Computer Physics Communications 254 (2020) 107337

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Evolutionary multi-objective optimization and Pareto-frontal uncertainty quantification of interatomic forcefields for thermal conductivity simulations☆,☆☆

Aravind Krishnamoorthy*,[1], Ankit Mishra[1], Nicholas Grabar[1], Nitish Baradwaj, Rajiv K. Kalia, Aiichiro Nakano, Priya Vashishta

Collaboratory of Advanced Computing and Simulations, University of Southern California, Los Angeles, CA 90089-0242, United States of America

ARTICLE INFO

Article history:
Received 18 June 2019
Received in revised form 26 March 2020
Accepted 9 April 2020
Available online 28 April 2020

Keywords:
Thermal conductivity
Molecular dynamics
Forcefield parameterization
Genetic algorithm
Multi-objective optimization

ABSTRACT

Predictive Molecular Dynamics simulations of thermal transport require forcefields that can simultaneously reproduce several structural, thermodynamic and vibrational properties of materials like lattice constants, phonon density of states, and specific heat. This requires a multi-objective optimization approach for forcefield parameterization. Existing methodologies for forcefield parameterization use ad-hoc and empirical weighting schemes to convert this into a single-objective optimization problem. Here, we provide and describe software to perform multi-objective optimization of Stillinger–Weber forcefields (SWFF) for two-dimensional layered materials using the recently developed 3rd generation non-dominated sorting genetic algorithm (NSGA-III). NSGA-III converges to the set of optimal forcefields lying on the Pareto front in the multi-dimensional objective space. This set of forcefields is used for uncertainty quantification of computed thermal conductivity due to variability in the forcefield parameters. We demonstrate this new optimization scheme by constructing a SWFF for a representative two-dimensional material, 2H-$MoSe_2$ and quantifying the uncertainty in their computed thermal conductivity.

Program summary

Program Title: MOGA-NSGA3
Program Files doi: http://dx.doi.org/10.17632/pbc6nb7hp9.1
Licensing Provisions: GNU General Public License 3
Programming Language: C++
Nature of problem: Interatomic forcefields used for molecular dynamics simulations of thermal conductivity must be parameterized to accurately capture structural and vibrational properties of the material system being modeled. Therefore, these forcefields must be simultaneously optimized against several ($n \geq 5$) material properties. However, such parameterization is difficult using existing forcefield parameterization schemes, which are limited to optimization of a single or few ($n < 3$) objectives.
Solution method: We present software to perform evolutionary optimization of forcefields for thermal conductivity simulations using the recently developed 3rd generation non-dominated sorting genetic algorithm (NSGA-III). The algorithm's unique reference-point-based niching and non-dominated sorting schemes enable efficient exploration of higher-dimensional objective spaces while preserving diversity among forcefield populations. The best set of forcefields on the Pareto front are used for estimating uncertainty in computed thermal conductivity due to forcefield parameterization.

1. Introduction

Molecular Dynamics (MD) simulations can be used to model and quantify dynamic processes in nanomaterial systems such as suspended monolayer materials that are difficult to measure experimentally [1]. Interactions between atoms in these simulations are described by forcefields, which are empirical physically-motivated mathematical functions along with an associated set of parameters. Optimization of forcefields for thermal transport

☆ The review of this paper was arranged by Prof. D.P. Landau.

☆☆ This paper and its associated computer program are available via the Computer Physics Communication homepage on ScienceDirect (http://www.sciencedirect.com/science/journal/00104655).

* Corresponding author.
E-mail address: kris658@usc.edu (A. Krishnamoorthy).

[1] Authors contributed equally.

https://doi.org/10.1016/j.cpc.2020.107337
0010-4655/

involves identifying a set of forcefield parameters that will most accurately reproduce physical properties like lattice constants, elastic constants, and phonon frequencies that govern thermal transport. While this parameterization requires simultaneous optimization of multiple properties with different physical units, most existing parameterization schemes simplify this into a single objective optimization problem by using an ad-hoc empirical weighting schemes to construct a weighted sum of objectives. Forcefields constructed using such single objective schemes are highly dependent on the choice of objective weights, making it difficult to quantify the uncertainty associated with variability in forcefield parameters.

There have been limited attempts at multi-objective forcefield parameterization for complex forcefields like ReaxFF and COMB [2] using evolutionary approaches [3–5]. These methods have relied on algorithms such as NSGA-II, which are suitable only for optimization of 3 or fewer objectives, thus greatly limiting the applications for which forcefields can be optimized [3]. In this paper, we present a multi-objective optimization scheme based on the recent the 3rd generation non-dominated sorting genetic algorithm (NSGA-III) [6] for the parameterization of empirical forcefields of the Stillinger–Weber functional form [7] for modeling thermal transport in layered and two-dimensional materials. The NSGA-III algorithm represents a generational improvement over existing schemes [8–12] by enabling optimization of forcefields against significantly larger number of objectives (up to 11) and also includes a unique reference-point-based niching scheme for the preservation of diversity among forcefields and thus a more thorough sampling of the Pareto front necessary for accurate uncertainty quantification. NSGA-III also includes a normalization scheme to accommodate objectives which differ either in order of magnitude or physical units (e.g. lattice constants in nm vs. phonon frequencies in cm^{-1}).

Stillinger–Weber forcefields (SWFFs) are a family of short-ranged interatomic forcefields with simple functional forms suitable for simulating large systems ($>>10^6$ atoms) for long times ($>$ several *ns*). They also contain a moderately large set of empirical parameters that can be tuned to accurately capture second-order and higher-order interatomic force constants, which govern lattice thermal conductivity [13]. Due to this combination of properties, SWFFs have been widely used to compute the thermal conductivity of several two-dimensional and layered materials. While more complex forcefields with larger sets of tunable parameters may be used for thermal transport calculations, their flexibility in accurately describing a very diverse phase space of atomic configurations is not required for low-temperature non-equilibrium MD simulations, where atoms merely vibrate about their equilibrium positions [14]. Therefore, we choose to optimize Stillinger–Weber forcefields using our NSGA-III-based scheme. There is a large scatter in the computed values of thermal conductivity for the same two-dimensional material, MoS_2 [15]. While this scatter is partially attributable to variations in the MD simulation parameters (system size, choice of equilibrium or non-equilibrium simulations etc.), the majority of the variation arises from differences in the forcefields parameters, which differ based on choices made during the single-objective parameterization process.

2. Results

The SWFF for the 2H-$MoSe_2$ crystal models interactions between atoms using 2-body bonding terms and 3-body angle terms [7]. The total potential energy of the given system of N atoms located at [r_1, r_2, …, r_N] in the SWFF can be written as

$$E\left(r_1, r_2, \ldots r_N\right) = \sum_{i<j} V_2\left(r_{ij}\right) + \sum_{i<j<k} V_3\left(r_{ij}, r_{jk}, \theta_{ijk}\right)$$

where $r_{ij} = r_j - r_i$. The 2-body term, V_2, is defined as

$$V_2\left(r_{ij}\right) = A\left(\frac{B}{r_{ij}^4} - 1\right) exp\left(\frac{\gamma}{r_{ij} - r_{cut}}\right)$$

The two-body term is defined by 3 optimizable parameters, A, B and γ.

The 3-body term, V_3 around a central atom i is given by three optimizable parameters, λ, γ_1 and γ_2 and has the following functional form. Geometric parameters, including interaction cut-off distances, r_{cut}, r_{cut1} and r_{cut2} and equilibrium angles, θ_0 are held fixed during parameterization.

$$V_3\left(r_{ij}, r_{ik}, \theta_{ijk}\right) = \lambda exp\left(\frac{\gamma_1}{r_{ij} - r_{cut1}} - \frac{\gamma_2}{r_{ik} - r_{cut2}}\right)(cos\theta - cos\theta_0)^2$$

The forcefield is parameterized against the lattice-constant, elastic constants, and full phonon dispersion curves at three values of lattice strain. Ground-truth values of the lattice and elastic constants are taken from quantum-mechanical density functional theory (DFT) calculations of the 2H and 1T' monolayer crystal. Mode Grüneisen parameters are calculated by comparing the phonon dispersion curve of the relaxed (i.e. strain-free) monolayer with monolayers under 5% compressive and tensile biaxial strains. The SWFF is parameterized to minimize 5 objectives — the root-mean-square (RMS) distance between the DFT value and forcefield-computed values of lattice constant, elastic constant and three phonon dispersion curves (one for each strain state). The error associated with phonon dispersions is calculated as the RMS difference in the predicted and ground-truth phonon frequencies for each band at each q-point along selected high-symmetry directions (Fig. 1d). The error is equal to a weighted sum of RMS frequency difference values for each band. The weights are chosen to give a greater priority to acoustic vibrational modes. Specifically, in our optimization, the relative weights assigned to the three lowest-energy bands are 30 times the values assigned to the remaining (higher-frequency optical) bands. This choice ensures a more accurate reproduction of acoustic bands by the optimized forcefields and is motivated by previous observations that acoustic phonons are responsible for the majority of lattice thermal transport in semiconducting and insulating systems [16]. Details about the DFT calculations for computing ground-truth values are given in the Methods section.

2.1. Parameterization of 2H $MoSe_2$

The monolayer 2H phase of $MoSe_2$ has a honeycomb crystal structure with hexagonal unit cell composed of a layer of 6-fold coordinated Mo atoms sandwiched between two planes of 3-fold coordinated Se atoms in an ABA stacking sequence (Fig. 1). The crystal structure consists of two atom types — Mo and Se, and can be effectively described by three two-body interactions — Mo-Mo, Se-Se and Mo-Se along with two three-body interaction — Mo-Se-Mo and Se-Mo-Se. Excluding the interaction cutoff distances, r_{cut}, r_{cut1} and r_{cut2}, the SWFF for the $MoSe_2$ system can be characterized by 13 design variables, which must be parameterized to minimize 5 objectives (lattice constant, elastic constant, phonon dispersion curves at three values of lattice strain).

Fig. 2 shows the reduction in the computed errors of all five objectives during forcefield optimization. The dark band indicates the variance in the measured objective between the different forcefields on the Pareto front at each generation. It is noticeable that the lattice constant and elastic moduli are sensitive to the variation of forcefield parameters, as indicated by the relatively large variance in computed errors. In contrast, the computed errors in phonon frequencies at all strain levels falls

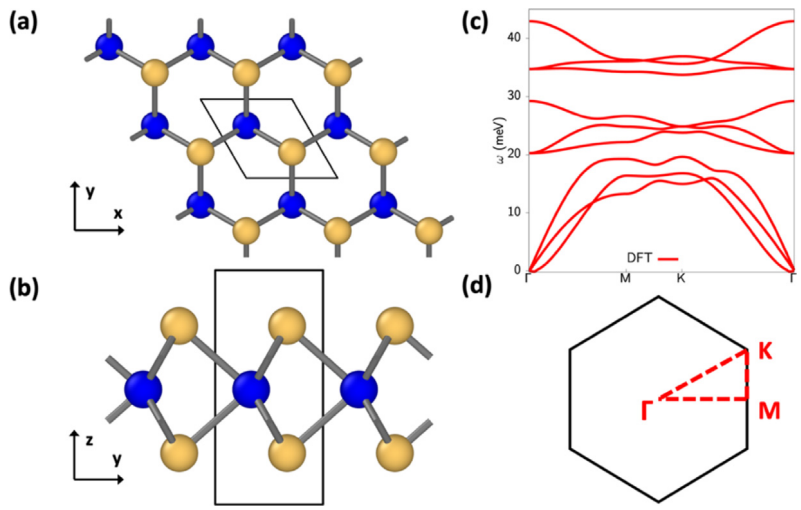

Fig. 1. (**a, b**) Crystal structure of $MoSe_2$ showing two distinct atom types and (**c**) Phonon dispersion of the unstrained monolayer along high-symmetry directions (**d**) in the Brillouin zone.

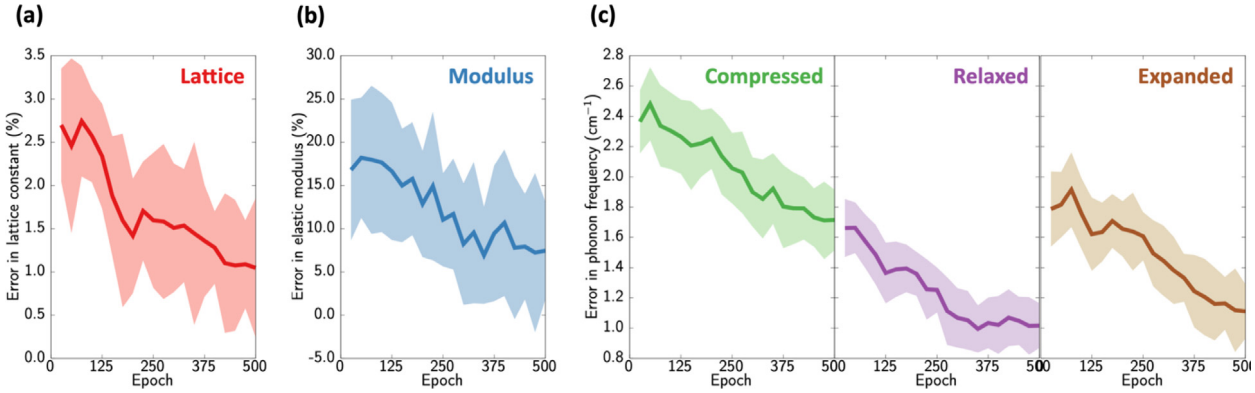

Fig. 2. Computed error in lattice constant (a), elastic modulus (b) and three phonon dispersion curves (c). The dark line represents median of the Pareto front for each epoch, while the band quantifies the variance among forcefields on the front.

Table 1
Two-body terms of the Stillinger–Weber forcefield.

Interacting atoms	A	γ	B	r_{min}	r_{max}
Se-Se	2.738	1.564	40.208	0.00	4.493
Mo-Se	5.459	0.497	13.307	0.00	3.345
Mo-Mo	2.899	0.406	38.732	0.00	4.493

The two-body interaction term has the form $V_2 = A\left(\frac{B}{r_{ij}^p} - 1\right) exp\left(\frac{\gamma}{r_{ij}-r_{cut}}\right)$.

Table 2
Three-body angle bending terms in the Stillinger–Weber forcefield.

Triplet	λ	θ_0	$\gamma_1 = \gamma_2$	r_{12}^{max}	r_{13}^{max}	r_{23}^{max}
Me-Mo-Se	53.107	80.154	2.036	3.345	3.345	4.493
Mo-Se-Mo	19.010	80.154	5.921	3.345	3.345	4.493

The three-body interaction term has the form $V_3 = \lambda exp\left(\frac{\gamma_1}{r_{ij}-r_1} - \frac{\gamma_2}{r_{ik}-r_2}\right)(cos\theta - cos\theta_0)^2$.

within a narrow band, indicating that all the forcefields on the front accurately reproduce the phonon dispersion curves.

The outcome of the genetic algorithm optimization is highly dependent on the choice of optimization metaparameters: population size, crossover and mutation ratios, etc. The forcefields in Fig. 2 were generated by the NSGA-III with a population size of 300 examples. At every epoch, crossover was performed on 80% of the population using a simulated binary crossover operator. Mutation was simulated by choosing forcefield parameters from a uniform distribution between pre-specified lower- and upper-bounds for each decision variable. Tables 1 and 2 show a representative forcefield for 2H $MoSe_2$ on the Pareto front of the final epoch. Nine other representative forcefields from the same front are provided in the SI.

3. Computational implementation

In addition to lower and upper bounds on each decision variable, forcefield optimization using NSGA-III algorithm also requires specification of a number of meta-parameters. Some of these hyper-parameters such as population sizes, crossover and mutation rates and magnitude of mutation are generic to evolutionary algorithms and other parameters like number of hyperplane reference points are specific to the NSGA-III method. The NSGA-III module (nsga3.c) completes a single iteration of Pareto optimization based on the input population(s). This single iteration approach allows for generality and optimization of objective functions without explicit functional forms. The output is two files: a new child population of variables and a copy of the previous population. This second population will be used in the

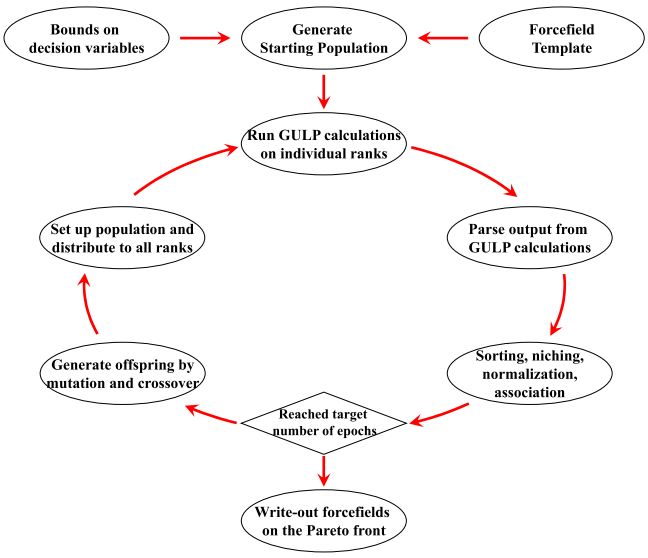

Fig. 3. Flow chart showing the various components of the workflow used for implementation of parallel genetic algorithms.

next iteration as well as the new population to improve retention of successful forcefields between generations.

Implementation of successive iterations for forcefield optimization is accomplished via a wrapper workflow script (moga.c). Fig. 3 shows the flow of control in the NSGA-III-MOGA workflow. The implementation of the workflow is parallel, employing the Message Passing Interface (MPI) library [17,18]. Each process evaluates the four objectives functions for a single forcefield in the current population. This is accomplished via the task runner software GULP [19]. Processes are assigned a line in the population file. The data is then parsed into the input format required for GULP. The resulting evaluations are collected serially into a single file and an iteration of Pareto optimization is performed.

NSGA-III implements a unique reference-point based framework that emphasizes population members that are non-dominated, yet close to a set of supplied reference points. The essential components of the previous iterations of the NSGA algorithm [20] were classification of the population into non-dominated levels (non-dominated sorting), selection of the 'mating' population according to rank and identification of the Pareto front (selection), and generation of the next population via genetic operations (crossover and mutation). NSGA-III includes these steps as well as the following: normalization of the values of the objective functions (normalization), association of members of the normalized population to the nearest reference point on the hyperplane (association), and selection of members among the Pareto front to maximize the diversity of the population relative to the reference points (niching).

The numerical complexity of the NSGA-III algorithm is $O(MN^2)$ where M is the number of objective and N is the number of populations, hence the algorithm scales as N^2 based on the population size. The implementation of the workflow as well as the genetic algorithm code is done in C++ to ensure the modularity and extending its plug-in capability of variety of MD simulation engines such as LAMMPS, GULP etc.

3.1. Implementation of NSGA-III

The current implementation of NSGA-III is an adaptation of the MATLAB code from the reference paper of Deb et al. [6]. Briefly, the unique aspects of the algorithm consist of three operations – normalization, association, and niching, which are described below, along with the algorithm in its entirety, in pseudocode.

Algorithm $Association(S_t, Z^r)$

Input: S_t: $populationsetsuchthatits \leq N$
Z^r: $referencesetchoosenforsamplingofparetofront$

Output:
π_s: $nearestreferencepoints$
d_s: $distancebetweens \wedge referencepoint\pi_s$

Steps:
1. $Computationofreferenceline$
   $forz\epsilon Z_r$
      $Computew = z$
   $endfor$
2. $Computationof\pi_s \wedge d_s$
   $fors\epsilon S_t$
      $forw\epsilon Z_r$
         $d^{\perp}(s,w) = s - w^T s/ \quad |w| \vee$
      $endfor$
   $endfor$
   $\pi(s) = w: argmin_{w\epsilon Z^r} d^{\perp}(s,w)$
   $d(s) = d^{\perp}(s, \pi(s))$

Algorithm $Niching(K, \rho_j, \pi, d, Z^r, F_l)$

Input: K: $Numberofmembersstillneedbeincluded F_l$

ρ_j: $Nichecount$
π_s: $Nearestreferencepoint \forall s\epsilon S_t$
d_s: $Distanceofs\epsilon S_t respective \pi_s$

Output:
P_{t+1}: $Finalpopulationofforn extgeneration$

Steps:
1. $Computationof P_{t+1}$
   $whilek \leq K$
      $J_{min} = \{j: argmin_{j\epsilon Z^r} \rho_j\}$
      $j = random(J_{min})$
      $I_j = \{s: \pi(s) = j, s\epsilon F_l\}$
      $if I_j \neq \phi$
         $if \rho_j = 0$
            $P_{t+1} = P_{t+1} U \left(s: argmin_{S\epsilon I_j} d(s)\right)$
         $else$
            $P_{t+1} = P_{t+1} U random(I_j)$
         $endif$
         $\rho_j = \rho_j + 1, F_l = F_l$
         $k = k + 1$
      $else$
         $Z^r = Z^r / \{j\}$
      $endif$
   $endwhile$

4. Discussion

In contrast to single-objective schemes, which produce a single optimized forcefield that minimizes the given objective, multi-objective schemes converge to a set of optimal forcefields that simultaneously minimize errors associated with all objectives. In the presented NSGA-III-based scheme, this is represented by the Pareto front of mutually non-dominated, reference-point based

Algorithm Non-dominated-Sort-Genetic Algorithm-III ($\mathbf{P}_t$, $\mathbf{P}_{t-1}$)

Input:

$\mathbf{P}_t$: population of the t^{th} generation of size N, containing n tuple element each
$\mathbf{P}_{t-1}$: population of the previous generation

Output:

$\mathbf{P}_{t+1}$: Population for next generation P_{t+1}

Steps:

1. Initialization and non-dominated sorting

$S_t = \emptyset, i = 1$
$R_t = P_t U P_{t-1}$
$\{F_i\} = Non - dominatedSort(R_t)$

2. Loop until size of current population exceeds N

while ($|S_t| < N$)
$S_t = S_t U F_i$
$i = i + 1$
end while

3. Include more genes until $|S_t| = N$; last front to be included is $F_l = F_i$,

if)
$P_{t+1} = S_t$
$else$
$1. P_{t+1} = j = 1 l = 1 F_j$
K more points to be chosen from F_l such that $K = N - P_{t+1} \vee$
2.Normalization of objectives and creation of reference set Z_r
$Normalize$)
3. Association of set S with reference points
$\pi_s, d_s = Associate(S_t, Z_r)$
$\pi_s: closestreferencepoint$
$d_s: distancebetweenreferencepoint \wedge s$
4. Compute niche count of reference point $j \epsilon Z^r: \rho_j = \sum_{s \epsilon S/F_l} (\pi_s | \quad | j)$
5. Choose K members one at a time from F_l to construct P_{t+1}
$P_{t+1} = Niching\left(K, \rho_j, \pi, d, Z^r, F_l, P_{t+1}\right)$
end if

forcefields. This Pareto-frontal breakdown of different forcefields for each generation provides a natural way to establish one of the primary sources of uncertainty in molecular dynamics simulations — namely the uncertainty in forcefield parameters. Variability in the computed thermal conductivity between these forcefields gives us the effective epistemic error due to forcefields parameters. The diversity preservation properties of NSGA-III, including niching and association to reference points on the hyperplane in the objective space, provide a more accurate estimate of the uncertainty in thermal conductivity due to errors in forcefield parameters. This Pareto-frontal uncertainty quantification approach offers an alternative method to estimate the errors in forcefield parameters [3,21–24], to complement the predominantly Bayesian approaches employed in prior studies [25].

Fig. 4 shows the computed thermal conductivity of 10 representative forcefields on the Pareto front of the final generation (Table 1 and SI). The computed κ shows a relatively tight distribution around 31.2 W/mK for a system of size 3200 × 100 Å in spite of a relatively large variation in the forcefield parameters. This error (σ = 3.07 W/mK) is indicative of the epistemic error in the forcefield parameters among a set of forcefields that are optimized to reproduce the physical properties relevant for thermal transport. System level quantum corrections and isotopic scattering corrections will affect the computed thermal conductivity values from each forcefield identically and would thus have no significant effect on the calculated variance of thermal conductivity.

5. Conclusion

In summary, we describe software to perform multi-objective optimization of atomistic forcefields for molecular dynamics simulations using the 3rd generation non-dominated-searching-based genetic algorithm (NSGA-III), which allows for the optimization of forcefields to reproduce up to 11 physical properties. We demonstrate this multi-objective evolutionary optimization workflow to construct empirical Stillinger–Weber forcefields to reproduce the lattice constant, elastic moduli and phonon dispersion curves at different values of lattice strain for a representative two-dimensional material — 2H-$MoSe_2$. The epistemic uncertainty in computed thermal conductivity is computed as the standard deviation of thermal conductivity computed from all forcefields on the Pareto front. Our NSGA-III-based workflow provides a better estimate of this uncertainty than previous work due to a better sampling of the Pareto front from the reference-point based preservation of forcefield diversity. This relatively large parameter and objective space accessible by the NSGA-III algorithm suggests that this method can be extended to more complex forcefields such as ReaxFF and COMB with significantly larger parameter sets.

Algorithm $Normalize$)

Input: $S_t: populationsetsuchthatits \leq N$
$Z^r: referencesetchoosenforsamplingofparetofront$
$Z^s: structuredpoints$
$Z^a: suppliedpoints$

Output:
$f^n, Z_r pointsonreferencenormalizedhyper-plane$

Steps:

1. $Computationofidealtranslatetheobjectivesfollowedbycomputationofextremepoints$
$forj = 1M$
$IdealPointComputation: z_j^{min} = min_{s\epsilon S_t} f_j(s)$
$ObjectiveTranslation: f_j'(s) = f_j(s) - z_j^{min} \forall s\epsilon S_t$
$ExtremePointComputation: z_j^{max} = s: argmin_{s\epsilon S_t} ASF(s, w_j) wherew_j = (\epsilon, \epsilon, .., \epsilon)^T$
$\epsilon = 10^{-6} \wedge w_j = 1$
$endfor$
2. $Normalizationofobjectivesf^n basedoncomputedinterceptsa_j forj = 1,2,3, \ldots.., M$
$fori = 1M$
$f_i^n = \frac{f_i'(x)}{a_i - z_i^{min}} = \frac{f_i(x) - z_i^{min}}{a_i - z_i^{min}}$
$endfor$

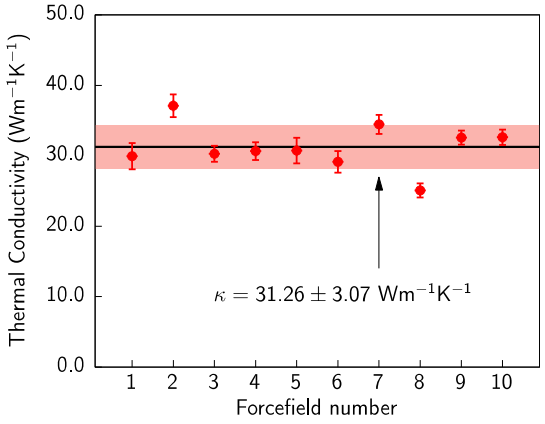

Fig. 4. Computed thermal conductivity values using 10 representative forcefields from the final Pareto front after 500 epochs for 2H $MoSe_2$ crystals. Error-bars on individual points represent uncertainty in computed thermal conductivity for each forcefield due to variations in the slope of the temperature profile.

6. Methods

6.1. NSGA-III forcefield parameterization

Parameterization was performed with population size of 300–500 for 500 generations to achieve convergence. Parameterization is performed in two stages. We initiate the first stage with randomly-generated values for each of the 13 decision variables with magnitudes similar to previously published values for the isomorphous MoS_2 crystal [15]. In this stage, we allow a perturbation of 50% for each parameter around the initial value. The mutation and crossover rate are also kept high initially at 60% and 40% respectively to ensure that the parameter space is well-sampled resulting in high diversity in the forcefield population. The forcefields on the Pareto front in this stage are chosen as the initial population for the second stage of the parameterization, where we impose a smaller perturbation of 20% with mutation and crossover rate of 80% and 20% respectively. This two-stage strategy ensures both a broader search of the parameter space as well as convergence to the global optimum set of forcefields. The MOGA scripts, *moga.c* and *ga.c*, can be compiled with GNU or Intel C++ compilers. The workflow and NSGA-III code are completely standalone in nature and do not require any external libraries or module to be downloaded for execution of force field optimization. The method uses GULP [19] for computing the 5 objectives from the forcefield.

6.2. Density functional theory simulations

Optimization of 2H $MoSe_2$ crystal structures were done quantum mechanically using density functional theory (DFT) with the projector augmented wave (PAW) [26] method implemented in the Vienna Ab initio Simulation Package (VASP) [27,28]. Exchange and correlation effects are calculated using the Perdew–Burke–Ernzerhof form of the generalized gradient approximation to the exchange–correlation functional [29]. Valence electron wave functions are constructed using a plane wave basis set containing components up to a kinetic energy of 450 eV and the reciprocal space is sampled at the Γ point with a 0.1 eV Gaussian smearing of orbital occupancies. DFT simulations were performed on systems containing 108 atoms, corresponding to 36 formula units of $MoSe_2$, in a simulation cell measuring 19.73 Å × 19.73 Å along the crystalline *a*- and *b*-directions. A vacuum of 15 Å is added along the *c*-axis of all simulation cells to remove spurious image interactions. Calculations were performed till each self-consistency cycle is converged in energy to within 1×10^{-5} eV/atom and forces on ions are under 1×10^{-2} eV/Å.

Dynamical stability was inferred from the frequency of normal vibration modes of the 2H crystal structure. The Hessian matrix was generated within the formalism of density functional perturbation theory and dispersion relations for normal

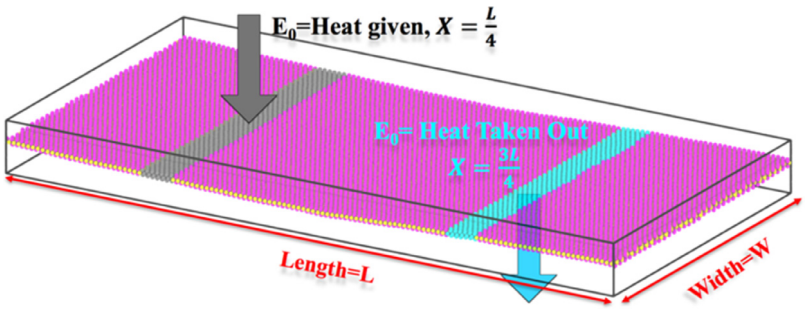

Fig. 5. Schematic of the NEMD simulations for measuring Thermal Conductivity of 2D Materials.

vibration modes were calculated using the open-source phonopy package [30].

6.3. Non-equilibrium molecular dynamics (NEMD) simulations for thermal conductivity calculations

Thermal conductivity is calculated by the so-called 'direct' method using non-equilibrium molecular dynamics simulations. In this method, a fixed and predefined heat flux, $\widehat{E_0}$, is added to the kinetic energy of atoms at $x = L/4$, where L is the dimension of the system along the heat-transfer direction, and an identical heat flux is removed from the atoms at $x = 3L/4$. At steady state, these thermal fluxes will establish a temperature gradient between $L/4$ and $3L/4$, which is related to the thermal conductivity of the materials (as determined by the forcefield) through the Fourier law of heat conduction. This method [31,32] provides us with linear temperature profiles as opposed to the Muller Plathe method [33]. See reference [15] for more details.

Specifically, $\kappa = \frac{1}{2A} \cdot \widehat{E_0} \cdot \left(\frac{dT}{dx}\right)^{-1}$, where A is the cross-sectional area of heat transfer, taken to be equal to 6.5 Å, one-half of the out-of-plane lattice constant in bulk 2H-$MoSe_2$ crystals (see Fig. 5).

In the systems that we used for testing the forcefields, we use a system of size 3200 × 100 Å on the x and y directions. All the simulations are carried out using LAMMPS (Large Scale Atomic/Molecular Massively Parallel Simulator) [34]. In these simulations periodic boundary conditions are applied in all three directions with a vacuum of about 100 Å in the z direction to simulate a thin film. The MD simulations use a timestep of 2 femtoseconds. Conjugate Gradient relaxation is done initially to obtain the correct box size. Following which the atoms are given a gaussian distribution of velocities equivalent to 300 K and the system is thermalized as a constant energy ensemble for 20 picoseconds (ps). This process is repeated 5 times. The system is then thermalized at 300 K as a constant temperature ensemble for 200 ps after which the system is ready for carrying out thermal conductivity calculations. NEMD simulations are performed for 12 nanoseconds (ns) with a timestep of 2 femtoseconds (fs) and an imposed heat flux of 0.2 eV/ps. The temperature gradient is computed over the last 6 ns of the NEMD simulation, to ensure that a steady state is established.

Software distribution

The source code and documentation for the NSGA-III-MOGA code is available on GitHub at https://github.com/USCCACS/MOGA-NSGA3.The software is free for use and distributed under the GNU General Public License v3.0.

Acknowledgments

This work was supported as part of the Computational Materials Sciences Program funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award Number DE-SC0014607.

Declaration of competing interest

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

References

[1] S.C. Lin, M.J. Buehler, Carbon 77 (2014) 351–359.
[2] T. Liang, et al., Annu. Rev. Mater. Res. 43 (1) (2013) 109–129.
[3] A. Mishra, et al., NPJ Comput. Mater. 4 (1) (2018) 42.
[4] A. Jaramillo-Botero, S. Naserifar, W.A. Goddard, J. Chem. Theory Comput. 10 (4) (2014) 1426–1439.
[5] J.P. Larentzos, et al., J. Chem. Theory Comput. 11 (2) (2015) 381–391.
[6] K. Deb, H. Jain, IEEE Trans. Evol. Comput. 18 (4) (2014) 577–601.
[7] F.H. Stillinger, T.A. Weber, Phys. Rev. B 31 (8) (1985) 5262–5271.
[8] J. Hunger, G. Huttner, J. Comput. Chem. 20 (4) (1999) 455–471.
[9] J. Wang, P.A. Kollman, J. Comput. Chem. 22 (12) (2001) 1219–1228.
[10] M.V. Ivanov, M.R. Talipov, Q.K. Timerghazin, J. Phys. Chem. A 119 (8) (2015) 1422–1434.
[11] F. Leonarski, et al., Genetic Algorithm Optimization of Force Field Parameters: Application To a Coarse-Grained Model of RNA, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
[12] F. Leonarski, et al., J. Chem. Theory Comput. 9 (11) (2013) 4874–4889.
[13] Z. Liang, et al., J. Appl. Phys. 118 (12) (2015).
[14] A. Rohskopf, et al., NPJ Comput. Mater. 3 (27) (2017).
[15] A. Krishnamoorthy, et al., AIP Adv. 9 (3) (2019).
[16] A.N. Gandi, U. Schwingenschlogl, Europhys. Lett. 113 (3) (2016) 36002.
[17] F. MPI, MPI: A Message-Passing Interface, Oregon Graduate Institute School of Science \ & Engineering, 1994.
[18] M.P. Forum, MPI: A Message-Passing Interface Standard, University of Tennessee, 1994.
[19] J.D. Gale, A.L. Rohl, Mol. Simul. 29 (5) (2003) 291–341.
[20] K. Deb, et al., IEEE Trans. Evol. Comput. 6 (2) (2002) 182–197.
[21] Z. Xiaowang, in: M.F. Stephen, H. Jan Peter (Eds.), Uncertainty Quantification and Model Calibration, IntechOpen, Rijeka, 2017, pp. 89–111.
[22] F. Cailliez, P. Pernot, J. Chem. Phys. 134 (5) (2011).
[23] A.P. Moore, et al., Comput. Mater. Sci. 126 (2017) 308–320.
[24] P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, J. Chem. Phys. 137 (14) (2012) 144103.
[25] S.L. Frederiksen, et al., Phys. Rev. Lett. 93 (16) (2004) 165501.
[26] P.E. Blöchl, Phys. Rev. B 50 (24) (1994) 17953–17979.
[27] G. Kresse, J. Furthmuller, Phys. Rev. B 54 (16) (1996) 11169–11186.
[28] G. Kresse, J. Furthmuller, Comput. Mater. Sci. 6 (1) (1996) 15–50.
[29] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (18) (1996) 3865–3868.
[30] A. Togo, I. Tanaka, Scr. Mater. 108 (2015) 1–5.
[31] T. Ikeshoji, B. Hafskjold, Mol. Phys. 81 (2) (1994) 251–261.
[32] P. Wirnsberger, D. Frenkel, C. Dellago, J. Chem. Phys. 143 (12) (2015) 124104.
[33] F. Muller-Plathe, Phys. Rev. E 59 (5) (1999) 4894–4898.
[34] S. Plimpton, J. Comput. Phys. 117 (1) (1995) 1–19.