www.pops.gatech.edu.
50. Turing, A. M. Computing machinery and intelligence. *Mind* **59**, 433–460 (1950).
51. Fogel, D. B. An introduction to simulated evolutionary optimization. *IEEE Trans. Neural Netw.* **5**, 3–14 (1994).
52. Coley, D. A. *An Introduction to Genetic Algorithms for Scientists and Engineers*. (World scientific, 1999).
53. Schaffer, J. D., Whitley, D. & Eshelman, L. J. *International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN-92)* 1–37 (IEEE, 1992).
54. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
55. Takuru, M., Takuma, S., Takuma, H., Keivan, E. & Junichiro, S. Importance of local force fields on lattice thermal conductivity reduction in PbTe 1–x Se x alloys. *Europhys. Lett.* **102**, 46002 (2013).
56. Zhou, F., Nielson, W., Xia, Y. & Ozoliņš, V. Lattice anharmonicity and thermal conductivity from compressive sensing of first-principles calculations. *Phys. Rev. Lett.* **113**, 185501 (2014).
57. Fennell, C. J. & Gezelter, J. D. Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. *J. Chem. Phys.* **124**, 234104 (2006).
58. Girifalco, L. A. & Weizer, V. G. Application of the Morse potential function to cubic metals. *Phys. Rev.* **114**, 687–690 (1959).
59. Glassbrenner, C. J. & Slack, G. A. Thermal conductivity of silicon and germanium from 3 K to the melting point. *Phys. Rev.* **134**, A1058–A1069 (1964).
60. Inyushkin, A. V., Taldenkov, A. N., Gibin, A. M., Gusev, A. V. & Pohl, H. J. On the isotope effect in thermal conductivity of silicon. *Phys. Status Solidi C* **1**, 2995–2998 (2004).
61. Ozhogin, V. I. et al. Isotope effect in the thermal conductivity of germanium single crystals. *J. Exp. Theoret. Phys. Lett.* **63**, 490–494 (1996).
62. Nilsson, G. & Nelin, G. Phonon dispersion relations in Ge at 80 K. *Phys. Rev. B* **3**, 364–369 (1971).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017