ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

Optimization of Thermal Conductance at Interfaces Using Machine Learning Algorithms

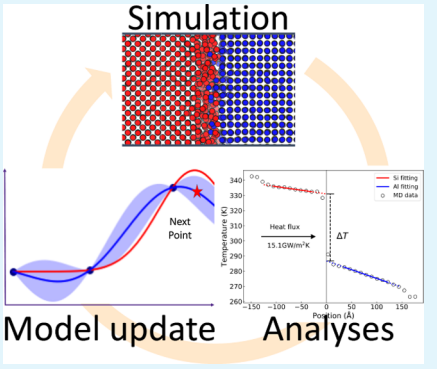
Sabiha Rustam,* Malachi Schram,* Zexi Lu, Anne M. Chaka, W. Steven Rosenthal, and Jim Pfaendtner

Cite This: *ACS Appl. Mater. Interfaces* 2022, 14, 32590−32597

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Optimization of thermal transport across the interface of two different materials is critical to micro-/nanoscale electronic, photonic, and phononic devices. Although several examples of compositional intermixing at the interfaces having a positive effect on interfacial thermal conductance (ITC) have been reported, an optimum arrangement has not yet been determined because of the large number of potential atomic configurations and the significant computational cost of evaluation. On the other hand, computation-driven materials design efforts are rising in popularity and importance. Yet, the scalability and transferability of machine learning models remain as challenges in creating a complete pipeline for the simulation and analysis of large molecular systems. In this work we present a scalable Bayesian optimization framework, which leverages dynamic spawning of jobs through the Message Passing Interface (MPI) to run multiple parallel molecular dynamics simulations within a parent MPI job to optimize heat transfer at the silicon and aluminum (Si/Al) interface. We found a maximum of 50% increase in the ITC when introducing a two-layer intermixed region that consists of a higher percentage of Si. Because of the random nature of the intermixing, the magnitude of increase in the ITC varies. We observed that both homogeneity/heterogeneity of the intermixing and the intrinsic stochastic nature of molecular dynamics simulations account for the variance in ITC.



KEYWORDS: *Bayesian optimization, molecular dynamics simulation, Si/Al interface, interatomic mixing, interfacial thermal conductance*

1. INTRODUCTION

As the size of electronic devices decreases, controlling heat transfer becomes a critical issue. For some devices such as thermoelectrics, retention of heat is a desirable characteristic to improve the figure-of-merit.[1,2] In contrast, for the majority of micro- and nanoscale devices as well as cryogenic detectors and quantum information systems, heat dissipation is important to maintain function.[3,4] For nanoscale devices in particular, when the surface-to-volume ratio is high, maximizing heat transfer across solid−solid interfaces is essential. This corresponds to maximizing phonon transport across interfaces. Hence, identifying ways to vary interfacial phonon scattering that can lead to controlling the thermal conductivity (Kapitza conductance) in nanosystems can provide necessary insights for tailoring thermal performance in devices and guidance for materials synthesis.[5]

In macroscopic devices, interfacial thermal transport depends primarily on the interfacial structure and contact areas (vibrational and lattice mismatch of the two materials are nanoscale factors). At the atomic level, however, roughness at the interface has been demonstrated to influence thermal transport.[5−11] Roughness can entail a wide range of phenomena, including compositional intermixing, nanopatterning, amorphous structures, and defects such as impurities, vacancies, and dislocations. Typically, roughness at an interface has been demonstrated to decrease thermal transport, both experimentally[5−7] and computationally.[8−11] Hopkins and colleagues used high-vacuum thin-sputter deposition to demonstrate that at a Cr/Si boundary thermal conductivity varies as a function of the thickness of the intermixing layer as well as the rate of compositional change.[6] Duda and Hopkins found that chemical etching can be used to control roughness between thin aluminum films and silicon substrates, and the impact on thermal conductivity measured by time-domain thermoreflectance was exponential with respect to the root-mean-square roughness of the interface.[7] In some computational models, Kapitza conductance was found to be reduced by amorphous or disordered regions[9] or by large mass impurities.[11] Other computational models, however, have determined that the thermal conductance can be increased by bridging the phonon mismatch between two materials. This bridging can occur by several methods such as intermixing,[12,13] introduction of impurity atoms of an intermediate mass,[11,14] or

Received: November 30, 2021
Accepted: June 27, 2022
Published: July 8, 2022

deposition of an interfacial film with median properties.[15] Increasing thermal conductance across an interface requires precise control of the interfacial structure, which is becoming increasingly attainable due to advanced synthetic methods such as molecular beam epitaxy. Despite these examples of atomic arrangements increasing thermal conductance, an optimum arrangement has not yet been determined because of the large number of potential atomic configurations and the significant computational cost of evaluation.

With current technological advances in the field of machine learning (ML) and computing science, the accessibility of ML methods for molecular simulations has increased significantly in the past decade.[16−18] Despite the variety of ML tools that have been applied to molecular simulation, there are still significant challenges in creating a complete pipeline for the simulation and analysis of atomistic systems. In particular, improvements in scalability and transferability are necessary for closing this gap, especially for large computational systems.

We developed an efficient and scalable framework to explore a large number of atomistic configurations to optimize thermal conductivity at the silicon and aluminum (Si/Al) interface. Silicon and aluminum were chosen as the materials because of their widespread use in micro- and nanoelectronics, detectors for dark matter and high-energy physics,[19] and resonator applications in quantum computing devices.[20] In addition, high-quality Si/Al interfaces can be grown by using techniques such as molecular beam epitaxy and ultrahigh-vacuum evaporation, where the abruptness of the interface and degree of intermixing can be controlled.[21] Only phonon transport was considered in this study as the interfacial electronic effects either can be separated from phonons[22,23] or can be neglected in nanosized thin films.[24] We used a single objective Bayesian optimization technique to optimize heat transfer at the semiconductor:metal interface by tuning the depth and fraction of compositional intermixing. A Gaussian process-based ML algorithm was used to search the configuration space, and nonequilibrium molecular dynamics (NEMD)[25] was used to evaluate the thermal conductance.[26]

Bayesian optimization has been used extensively to optimize a scalar figure-of-merit or property of interest—such as isotropic thermal conductance—when at any set of input parameters the property remains uncertain and therefore can only be known up to some distribution.[27−29] The distribution of the property as a function of the input parameter is termed the response function. An upper confidence bound (UCB)[30] approach is called for when the purpose is to identify a subset of possible parameters that maximizes the response function. However, it is computationally expensive to use molecular dynamics (MD) simulations to estimate thermal conductance, so a parallel batch sampling variation is developed for this study in which several choices for the input parameters are evaluated simultaneously in a high-performance computing (HPC) environment. Typically, in statistical sampling, as the number of parameters tested increases, the knowledge of the thermal conductance distribution improves. A standard learning rate is applied to the UCB cutoff so that as the number of parameters tested increases, the level of confidence in the UCB likewise increases, and the uncertainty in the optimal material parameters decreases.

2. METHODOLOGY

2.1. Simulation Domain. MD is a powerful technique to simulate heat transfer processes, as it intrinsically includes all lattice vibrations with well-developed interatomic potentials. In its classical limits, MD is an ideal method for predicting thermal boundary resistance.[31,32] We chose the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package[33] for integration with ML because of its high-efficiency computation parallelization and high customizability. We used the modified embedded atomic method (MEAM) potential of Jelinek et al., which has been reported to reproduce Al and Si properties with high accuracy for all pairwise atomic interactions.[34]

2.1.1. System Setup. The schematic of the periodic simulation cell is shown in Figure 1. The system consists of

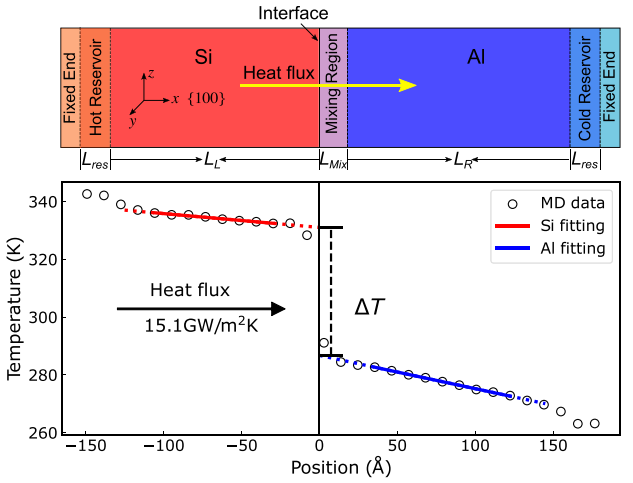


Figure 1. Schematic of the simulation cell (top) and example of fitted temperature profile (bottom). Solid lines are the fitting to actual data points, and dashed lines are the extrapolation of the fitted lines. Each point on the temperature profile is obtained through kinetic energy of the group of atoms in a 1.1 nm cross-sectional bin in the heat transfer direction averaged over the number of atoms in the bin over the data collection period.

hot/cold reservoirs and fixed boundaries at both ends to break periodicity in the heat transfer direction, while maintaining periodic boundary conditions in the y and z directions. The reservoirs have the same composition and periodicity as the sample. The sample region is defined as everything except for the reservoirs and fixed boundary regions. The Si region is 30 unit cells in length (lattice constant a_{Si} = 5.43 Å), and the Al region is 45 unit cells in length (a_{Al} = 4.05 Å). The two regions are in contact with the mixing region that consists of 3 unit cells of Al, forming a single interface parallel to the (100) crystallographic plane. The cross-sectional area of the simulation box, which is parallel to the interface, consists of 6 × 6 Si unit cells and 8 × 8 Al unit cells, such that the lattice mismatch is small (0.56%) to minimize interfacial strain. As the bulk of the materials were fully relaxed prior to NEMD simulations, the straining effects are assumed to be negligible. Therefore, we omit the investigation of straining effect on interfacial thermal transport properties. The mixing region bounds the optimization space to 6 atomic layers of Al, with mixing ratio defined as the ratio of Si to Al. For an intermixing arrangement of [l, r], where l is the mixing depth (atomic layers) and r is the mixing ratio, the initial configuration is produced by obtaining the indices of all Al atoms in the mixing region l and randomly selecting r% of indices to replace the atom type with Si atom. A unity mixing fraction means that all

https://doi.org/10.1021/acsami.1c23222
ACS Appl. Mater. Interfaces 2022, 14, 32590−32597

atoms in the mixing region have been swapped from Al to Si while still initially in an FCC configuration. The minimum mixing fraction is set to 0.01 to ensure there are at least 1−2 Si atoms intermixed into the Al structure, while the maximum mixing fraction is set to 0.99 to ensure there are at least 1−2 Al atoms present in the mixing region.

2.1.2. Direct Method. The NEMD heat transfer simulation is set up by using the direct method,[26] where a fixed amount of energy $\Delta\epsilon$ is added to atoms in the hot reservoir in the Si substrate and the same energy is subtracted from atoms in the cold reservoir in the Al substrate, creating a heat flux $J = \Delta\epsilon/A_c\Delta t$ (A_c is the cross-sectional area; Δt is the time step of MD). This establishes a temperature gradient in the system, which when combined with knowledge of the energy transfer rate enables us to compute the interfacial thermal conductance (ITC) as the inverse of the Kapitza resistance:[35]

$$G = \frac{1}{R_k} = \frac{J}{\Delta T} \tag{1}$$

where J is the heat flux (in W/m^2) and ΔT is the interfacial temperature gradient. The interface has a variable thickness due to atomic intermixing; thus, we treat the maximum Si/Al mixing region as the interfacial region and evaluate ΔT at its boundaries, so that G can manifest all the influence from intermixing and represent the overall thermal performance of the system when the length is fixed. The method of applying heat flux involves scaling the velocities of the atoms in each of the reservoirs to add or remove a desired amount of kinetic energy. We defined thickness of the reservoirs to consist of four unit cells in each material such that a sufficient number of atoms are present in the reservoirs, reducing the energy scaling per atom.

2.1.3. Data Collection and Analysis. During MD simulations, the sample and reservoirs are initially equilibrated at a uniform temperature of 300 K by randomly generating the atomic velocities according to a Gaussian distribution. This is followed by a constant number of atoms, volume, and energy (NVE, microcanonical) simulation, where a heat flux of 15.1 GW/m^2 is applied. The magnitude of the heat flux is large enough to create a distinct temperature profile along the substrate without developing nonlinearities throughout the sample due to the temperature dependence of the thermal conductivity.[13] Immediately after the heat flux is applied, the simulation runs for a period of 4 ns with 1 fs timesteps (4 million steps) to allow the sample to reach a steady-state condition. Atomic energies are then collected at every time step for an additional 2 ns (2 million steps), which are averaged to obtain the temperature profile along the x-axis. Considering that the mixing regions of the systems have different thicknesses, the temperature change at the interface is evaluated by extrapolating the linear part of the temperature profile in each material to the interface, where ΔT is taken as the difference of the two values, as shown in Figure 1b. We found that the distribution of temperature drops across the mixing region was sensitive to the choice of regions supporting each linear fit. From a preliminary data set, the end points for each linear region were optimized to ensure the fitting function works for all the systems with various intermixing arrangements while maintaining the quality of the linear fit (with $R^2 \geq 0.98$). For a given heat flux, the change in the ITC is monitored by keeping track of ΔT.

2.1.4. Uncertainty in MD Simulations. Small deviations can have large effects during high-throughput optimization over a large sample space. With randomness and uncertainties in MD simulations, assessing the uncertainty upfront is important. Previous studies have demonstrated that the MD-predicted thermal boundary resistance (R_k) depends on the size of the system L_s, where $L_s = L_L + L_{mix} + L_R$, and the thermal reservoir (L_{res}) used in the simulation.[13,26,32] Therefore, we kept L_s and L_{res} constant over all simulations. We indeed found that L_s changes slightly during the simulation with varying intermixing depth and percentage. However, the magnitude of this change is <0.5% in all cases, which is sufficiently small and can be neglected. The size dependence of R_k on L_s, which in principle can only be eliminated when the system is infinitely large, still exists and can potentially affect the optimization of intermixing. According to the study of Schelling et al., for the same interface there is a linear correlation between R_k and $1/L_s$;[26] thus, it was expected that in our systems, each distinct intermixing ratio or pattern warrants a new interface with a different linear correlation. However, by conducting a convergence study we find that the linear correlation is consistent for the same intermixing ratio and region regardless of the random atomic arrangement (Figure S1). Therefore, linear extrapolation can still be applied to extend our results to systems of varied lengths. Still, a convergence study is needed for the size effect on the optimization of intermixing, which is introduced in the Results section.

Significant progress in simulation methodology and computing power has enabled simulations of increasingly large systems and long time scales.[16−18,36,37] However, the accuracy of these simulations in predicting physical observables remains variable and relies on both accurate energy functions and effective sampling. With accurate energy functions, adequate sampling of a desired phase space is essential in achieving low associated error of estimated thermodynamic and kinetic properties. Simulations generally consist of equilibration and production phases, where the systems evolve during equilibration to be representative of the desired ensemble before data are collected during production run to derive physical properties. It is necessary to thoroughly investigate the effect of equilibration and data collection periods to determine the appropriate time scale to ensure both the accuracy and efficiency of the simulations while comparing the results of independent MD in a self-consistent manner. To isolate the effect of intermixing, both pristine and intermixed interfaces are examined in this particular study. We chose 15 different combinations of $[\tau_{ss}, \tau_{av}]$ to examine the sensitivity of ITC to the simulation time, where τ_{ss} is the time allowed for the temperature to reach steady state after introducing the heat flux and τ_{av} is the time used for collecting and averaging atomic energies to obtain the temperature profiles.

2.2. Machine Learning Domain. *2.2.1. Bayesian Optimization.* Bayesian optimization is a sample-efficient and gradient-free approach to global optimization black-box functions. The goal of this technique is to bound the maxima of the random objective function, $f(\mathbf{x};\xi)$ (where ξ represents the random input upon which f depends) by using the fewest number of samples that satisfy the accuracy requirement or the computational budget.[29,38,39]

Bayesian optimization[40] consists of two main components: (1) a Bayesian statistical model for modeling the objective function and (2) an acquisition function to decide the next point to sample. The statistical model provides a Bayesian posterior probability distribution that describes potential values for $\mathbf{x}$ at a candidate point x. Each time we observe f

https://doi.org/10.1021/acsami.1c23222
ACS Appl. Mater. Interfaces 2022, 14, 32590−32597

at a new point, this posterior distribution is updated. The acquisition function measures the value that would be generated by evaluation of the objective function at a new point x based on the current posterior distribution over f.

The Gaussian process (GP) is a generalization of the Gaussian probability distribution. It is a convenient and powerful prior distribution on functions and are used in majority of the Bayesian optimization work. GPs are defined by their mean ($\mathbf{m}$) and covariance/kernel functions ($k(x_i,x_j)$). $\mathbf{m}$ is almost always assumed a constant value (often assumed 0) which simplifies the necessary equations for conditioning. We parametrized the covariance kernel as a composite of a two-dimensional radial basis function and added a white noise kernel to account for the noise from MD simulations. Specifically

$$k(x_i, x_j) = C\left(\exp\left[\frac{-d(x_i, x_j)^2}{2{l_1}^2}\right] + \exp\left[\frac{-d(x_i, x_j)^2}{2{l_2}^2}\right]\right) + \delta(x_i - x_j) \quad (2)$$

Here we define C as a constant term, $d = x_i - x_j$, and l_i is the parameter that sets the characteristic length scale of the process for each dimension.

The second component of Bayesian optimization is an acquisition function, which defines how to explore/exploit the parameter space for a defined figure-of-merit. Normally, the high values of acquisition functions correspond to potentially high values of the objective function, whether for high predicted mean value, the greater uncertainty, or both.[41] The upper confidence bound (UCB) is a popular acquisition function; it calculates the most optimistic improvement at a given point, weighted by a scalar multiplier, β, of the noise estimation, as follows:

$$\mathrm{UCB}(\mathbf{x}) = \mu(\mathbf{x}) + \beta\sigma(\mathbf{x}) \quad (3)$$

The value of β balances between exploiting the supposed location of the maximum where the posterior mean $\mu(\mathbf{x})$ is high and exploring uncertain regions where the posterior variance $\sigma(\mathbf{x})$ is high. In sequential optimization, argmax UCB($\mathbf{x}$) is then taken as the next sampling point. Because the time-consuming nature of the MD simulations necessitates batch evaluation to achieve optimization within a reasonable time frame, we implemented a quantile-based adaptive sampling algorithm based on UCB, as described in Algorithm 1, to select a batch of sampling points. As opposed to using the maximum UCB value to determine the next sampling point, we randomly sampled a batch of test points from the quantile of the GP above the UCB, hereafter called the quantile region. As the optimization proceeds, we restrict the quantile region further to the upper P_Q percent of heat transfer rate values, $1/\Delta T$, in the quantile region by evaluating the threshold Q, which is the lower bound on the heat transfer rate value in this restricted region. Considering higher variance in ΔT values acquired from the MD simulations for the intermixed cases, $\beta = 5$ is taken to ensure the GP regions with high uncertainties are not excluded from the reduced sampling region at the earlier stage of the optimization.

2.2.2. MPI Assisted Parallel Optimization Framework. Although Bayesian optimization is an efficient technique, the computational cost of MD is still comparatively high. To accelerate the exploration/exploitation process, we developed a Message Passing Interface (MPI) based framework that can execute multiple concurrent LAMMPS simulations and update

Algorithm 1: Quantile-based adaptive sampling

Input: f: function to be optimized
$\mathcal{X}$: sampling space
$\mathcal{D}$: initial dataset $\{(x_1, y_1), ..., (x_n, y_n)\}$
T: number of iterations
b: batch size
P_Q: percentage of the quantile region //starts with 100%
Q: quantile threshold value
ε : decay rate

```
GP = model_init(D)
for t in 1,...,T, do
    UCB(x) = μ_{t-1}(x) + βσ_{t-1}(x) // Calculate UCB using full sampling space ;
    Q = lower bound of the top P_Q quantile of UCB(x);
    R = {x ∈ X|UCB(x) ≥ Q} // Update quantile region ;
    for i in 1,...,b, do
        x_i = random.choice(R) // randomly select sampling points from quantile
          region ;
        Query x_i to obtain y_i;
    end
    D = D ∪ {(x_1, y_1), ..., (x_b, y_b)};
    GP.update_model;
    P_Q = P_Q × ε;
end
```

a central Bayesian optimization model, as shown in Figure 2. Child processes are created with the MPI spawning function.

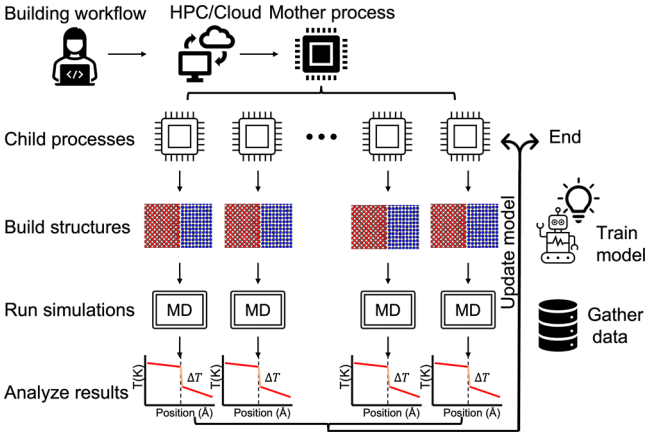


Figure 2. Schematic diagram of the MPI-based framework used to achieve batch optimization. Each of the child processes is initiated with **MPI_SPAWN** and interfaces with LAMMPS by using the OpenAI gym environment.

The spawned processes have their own **MPI_COMM_WORLD** consisting of identical copies of the MPI program specified by the command, and in addition they have a predefined intercommunicator **MPI_COMM_PARENT**.[42] The open-source Python library OpenAI Gym[43] is used in this work for creating environment instances and interacting with them. It provides a standard application programming interface to communicate between Bayesian optimization algorithm and the environment in which sampling points are queried by building the intermixing structure, obtaining temperature profile with LAMMPS and performing post-simulation analyses to evaluate ΔT. A more detailed description of the OpenAI Gym environment is included in the Supporting Information.

3. RESULTS AND DISCUSSION

As detailed in the methodology section, we used a GP model as the probabilistic surrogate for the black-box objective function we seek to optimize which describes the relationship between intermixing and ITC. ITC is calculated as $G = J/\Delta T$. For systems with mixing regions no larger than two atomic layers, we obtained ITC between 0.35 and 0.48 GW/(m^2 K), which falls in the range predicted by mismatch models and

https://doi.org/10.1021/acsami.1c23222
ACS Appl. Mater. Interfaces 2022, 14, 32590−32597

similar simulations.[44] These theoretical values are higher than experimental measurements,[45] likely due to the ideal conditions in simulations without oxidation. In addition to the idealized interfaces, the choice of the interatomic potentials as well as the finite system size effect can also lead to the discrepancy between experiment and simulations. Because we fixed the heat flux J for all systems, we defined the figure-of-merit for the optimization as the inverse of the temperature drop across the interface, $1/\Delta T$, which is used as an analogue for thermal conductivity. Snapshots of the GP-UCB hypersurface throughout the optimization are shown in Figure 3.

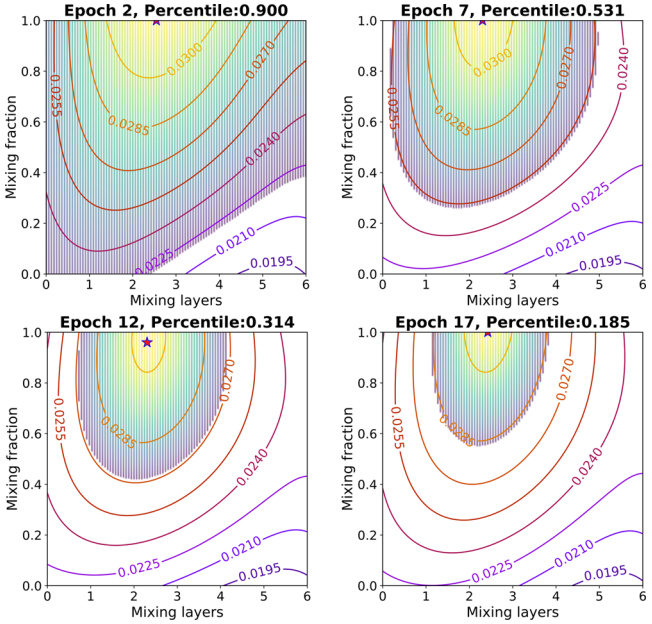


Figure 3. Illustration of the GP-UCB surface evolution throughout the optimization. The contours indicate the topography of the GP-UCB surface in terms of the heat transfer rate, $1/\Delta T$. The shaded region is the quantile region at a given epoch, which shrinks during optimization as the learning rate forces P_Q to decrease (called "percentile" in each subplot) and consequently the heat transfer rate threshold Q to increase.

Throughout the optimization, the shape of the GP hypersurface evolves and converges while the quantile region continues to shrink with each iteration. A batch of 10 sampling points selected from the evolving quantile region is evaluated (with NEMD simulation) per iteration to accelerate the exploration/exploitation process. Resulting $1/\Delta T$ values obtained from these simulations show large variance especially at early iterations of the optimization, as shown in Figure S2a, due to the fact that we are evaluating the points in the quantile region with low a quantile threshold value. With each iteration of the optimization, as the quantile region shrinks, points with higher predicted GP values are evaluated, and the mean $1/\Delta T$ value of the batch also increases. The GP predicted maximum $1/\Delta T$ values for each iteration plotted in Figure S2b show the convergence of the GP model only after six iterations with minimal variance. A convergence study shows that optimizations of systems with different lengths tend to reach the same intermixing percentile region (Figure S3). This, together with the applicability of linear extrapolation of our results, suggests that our current data set is capable of representing a wider range of Si/Al systems of varying lengths. In the following paragraphs we will look into the sources of the large variance from MD data and discuss the significance of the intrinsic uncertainties to the model training and evaluation.

A snapshot comparison of three mixing profiles (pristine, optimal, and negative impact) is shown in Figure 4. The results from the optimization scan show that the ITC increases by as much as 50% relative to the pristine interface with a mixing region of two atomic layers at 93% intermixing. Previous work based on the atomistic Green's function has demonstrated that interfacial atomic intermixing can enhance the elastic phonon transmission by providing a region with bridging phonon density of states that overlaps with both materials, thus increasing the ITC,[12] while more recent work based on classical MD suggests that enhanced inelastic phonon transport contributes significantly at such deformed interfaces.[46] We attribute the ITC enhancement in our study to both factors, considering the similarity between our modeling environment and that of Lu et al.[46] However, further increasing the size of the mixing region beyond the apex values reduces the ITC. This can be explained by increased Anderson localization in

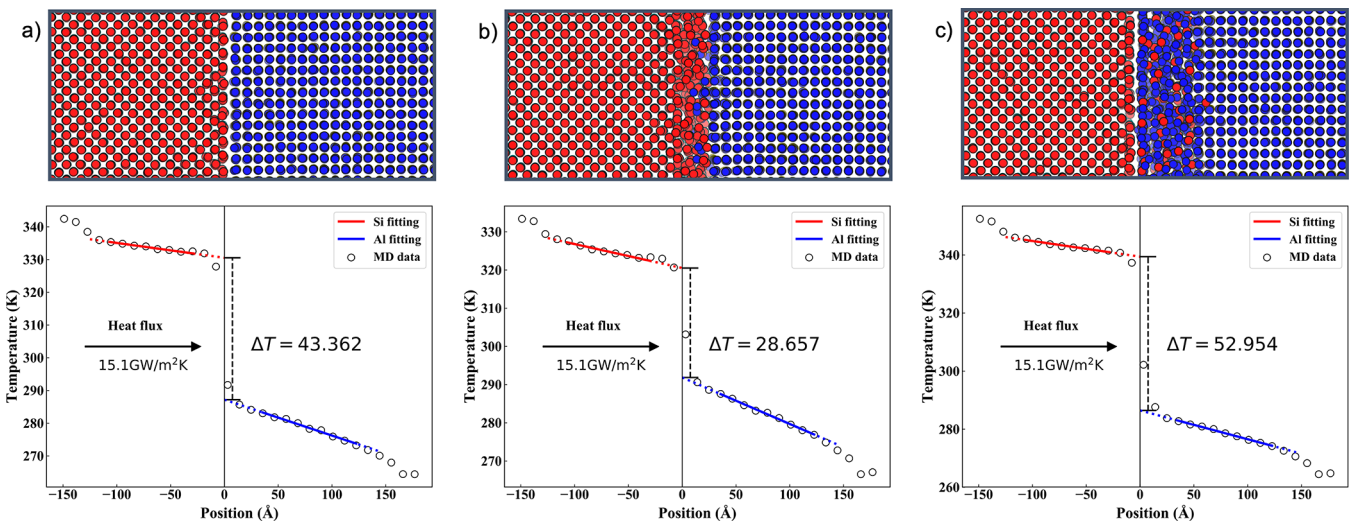


Figure 4. Temperature profiles of three interfacial mixing configurations: (a) pristine interface, (b) two-layer intermixing at 93%, and (c) six-layer intermixing at 50%.

https://doi.org/10.1021/acsami.1c23222
ACS Appl. Mater. Interfaces 2022, 14, 32590−32597

the disordered mixing region,[47,48] which negates the increase introduced by enhanced elastic and inelastic phonon transport. It is noteworthy that even with 50% increase, the ITC does not reach the value predicted by the acoustic mismatch model.[44]

A 50% increase of the ITC, despite the significance, is not always guaranteed. This can be attributed to two factors. One is the intrinsic errors from the MD simulations which can be minimized but cannot be eliminated. The other is the random positioning of the intermixed atoms leading to different intermixed structures for the same intermixing length and ratio. NEMD is a powerful technique to simulate heat transfer processes. Still, no single method is perfect. With systematic and random sources of errors accruing in MD simulations, assessing the uncertainties is important to support active-learning strategies. In our work, we carefully evaluated the system design, parametrization, and analysis of simulation result to minimize the variation in the results to be less than the quantity of interest. As stated in the Setup section, with length of the system and thermal reservoir kept constant over all simulations, resulting ITCs are directly comparable. Simulation duration is another important parameter. We examined 15 different combinations of $[\tau_{ss}, \tau_{av}]$ for both pristine and intermixed structures, and the distributions of the resulting ΔT values are plotted in Figure 5. An additional figure with different ordering (varying τ_{av} for τ_{ss}) is included in the Supporting Information (Figure S4).

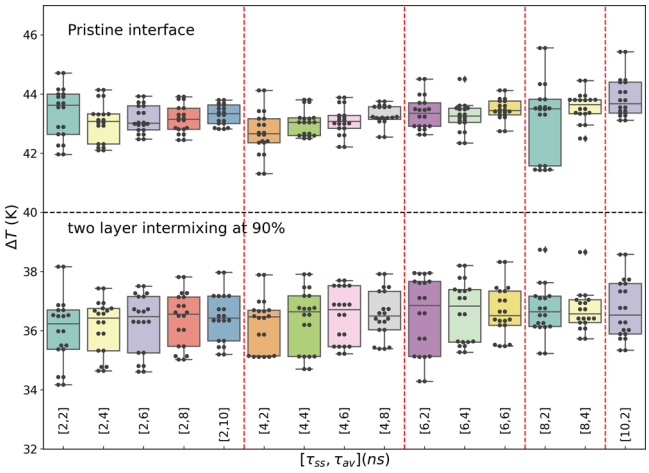


Figure 5. Various run times and distribution of resulting ΔT for pristine and intermixed interfaces. Each box plot contains 20 data points. Vertical dashed red lines are used to separate each set of τ_{ss}.

The relative variance (variance divided by mean) of ΔT ranges from 0.27% to 4.0% for the pristine interface and 1.69% to 4.72% for the intermixed cases. A detailed statistical description of each entry is listed in the Supporting Information (Table S1). For a fixed τ_{ss}, the variance of ΔT decreases with increasing data collection time τ_{av}. In theory, the uncertainties can be reduced to a certain degree (the intrinsic uncertainty of MD from its stochastic nature) with increased computational time. However, this increases the cost of the simulation significantly and the low cost-effectiveness destroys the ML workflow. For our system, each additional nanosecond of NVE integration adds ~256 core hours to the simulation. Increasing the $[\tau_{ss}, \tau_{av}]$ from [4 ns, 2 ns] to [4 ns, 8 ns] doubles the computational cost from ~1700 to ~3400 core hours, while lowering the relative variance roughly from 1.37% to 0.27% for the pristine interface and only from 2.21% to 1.82% for the intermixed case. Considering the trade-off between computational cost and accuracy, and that each data point requires a separate MD simulation, we equilibrate the systems for 4 ns before collecting data for an additional 2 ns to calculate the temperature gradient.

Comparing the pristine interface with the intermixed, we observed greater variance in ΔT for the intermixed cases. Because the positions of the swapped atoms in the intermixing layers are assigned randomly when generating the intermixed structure, homogeneity/heterogeneity of the intermixed region is not controlled. Lu et al.[46] also found the distribution of disorder to affect the ITC. Swapping atoms at the boundary layer rather than in layers further from the interface improves ITC enhancement, as this provides better atomic intermixing and more inelastic scattering sites at the interface. Introducing changes away from the interface has a reduced effect on ITC and introduces extra inelastic scattering in the bulk phase of the material, which hinders phonon transport.[46] It is important to highlight that despite the greater variation in results, there is no overlap between the temperature gradients of intermixed versus pristine interfaces, clearly indicating the improvement of thermal conductivity with intermixing. We note that the percentage of enhancement also depends on the total length of the system, as is indicated by the size effect on ITC (Figure S1); thus, a 50% increase is not guaranteed for systems of sizes other than those used in this study.

4. CONCLUSIONS

This study highlights the advantages of Bayesian optimization technique with built-in uncertainties leveraging the errors from MD to yield statistically robust and reproducible results. We pushed the model to explore more in the higher confidence region using an adaptive quantile-based confidence interval approach (1) to prevent early lock-in on local maxima, (2) to increase the confidence in the result, and (3) to generate multiple nonoverlapping plausible query points to expedite the optimization. Parallel evaluation of candidate structures with MD simulations, made possible by dynamic spawning of jobs through the Message Passing Interface (MPI), ensured viability of optimization within a reasonable time frame. This framework—seamlessly incorporating MD simulations with ML algorithms—is easily transferrable to any system with well-developed interatomic potentials. The GP model suggests that the ITC of the Si/Al junction can be enhanced by introducing an intermixed region that consists of three or fewer atomic layers containing 80% or higher ratio of Si. Further extending the intermixing region increases Anderson localization in the disordered mixing region and negatively impacts the ITC. Because of the random nature of intermixing, the magnitude of increase in the ITC varies. Further efforts are ongoing to better understand the relationship between ITC and distribution of the intermixing.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsami.1c23222.

System size dependence of ΔT on $1/L$, description of the OpenAI Gym environment, GP model convergence result, results from the smaller system size showing the similar convergence of the GP model, additional figure

https://doi.org/10.1021/acsami.1c23222
ACS Appl. Mater. Interfaces 2022, 14, 32590−32597

of the uncertainty analysis result with different ordering (varying τ_{av} for τ_{ss}), and a table containing the statistical descriptions of the uncertainty analysis result (PDF)

AUTHOR INFORMATION

Corresponding Authors

Sabiha Rustam − *Department of Chemical Engineering, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0002-2424-8153; Email: sabiha3@uw.edu

Malachi Schram − *Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606, United States;* Email: schram@jlab.org

Authors

Zexi Lu − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States;* orcid.org/0000-0002-8343-9902

Anne M. Chaka − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States;* orcid.org/0000-0001-7872-7590

W. Steven Rosenthal − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States*

Jim Pfaendtner − *Department of Chemical Engineering, University of Washington, Seattle, Washington 98195, United States; Pacific Northwest National Laboratory, Richland, Washington 99352, United States;* orcid.org/0000-0001-6727-2957

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.1c23222

Author Contributions

S.R., M.S., Z.L., and A.M.C designed the research. S.R. and M.S. executed the research and analyzed the results. The manuscript was written through contributions of all authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the Pacific Northwest National Laboratory (PNNL) Lab Directed Research and Development (LDRD) program as part of the Chemical Dynamics Initiative and under the Data Intensive Research Enabling Clean Technology (DIRECT) National Science Foundation (NSF) National Research Traineeship program (DGE-1633216). The research was performed by using computational resources provided by PNNL Institutional Computing (PIC). PNNL is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC05-76RL01830. Work performed while Malachi Schram was at Pacific Northwest National Laboratory.

REFERENCES

(1) Pan, L.; Bogy, D. B. Heat-Assisted Magnetic Recording. *Nat. Photonics* **2009**, *3*, 189−190.

(2) Zhang, X.; Zhao, L.-D. Thermoelectric Materials: Energy Conversion Between Heat and Electricity. *Journal of Materiomics* **2015**, *1*, 92−105.

(3) Xu, X.; Zhang, Q.; Hao, M.; Hu, Y.; Lin, Z.; Peng, L.; Wang, T.; Ren, X.; Wang, C.; Zhao, Z.; Wan, C.; Fei, H.; Wang, L.; Zhu, J.; Sun, H.; Chen, W.; Du, T.; Deng, B.; Cheng, G. J.; Shakir, I.; Dames, C.; Fisher, T. S.; Zhang, X.; Li, H.; Huang, Y.; Duan, X. Double-Negative-Index Ceramic Aerogels for Thermal Superinsulation. *Science* **2019**, *363*, 723−727.

(4) Chen, R.; Lee, J.; Lee, W.; Li, D. Thermoelectrics of Nanowires. *Chem. Rev.* **2019**, *119*, 9260−9302.

(5) Hopkins, P. E.; Duda, J. C.; Petz, C. W.; Floro, J. A. Controlling Thermal Conductance through Quantum Dot Roughening at Interfaces. *Phys. Rev. B* **2011**, *84*, 035438.

(6) Hopkins, P. E.; Norris, P. M.; Stevens, R. J.; Beechem, T. E.; Graham, S. Influence of Interfacial Mixing on Thermal Boundary Conductance Across a Chromium/Silicon Interface. *Journal of Heat Transfer* **2008**, *130*, 062402.

(7) Duda, J. C.; Hopkins, P. E. Systematically Controlling Kapitza Conductance via Chemical Etching. *Appl. Phys. Lett.* **2012**, *100*, 111602.

(8) Shao, C.; Rong, Q.; Hu, M.; Bao, H. Probing Phonon-Surface Interaction by Wave−Packet Simulation: Effect of Roughness and Morphology. *J. Appl. Phys.* **2017**, *122*, 155104.

(9) Ih Choi, W.; Kim, K.; Narumanchi, S. Thermal Conductance at Atomically Clean and Disordered Silicon/Aluminum Interfaces: A Molecular Dynamics Simulation Study. *J. Appl. Phys.* **2012**, *112*, 054305.

(10) Ran, X.; Guo, Y.; Wang, M. Interfacial Phonon Transport with Frequency-Dependent Transmissivity by Monte Carlo Simulation. *Int. J. Heat Mass Transfer* **2018**, *123*, 616−628.

(11) Duda, J. C.; English, T. S.; Piekos, E. S.; Beechem, T. E.; Kenny, T. W.; Hopkins, P. E. Bidirectionally Tuning Kapitza Conductance through the Inclusion of Substitutional Impurities. *J. Appl. Phys.* **2012**, *112*, 073519.

(12) Tian, Z.; Esfarjani, K.; Chen, G. Enhancing Phonon Transmission Across a Si/GE Interface by Atomic Roughness: First-Principles Study with the Green's Function Method. *Phys. Rev. B* **2012**, *86*, 235304.

(13) Landry, E. S.; McGaughey, A. J. H. Effect of Interfacial Species Mixing on Phonon Transport in Semiconductor Superlattices. *Phys. Rev. B* **2009**, *79*, 075316.

(14) Stevens, R. J.; Zhigilei, L. V.; Norris, P. M. Effects of Temperature and Disorder on Thermal Boundary Conductance at Solid-Solid Interfaces: Nonequilibrium Molecular Dynamics Simulations. *Int. J. Heat Mass Transfer* **2007**, *50*, 3977−3989.

(15) English, T. S.; Duda, J. C.; Smoyer, J. L.; Jordan, D. A.; Norris, P. M.; Zhigilei, L. V. Enhancing and Tuning Phonon Transport at Vibrationally Mismatched Solid-Solid Interfaces. *Phys. Rev. B* **2012**, *85*, 035438.

(16) Moore, G. E. Cramming More Components onto Integrated Circuits. *Proceedings of the IEEE* **1998**, *86*, 82−85.

(17) Vendruscolo, M.; Dobson, C. M. Protein Dynamics: Moore's Law in Molecular Biology. *Curr. Biol.* **2011**, *21*, R68−R70.

(18) Dror, R. O.; Jensen, M. Ø.; Borhani, D. W.; Shaw, D. E. Exploring Atomic Resolution Physiology on a Femtosecond to Millisecond Timescale Using Molecular Dynamics Simulations. *J. Gen. Physiol.* **2010**, *135*, 555−562.

(19) Yen, J.; Shank, B.; Young, B.; Cabrera, B.; Brink, P.; Cherry, M.; Kreikebaum, J. M.; Moffatt, R.; Redl, P.; Tomada, A.; Tortorici, E. Measurement of Quasiparticle Transport in Aluminum Films Using Tungsten Transition-Edge Sensors. *Appl. Phys. Lett.* **2014**, *105*, 163504.

(20) Koch, J.; Yu, T. M.; Gambetta, J.; Houck, A. A.; Schuster, D. I.; Majer, J.; Blais, A.; Devoret, M. H.; Girvin, S. M.; Schoelkopf, R. J. Charge-Insensitive Qubit Design Derived from the Cooper Pair Box. *Phys. Rev. A* **2007**, *76*, 042319.

(21) Alexandera, A.; McSkimming, B. M.; Arey, B.; Arslan, I.; Richardson, C. J. Nucleation and Growth of Metamorphic Epitaxial Aluminum on Silicon (111) 7×7 and $\sqrt{3} \times \sqrt{3}$ Surfaces. *J. Mater. Res.* **2017**, *32*, 4067.

(22) Majumdar, A.; Reddy, P. Role of Electron-Phonon Coupling in Thermal Conductance of Metal-Nonmetal Interfaces. *Appl. Phys. Lett.* **2004**, *84*, 4768−4770.

(23) Wang, Y.; Ruan, X.; Roy, A. K. Two-Temperature Non-equilibrium Molecular Dynamics Simulation of Thermal Transport Across Metal-Nonmetal Interfaces. *Phys. Rev. B* **2012**, *85*, 205311.
(24) Lu, Z.; Wang, Y.; Ruan, X. The Critical Particle Size for Enhancing Thermal Conductivity in Metal Nanoparticle-Polymer Composites. *J. Appl. Phys.* **2018**, *123*, 074302.
(25) Kubo, R.; Toda, M.; Hashitsume, N. *Statistical Physics II: Nonequilibrium Statistical Mechanics*; Springer Science & Business Media: 2012; Vol. *31*.
(26) Schelling, P. K.; Phillpot, S. R.; Keblinski, P. Comparison of Atomic-Level Simulation Methods for Computing Thermal Conductivity. *Phys. Rev. B* **2002**, *65*, 144306.
(27) Kaufmann, E.; Cappé, O.; Garivier, A. On Bayesian Upper Confidence Bounds for Bandit Problems. *Artificial intelligence and statistics* **2012**, 592−600.
(28) Auer, P. Using Confidence Bounds for Exploitation-Exploration Trade-Offs. *Journal of Machine Learning Research* **2002**, *3*, 397−422.
(29) Frazier, P. I. A Tutorial on Bayesian Optimization; *arXiv:1807.02811*. **2018**.
(30) Srinivas, N.; Krause, A.; Kakade, S. M.; Seeger, M. W. Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting. *IEEE Transactions on Information Theory* **2012**, *58*, 3250−3265.
(31) Müller-Plathe, F. A Simple Nonequilibrium Molecular Dynamics Method for Calculating the Thermal Conductivity. *J. Chem. Phys.* **1997**, *106*, 6082−6085.
(32) Landry, E. S.; McGaughey, A. J. H. Thermal Boundary Resistance Predictions from Molecular Dynamics Simulations and Theoretical Calculations. *Phys. Rev. B* **2009**, *80*, 165304.
(33) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(34) Jelinek, B.; Groh, S.; Horstemeyer, M. F.; Houze, J.; Kim, S. G.; Wagner, G. J.; Moitra, A.; Baskes, M. I. Modified Embedded Atom Method Potential for Al, Si, Mg, Cu, and Fe Alloys. *Phys. Rev. B* **2012**, *85*, 245102.
(35) POLLACK, G. L. Kapitza Resistance. *Rev. Mod. Phys.* **1969**, *41*, 48−81.
(36) Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, C. R.; Ierardi, D. J.; Kolossvary, I.; Klepeis, J. L.; Layman, T.; McLeavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y.; Spengler, J.; Theobald, M.; Towles, B.; Wang, S. C. Anton, a Special-Purpose Machine for Molecular Dynamics Simulation. *Communications of the ACM* **2008**, *51*, 91−97.
(37) Shaw, D. E.; Grossman, J. P.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H.; Forte, A.; Gagliardo, J.; Gill, G.; Greskamp, B.; Ho, C. R.; Ierardi, D. J.; Iserovich, L.; Kuskin, J. S.; Larson, R. H.; Layman, T.; Lee, L. S.; Lerer, A. K.; Li, C.; Killebrew, D.; Mackenzie, K. M.; Mok, S.; Moraes, M. A.; Mueller, R.; Nociolo, L. J.; Peticolas, J. L.; Quan, T.; Ramot, D.; Salmon, J. K.; Scarpazza, D. P.; Ben Schafer, U.; Siddique, N.; Snyder, C. W.; Tang, P. T. P.; Theobald, M.; Toma, H.; Towles, B.; Vitale, B.; Wang, S. C.; Young, C. *Anton 2: Raising the Bar for Performance and Programmability in a Special- Purpose Molecular Dynamics Supercomputer*. International Conference for High Performance Computing, Networking, Storage and Analysis, SC. **2014**; pp 41−53.
(38) Brochu, E.; Cora, V. M.; de Freitas, N. *A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning*. **2010**.
(39) Boender, C. G. E.; Mockus, J. Bayesian Approach to Global Optimization−Theory and Applications. *Mathematics of Computation* **1991**, *56*, 878.
(40) Mockus, J.; Tiesis, V.; Zilinskas, A. The Application of Bayesian Methods for Seeking the Extremum. *Towards Global Optimization* **1978**, *2*, 2.
(41) Hoffman, M.; Brochu, E.; de Freitas, N. Portfolio Allocation for Bayesian Optimization. *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence* **2011**, 327−336.
(42) Walker, D. W. *Standards for Message-Passing in a Distributed Memory Environment*. **1992**.
(43) Brockman, G.; Cheung, V.; Pettersson, L.; Schneider, J.; Schulman, J.; Tang, J.; Zaremba, W. OpenAI Gym, 2016.
(44) Zhang, Y.; Ma, D.; Zang, Y.; Wang, X.; Yang, N. A Modified Theoretical Model to Accurately Account for Interfacial Roughness in Predicting the Interfacial Thermal Conductance. *Frontiers in Energy Research* **2018**, DOI: 10.3389/fenrg.2018.0004.
(45) Hopkins, P. E.; Phinney, L. M.; Serrano, J. R.; Beechem, T. E. Effects of Surface Roughness and Oxide Layer on the Thermal Boundary Conductance at Aluminum/Silicon Interfaces. *IHTC* **2010**, 313−319.
(46) Lu, Z.; Chaka, A. M.; Sushko, P. V. Thermal Conductance Enhanced via Inelastic Phonon Transport by Atomic Vacancies at Cu/Si Interfaces. *Phys. Rev. B* **2020**, *102*, 075449.
(47) Anderson, P. W. Absence of Diffusion in Certain Random Lattices. *Phys. Rev.* **1958**, *109*, 1492−1505.
(48) Roy Chowdhury, P.; Reynolds, C.; Garrett, A.; Feng, T.; Adiga, S. P.; Ruan, X. Machine Learning Maximized Anderson Localization of Phonons in Aperiodic Superlattices. *Nano Energy* **2020**, *69*, 104428.