

Designing thermal functional materials by coupling thermal transport calculations and machine learning

Cite as: J. Appl. Phys. 128, 161102 (2020); doi: 10.1063/5.0017042

Submitted: 9 June 2020 · Accepted: 2 October 2020 ·

Published Online: 27 October 2020



View Online



Export Citation



CrossMark

Shenghong Ju,^{1,2,3,4} Shuntaro Shimizu,⁴ and Junichiro Shiomi^{4,5,a} 

AFFILIATIONS

¹China-UK Low Carbon College, Shanghai Jiao Tong University, Shanghai 201306, China

²School of Material Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

³Materials Genome Initiative Center, Shanghai Jiao Tong University, Shanghai 200240, China

⁴Department of Mechanical Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan

⁵Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), 4-1-8, Kawaguchi, Saitama 332-0012, Japan

^aAuthor to whom correspondence should be addressed: shomi@photon.t.u-tokyo.ac.jp

ABSTRACT

Advances in materials informatics (MI), which combines material property calculations/measurements and informatics algorithms, have realized properties in the nanostructures of thermal functional materials beyond what is accessible using empirical approaches based on physical instincts and models. In this Tutorial, we introduce technological procedures and underlying knowledge of MI combining thermal transport calculations and machine learning using an optimization problem of superlattice structures as an example (sample script available in the supplement). To provide fundamental guidance on how to use MI, we describe practical details about descriptors, objective functions, property calculators, machine learning (Bayesian optimization) algorithms, and optimization efficiencies. We then briefly review the recent successful applications of MI to design thermoelectric and thermal radiation materials. Finally, we summarize and provide future perspectives about the topic.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0017042>

I. INTRODUCTION

The rapid development of nanotechnology has introduced unprecedented opportunities in designing nanostructures for energy transport in the forms of phonons, electrons, and photons. Phonon transport becomes more ballistic when the material length scale is reduced to the nanoscale. Forming the appropriate nanostructure can control heat conduction. Such control has been widely realized in superlattices,^{1,2} phononic crystals,^{3–5} nanocrystals,^{6–8} and nanocomposites.^{9,10} Furthermore, when the length scale of the nanostructure periodicity exceeds the phase-relaxation lengths of the phonons that contribute to the thermal transport and the nanostructure-boundary roughness is smaller than the wavelengths,^{11,12} nanostructures can exhibit various coherent phonon transport phenomena such as constructive/deconstructive phonon interference,^{13–15} resonance,¹⁶ and localization.^{17,18} These phenomena can magnify the controllability of heat conduction.

A similar discussion can be made for heat conduction through interfaces. Various factors have been reported to tune the interfacial thermal conductance such as roughness,^{19,20} vacancy defects,²¹ lattice orientation,^{22,23} nanojunctions,²⁴ and interfacial adhesion or bonding.^{25,26} To further advance the field of thermal transport control by nanostructures, an important question is how much room is left to control for a specific class of nanostructures. That is, what is the optimal structure to minimize or maximize the thermal conductivity? However, this is a challenging question to answer because phonon transport is highly sensitive to the detailed atomistic configuration.

When phonon transport is coupled with other energy carrier transports such as electrons, transport property design becomes more complex. Considering thermoelectric applications as an example, the energy conversion efficiency of thermoelectric materials is characterized by the dimensionless figure of merit: $ZT = S^2 \sigma T / (\kappa_e + \kappa_{ph})$, where

S , σ , T , κ_e , and κ_{ph} are the Seebeck coefficient, electrical conductivity, absolute temperature, and thermal conductivities of phonons and electrons, respectively. Many studies have been dedicated to reduce the phonon thermal conductivity.^{7,27} However, increasing the figure of merit by altering only one component is difficult due to the general correlation between the electric and thermal transports.

In addition to the thermal transport in solids carried by phonons and electrons, photons can transfer heat via thermal radiation. By tuning the photon transport, wavelength- or directional-selective emitters/absorbers have been designed and applied to thermal photovoltaics,^{28,29} infrared heaters,³⁰ and thermal radiative cooling.^{31–33} Similar to phonon transport, when the structure length scale is comparable or smaller than the photon wavelength, micro-/nanostructures can induce interference and diffraction in the electromagnetic waves. This allows thermal radiation to be tailored. Although various structures such as nanowires, planar multilayers, periodic gratings, nanoparticles embedded in a matrix, and two-dimensional materials have been proposed, their optimization usually involves many variables: material type, layer thickness, grating size, particle size and distribution, etc. Hence, the number of possible candidates is enormous.

Along with theory, simulations, and experiments, materials informatics (MI)^{34–36} has become the fourth paradigm of science in the past few years. MI combines material property calculations/measurements and informatics algorithms. Not only has it guided experiments directly,³⁷ MI has a high efficiency in designing novel materials such as drugs,^{38,39} polymers,⁴⁰ thermoelectrics,^{41,42} magnetic tunnel junctions,⁴³ and thermal emitters/absorbers.^{44,45} Compared with traditional trial-and-error materials' research procedures, which are time-consuming and costly, MI has opened new avenues for accelerating the discovery, development, manufacturing, and deployment of advanced materials.

Many MI methods have been developed for energy transport optimization. However, they can be divided into two categories: database screening and structure designing. The main difference is that screening selects from available databases while designing creates novel materials or structures that do not yet exist.

Due to advances in computational materials science and computing power, more and more open-access material databases have been established. Examples include Materials Project,⁴⁶ AFLOW,⁴⁷ ICSD, OQMD,^{48,49} and NOMAD.⁵⁰ To select the most appropriate material from these databases for material design, high-throughput screening⁵¹ has been developed. Typical screening involves various machine learning algorithms such as neural networks,^{52,53} Gaussian process regression,⁵⁴ random-forest regression,⁵⁵ and transfer learning.⁵⁶

For structure design, an initial database is typically not available. Hence, training data are collected gradually, and the prediction model is updated over time via an iterative and closed-loop design process. Currently, various intelligent optimization algorithms such as the genetic algorithm,^{57,58} particle swarm optimization,⁵⁹ Bayesian optimization (BO),^{42,54,60–62} Monte Carlo tree search (MCTS),^{63,64} and quantum annealing (factorization machine)⁶⁵ are widely employed as black-box optimization tools. They realize automated materials' discovery based on an iterative process to select the optimal candidate from a massive number of potential structures.

This Tutorial introduces the automated design of thermal functional materials through a combination of thermal transport calculations and a Bayesian optimization algorithm. It offers guidance for researchers who are interested in efficiently designing novel materials for thermal transport using MI. This Tutorial is organized as follows. Section II explains the basis of MI using an example problem of designing a superlattice structure to obtain minimum thermal conductance. Section III discusses the optimization results, including the efficiency of optimization, the effect of the parameters on optimization, and new knowledge we can learn from MI. Section IV introduces recent applications of Bayesian optimizations on nanostructure designs for thermoelectrics⁴² and thermal radiation.^{44,45} Finally, Sec. V summarizes and gives perspectives for designing thermal functional materials using MI.

II. METHODOLOGY

Over the past few decades, various computational and experimental techniques for thermal transport have been developed at the nanoscale. Typically, the design is based on a physical model of thermal transport formulated in four dimensions (time and the three dimensions of space) either in a real or in a reciprocal representation, which is preferred for humans to comprehend. However, dimensional restrictions may become the limiting factor when identifying complicated structure–property relations. On the other hand, machine learning with a black-box model is advantageous in handling hyperspace of a large number of dimensions and may be superior in predicting complicated structure–property relations, though it may not immediately help humans understand the underlying physics. With advances in machine learning and data-mining methodologies, researchers are now able to construct a black-box model from computational or experimental data and identify or predict nanostructures or materials displaying the desired properties. Although these methods employ interpolation in a large-dimensional hyperspace, this can be viewed as an extrapolation in our usual four dimensions. Hence, it offers great opportunities for discovering, designing, and even understanding novel materials.

The key to this MI formalism is to build communications between materials' simulations/experiments and machine learning tools (Fig. 1). Generally, four essential elements are necessary for MI: objective function, descriptor selection, property calculator to evaluate the objective function, and an informatics optimization method. In the following, we introduce the general methodology to optimize the nanostructure for thermal transport.

A. Problem description and objective function

Before initiating any optimization, the optimization problem needs to be clearly defined with a specific objective function and the degrees of freedom to be optimized. Herein, we use the design of the superlattice structure as an example to demonstrate the methodology. Superlattice materials,^{66–69} which are important in applications such as thermoelectrics and semiconductor lasers, have attracted much attention in the past few decades due to their unique thermal and electrical transport properties.

Alternately, depositing two different materials on a substrate can synthesize a superlattice in a periodic or aperiodic fashion.

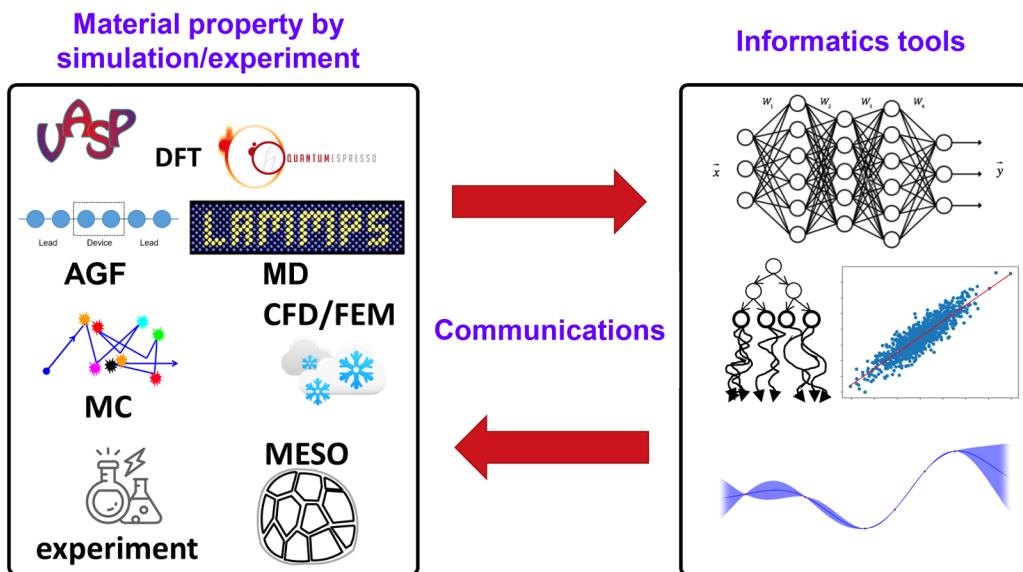


FIG. 1. Illustration of MI, which combines material property simulations/experiments and machine learning tools.

Periodic superlattices are well studied. Their thermal conductivity exhibits a minimum with respect to layer thickness due to the crossover between wave interference and particle scattering effects of phonons.⁷⁰ Recent studies have revealed that aperiodic superlattices have a lower thermal conductivity than periodic ones.⁷¹ Because the thickness of each layer in a superlattice differs, the

underlying physics becomes complicated. Consequently, designing an optimal superlattice structure from a vast number of candidates based on physical principles is a challenge.

Figure 2 illustrates a case where the design aims to realize an optimal structure to minimize thermal conductance at room temperature. The superlattice structure consists of silicon (Si) and

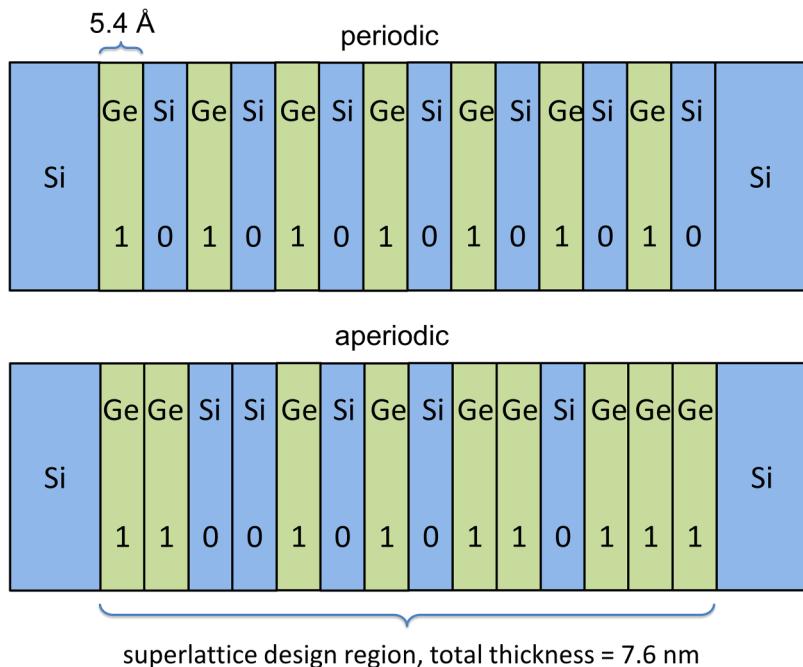


FIG. 2. Illustration of Si/Ge superlattice structures used for the structure-optimization problem to minimize the thermal conductance. Total thickness of the superlattice design region is 7.6 nm, which is divided into 14-unit layers that are either Si or Ge.

germanium (Ge) layers with a total thickness of 7.6 nm sandwiched by two Si leads. The superlattice structure is discretized into unit layers with thicknesses of 5.43 Å, which is the lattice constant of a silicon crystal. Each unit layer can be either Si or Ge. The objective is to find the optimal arrangement of Si and Ge layers to obtain the minimum thermal conductance. The objective function is the magnitude of thermal conductance or its reciprocal. To validate and discuss the optimization efficiency, the number of layers is intentionally kept to 14, which gives 16 384 possible structures. This number is small enough that all structures can be calculated for validation.

B. Descriptors

Descriptors, which are digital representations of the nanostructures, are used as inputs in machine learning models. Adopting an effective descriptor is important for successful machine learning. Various types of descriptors have been proposed in the literature for MI. Descriptors associated with the fundamental crystal properties are widely used since large amounts of data that are known to influence the transport properties in solid-state physics are available and compiled as databases. For instance, in database screening for high and low thermal conductivity crystals,^{51,54,72,73} the properties associated with the crystal structure are usually used as crystal descriptors. These include the lattice constants, the number of elements, the number of atoms in the cell, volume, density, formation energy, and space group number. In addition to independent parameters, secondary parameters, which are generated by mixing different parameters such as volume/atom, density/atom, and formation energy/atom, can also improve the prediction accuracy. Moreover, different types of elemental descriptors are important in screening. These include the atomic number, atomic weight, empirical covalent atomic radius, position in the periodic table, Pettifor chemical scale, and Pauling electronegativity. However, a descriptor does not have to be a physical property. Abstract quantities such as radial distribution function perform as well as or often even better than a physical property.

On the other hand, in the design of new nanostructures, descriptors must incorporate the spatial distribution of atoms. In the case of superlattice structure optimization, three types of descriptors are tested: digital flag numbers, the Coulomb matrix (CM), and the mass matrix (MM).⁷⁴

1. Digital flag numbers

Binary flags of “1” and “0,” which are convenient and intuitive, indicate the material state of each unit layer. Here, “1” indicates a Ge layer and “0” indicates a Si layer. For instance, the descriptor for the aperiodic superlattice structure shown in Fig. 1 is (110010110111). The 14 layers give rise to 16 384 possible candidates. It should be mentioned that decimal fraction values can also be used as descriptors. For instance, Guo *et al.*⁴⁴ used a discrete decimal fraction to describe the layer thicknesses in the multilayer structure introduced in Sec. IV.

2. Coulomb matrix

CM is a common descriptor in the field of molecular chemistry. It describes the Coulomb interaction between atoms.⁷⁵ For a

given atomistic structure, CM is calculated as

$$C = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{pmatrix}, \quad (1)$$

$$C_{ij} = \begin{cases} 0.5q_i^{2.4}, & (i = j), \\ \frac{q_i q_j}{r_{ij}}, & (i \neq j), \end{cases} \quad (2)$$

where q is the nuclear charge of the atom, r_{ij} is the distance between atoms i and j , and n is the number of atoms in the system. The diagonal elements in CM describe the self-interaction of each atom, which can be obtained by a polynomial fit of the atomic energies to the nuclear charge q , while the off-diagonal elements represent the Coulomb interactions between atoms i and j .

Because CM is symmetric, the actual descriptor vector can be obtained by adopting only the upper triangular part ($i \leq j$) in CM as

$$D_{\text{CM}} = (C_{11}, C_{12}, \dots, C_{1n}, C_{22}, C_{23}, \dots, C_{2n}, \dots, C_{nn})^T. \quad (3)$$

Therefore, the dimension size of CM is $n(n+1)/2$. Based on the Coulomb matrix, the eigenvalues of CM (EVCM)^{75,76} can be calculated and used as a descriptor,

$$D_{\text{EVCM}} = (\lambda_1, \lambda_2, \dots, \lambda_n)^T, \quad (4)$$

where λ_i , $i \in (1, 2, \dots, n)$, are the eigenvalues of C . In this way, the dimension size of EVCM can be reduced from $n(n+1)/2$ to n compared with CM, decreasing the search space. In general, compressing the dimension size of a descriptor is preferable for an optimization. However, over-compression may not sufficiently represent the actual structural information, negatively affecting the optimization. Another advantage of a CM descriptor is that it is size-independent when a cut-off for the interaction for a complicated irregular atomistic structure is applied.

3. Mass matrix

Similar to the definition of CM, the MM descriptor, which considers the mass information instead of charge, is defined as

$$M = \begin{pmatrix} 0 & M_{12} & \dots & M_{1n} \\ M_{21} & 0 & \dots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \dots & 0 \end{pmatrix}, \quad (5)$$

$$M_{ij} = \begin{cases} 0, & (i = j), \\ \frac{m_i m_j}{r_{ij}}, & (i \neq j), \end{cases} \quad (6)$$

where m is the atomic weight of the atom. The diagonal elements expressing the self-interaction are set to 0, considering that heat transport occurs via interatomic interactions. Similar with CM, the

final descriptors of MM and the eigenvalues of MM (EVMM) can be defined as

$$D_{\text{MM}} = (M_{12}, M_{13}, \dots, M_{1n}, M_{23}, \dots, M_{2n}, \dots, M_{(n-1)n})^T, \quad (7)$$

$$D_{\text{EVMM}} = (\mu_1, \mu_2, \dots, \mu_n)^T, \quad (8)$$

where μ_i , $i \in (1, 2, \dots, n)$, are the eigenvalues of M . MM should describe the thermal transport properties of the structure better by incorporating information about the atomic mass while inheriting the good features of CM. It is worth a mention that the expression of MM [Eq. (6)] was formulated simply as an extension of CM and was not derived based on physical principles, but descriptors do not necessarily need to be physical, and often unphysical descriptors can perform as well as the physical ones if not better.

C. Property calculator

The property calculator is used to calculate the objective function during optimization. It can be any open-source, commercial, or homemade tool for thermal transport property calculations. Examples include density functional theory,⁷⁷ molecular dynamics simulations,⁷⁸ Monte Carlo simulations,⁷⁹ finite element method, and electromagnetic simulations.^{80,81} When selecting the appropriate property calculator, the calculation time is a critical consideration because it needs to be balanced with the selection time (including modeling). This balance depends on the target system and the property as well as the general size and complexity of the system. For a given property calculation method, the calculation time increases as the system size, selection time, and the number of structure candidates increase. However, the selection time may be overcome by, for example, choosing the appropriate selection process and transforming it to a binary optimization problem (fractonization machine) to use quantum (or digital) annealers.⁶⁵ Although the property calculation can also be sped up by parallelization, the speed remains limited, particularly when using *ab initio* approaches. Therefore, the choice of calculator is critical when considering the objective of the research, and often compromise is necessary. For instance, in the example problem of superlattice optimization among 16 384 candidates, the atomistic Green's function (AGF) method is chosen because it can be performed within one minute. This time is comparable with the selection (modeling) time in Bayesian optimization, which is about a half a minute on a computer with an Intel Core i7-8700 CPU @ 3.20 GHz. The drawback of AGF is that it cannot capture anharmonic phonon-phonon scattering. Hence, this is used as a compromise for the sake of optimization speed because the influence of anharmonicity has been thought to be negligible in many cases since the interface distance is much smaller than the phonon mean-free-paths. It should be noted here that recent advances in the anharmonic AGF calculations⁸² have suggested notable contribution of anharmonicity to the interfacial conductance, and this aspect of compromise requires more attention in the future. Of course, the choice of calculator also depends on the amount of computer resources available and the acceptable wait time for the results. Heavier property calculators that include anharmonic effects such as molecular dynamics

and anharmonic lattice dynamics are feasible, but it is important to consider the balance in the problem design.

Here, we briefly describe the AGF method. For a more detailed description, refer to previous works.^{83,84} The phonon transmission function across the Si/Ge superlattice region $\Xi(\omega)$ is calculated by

$$\Xi(\omega) = \text{Trace}[\Gamma_L G^r \Gamma_R G^a], \quad (9)$$

where ω is the phonon frequency. G^r and G^a are the retarded and advanced Green's functions of the scattering region, respectively. The level broadening matrices $\Gamma_L = i(\Sigma_L^r - \Sigma_L^a)$ and $\Gamma_R = i(\Sigma_R^r - \Sigma_R^a)$ describe the rates of inflow from the left lead and outflow into the right lead, respectively. Σ_L and Σ_R are the self-energies, which are calculated from the surface Green's functions of the left and right leads, respectively. The surface Green's function calculated for semi-infinite regions reproduces the left and right Si leads with an infinite thickness (Fig. 2). Hence, the transmission can be calculated for the full phonon spectrum, including long-wavelength phonons.

With the two leads kept at different temperatures T_L and T_R , the heat current flowing through the device is given by the Landauer formula,⁸⁵

$$J = \frac{1}{2\pi} \int \hbar\omega [f_L(\omega, T_L) - f_R(\omega, T_R)] \Xi(\omega) d\omega, \quad (10)$$

where f_L and f_R are the Bose-Einstein distributions of phonons. In the limit of small temperature differences, the value of thermal conductance at average temperature T can be further obtained as

$$\sigma = \frac{\hbar^2}{2\pi k_B T^2 S} \int \omega^2 \Xi(\omega) \frac{e^{i\omega/k_B T}}{(e^{i\omega/k_B T} - 1)^2} d\omega, \quad (11)$$

where k_B is the Boltzmann constant and S is the cross-sectional area. As mentioned above, only the harmonic interaction is considered here. In principle, the anharmonic interaction can be included in the AGF calculation by introducing three-phonon scattering at interfaces via the many-body self-energy in the reciprocal space. However, the calculation is heavy.

To minimize the computational cost, the supercell size in the transverse direction perpendicular to heat conduction is set as a single unit cell (UC) (i.e., $5.43 \times 5.43 \text{ \AA}^2$) since the periodic boundary condition can represent the interface with an infinite cross section, and the thermal conductance is independent of the size of the transverse supercell. The lattice mismatch between Si and Ge is ignored because the stress effect on thermal transport due to the lattice mismatch is negligible.⁸³ To further reduce the computational time, the k -mesh size is optimized in the AGF calculation. Figure 3 shows the dependences of the phonon transmission function and thermal conductance on k -mesh size for a pure silicon system with a cross section size of $1 \text{ UC} \times 1 \text{ UC}$, where convergence is confirmed for a mesh size larger than 10×10 . This confirms that the k -mesh used in our calculation (20×20) is sufficient to capture the phonons in the full frequency spectrum. All the calculations are carried out using the commercial Atomistix ToolKit simulation package (ATK)⁸⁶ with Tersoff potentials.^{87,88}

D. Combining the Bayesian optimization and thermal transport calculation

The final step is to choose a suitable machine-learning algorithm and combine it with the property calculator to form closed-loop optimization. Among the various machine-learning algorithms, the Bayesian optimization employing Gaussian processes is well established as a prediction model for black-box optimizations. The combined process for the Bayesian optimization and the thermal-transport-property calculator is shown in [Algorithm 1](#) and [Fig. 4](#). The sample script can be found in the [supplementary material](#). Suppose the candidates are represented as d -dimensional vectors $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^d$, which are candidate descriptors. The dimensional size is typically smaller than 20 in most successful applications. In the

Bayesian optimization process, the transport properties of n candidates are initially calculated. Then, the next one to calculate is chosen. The purpose of the Bayesian optimization is to find the candidate with an optimal transport property using as few calculations/experiments as possible. A Bayesian regression function is learned from n pairs of descriptors and transport properties. This statistical prediction model is invariably a Gaussian process, which provides a Bayesian posterior probability distribution that describes the potential thermal conductance values for candidate point x . The best candidates are chosen based on the expected improvement criterion,⁸⁹ a widely used acquisition function that can explicitly encode a trade-off between exploitation (evaluating points with low mean) and exploration (evaluating points with high uncertainty). Besides the expected improvement, the maximum probability of improvement and upper confidence bound are two other popular acquisition functions⁸⁹ used in Bayesian optimization. Finally, the exact thermal conductance values are calculated for the chosen candidate and added to the training set. By repeating this procedure, the calculation of thermal conductance is scheduled optimally, and the best candidate can be found quickly.

Algorithm 1. Basic processes of Bayesian optimization

```

Determine vector descriptors for the optimization problem.
Randomly select  $n_0$  initial candidates and evaluate the target property.
While the calculated number of candidates or the optimization target is
not reached
do
    Update the posterior probability distribution using all available training data.
    Compute the acquisition function using the current posterior distribution.
    Select next possible optimal candidate  $x_n$  according to the maximization
    of the acquisition function over  $\mathbf{x}$ .
    Evaluate the property for candidate  $x_n$ .
    Increment n.
end while
Return the solution.

```

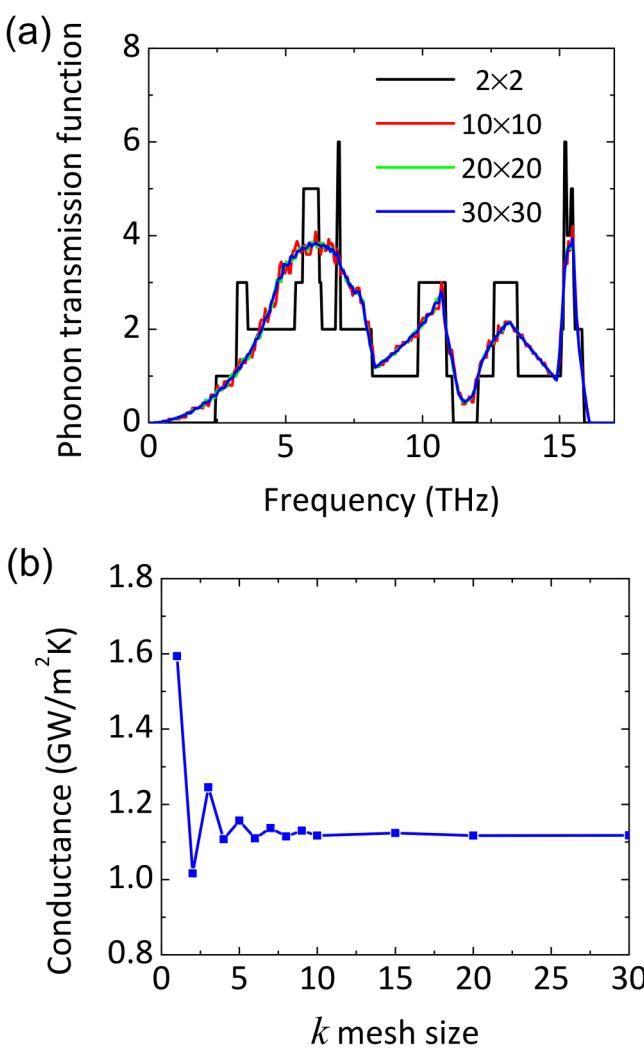


FIG. 3. Dependence of the (a) phonon transmission function and (b) thermal conductance on k -mesh size for a pure silicon system with a cross section size of 1 UC \times 1 UC.

Many Bayesian optimization tools have been developed. Examples include GPyOpt,⁹⁰ DiceOptim,⁹¹ Spearmint,⁹² GPflow,⁹³ Metrics Optimization Engine,⁹⁴ and the open-source Bayesian optimization library (COMBO).⁶¹ Here, we use COMBO⁶¹ to perform the Bayesian optimization process. COMBO is a black-box tool without user hyperparameter tuning for the kernel of Gaussian processes or an acquisition function. It has been successfully applied to optimizations in various energy transport problems. It is highly scalable and convenient for researchers from different research fields such as materials, physics, chemistry, and engineering due to the efficient protocol employing Thompson sampling, random feature maps, one-rank Cholesky update, and automatic hyperparameter tuning.

In COMBO, the prediction model is built by employing a Bayesian linear regression model combined with a random feature map Φ as

$$y = \mathbf{w}^T \varphi(\mathbf{x}) + \varepsilon, \quad (12)$$

where \mathbf{x} is a candidate descriptor, the values y are the thermal transport property evaluations for the candidates, \mathbf{w} is a D -dimensional weight vector, and ε is the noise subject to a

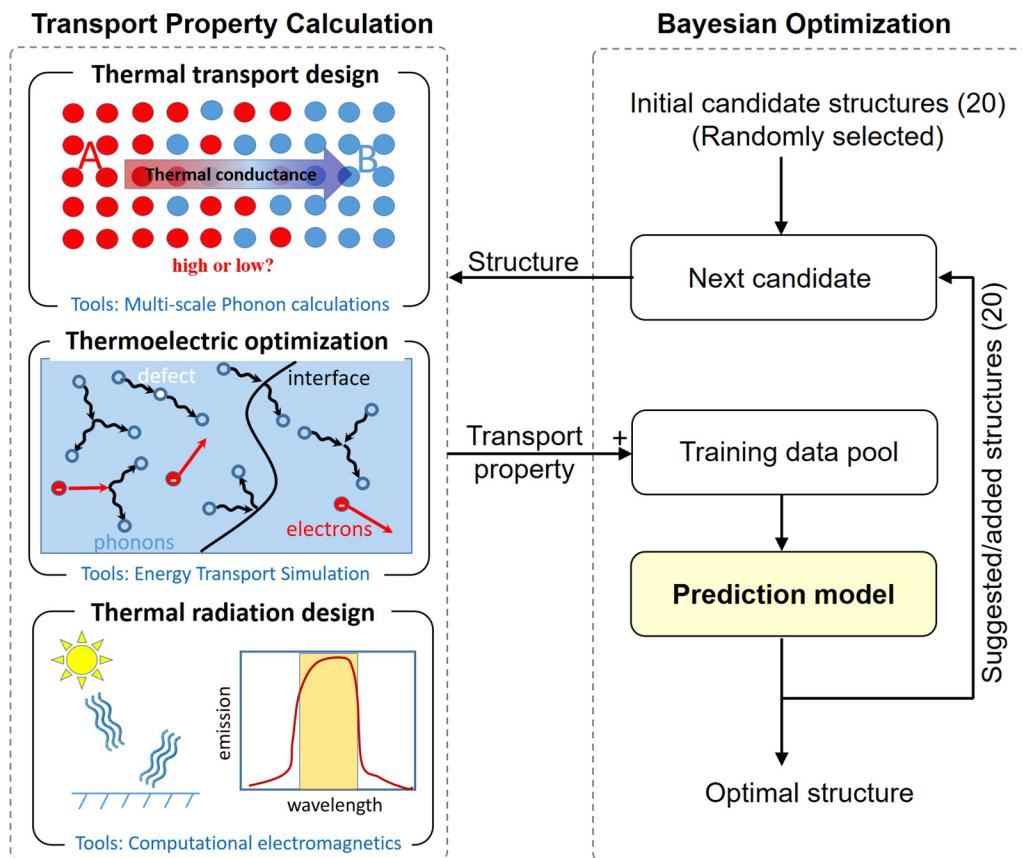


FIG. 4. General workflow of the optimal material design for thermal functions such as phonon transport, thermoelectrics, and thermal radiation.

normal distribution with mean 0 and variance σ . The selection of feature map Φ which approximates the distribution will affect the performance of Bayesian optimization crucially. Here, the random feature map means the feature map is built based on random samples. With the increasing number of random samples, the Bayesian linear model with the random feature map converges to a Gaussian process. In COMBO, the random feature map is chosen so that the inner product corresponds to the Gaussian kernel,⁹⁵

$$\varphi(\mathbf{x})^T \varphi(\mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{\eta^2}\right). \quad (13)$$

The model's performance depends on hyperparameters σ and η that are initialized according to the heuristic procedure by Yang *et al.*⁹⁶ In every round with 20 or so training examples added, the hyperparameters are periodically configured and updated by maximizing the type-II likelihood.⁹⁷ The [supplementary material](#) includes the full example script and supporting thermal conductance data for researchers to conveniently repeat our demonstration case for the superlattice structure design.

III. RESULTS AND DISCUSSION

A. Optimization efficiency of Bayesian optimization

Figure 5(a) shows the typical evolution of Bayesian optimization. As the number of calculated candidate structures increases, the thermal conductance value fluctuates, reflecting the stochastic feature in Bayesian optimization. The current minimum thermal conductance converges gradually in the calculations of 1000 structures, which is around 6.1% of the total number of candidates.

This value is somewhat larger (i.e., the optimization efficiency is lower) than the other cases that the authors and co-workers have previously reported such as interfacial Si/Ge alloy structure optimization.⁶⁰ One reason may be that the distribution of structures near the lowest thermal conductance is large in the histogram (Fig. 6), suggesting that there are many structures with similarly low thermal conductance values. Thus, more calculations are required to find the superlattice with the global minimum of thermal conductance. This differs from the case for alloy structure optimization, which has a smaller distribution near the lowest thermal conductance. These results suggest that optimization efficiency is not universal and depends on the details of the optimization

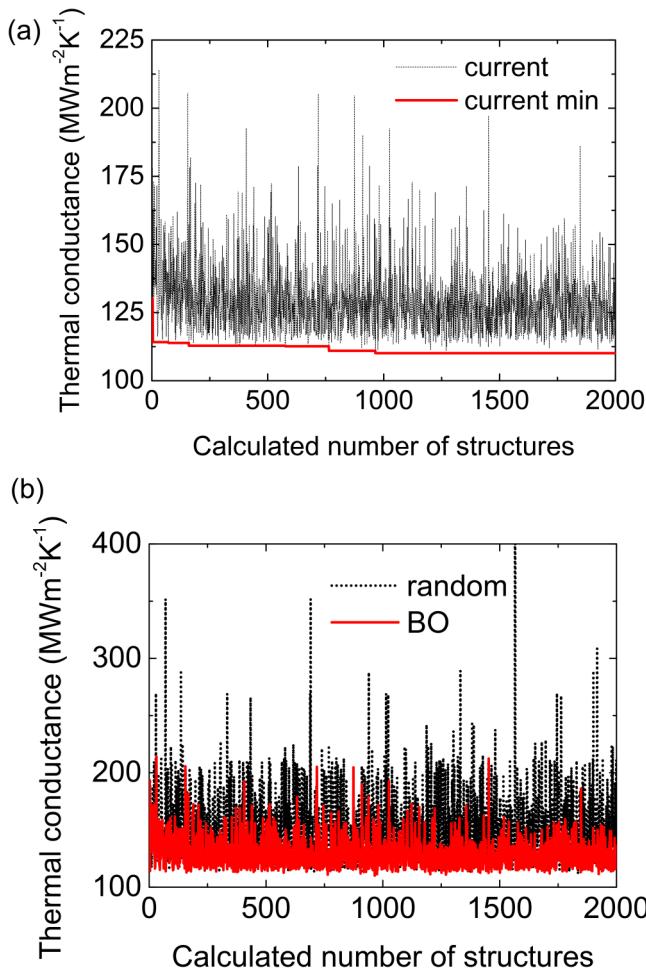


FIG. 5. (a) Evolution of the thermal conductance of current and the best-so-far superlattice structure as the number of calculated structures increases. (b) Comparison of optimization evolution between the Bayesian optimization and a random search.

problem. Regardless, machine learning performs significantly better than a random search to find structures with a lower thermal conductance [Fig. 5(b)].

To check optimization accuracy, the thermal conductances of all candidates are calculated. It is confirmed that the minimum thermal conductance and the corresponding superlattice structure are the same as those obtained by Bayesian optimization.

To evaluate how optimization efficiency depends on the algorithm, the Bayesian optimization is compared with a random search and a Monte Carlo tree search, which combines the generality of random simulation with a precision tree search in optimization. For each optimization algorithm, ten rounds of optimizations are conducted with different random seeds, which represent the different optimization experiments. Figure 7 shows the average convergence of each optimization method. The efficiency of Bayesian optimization is

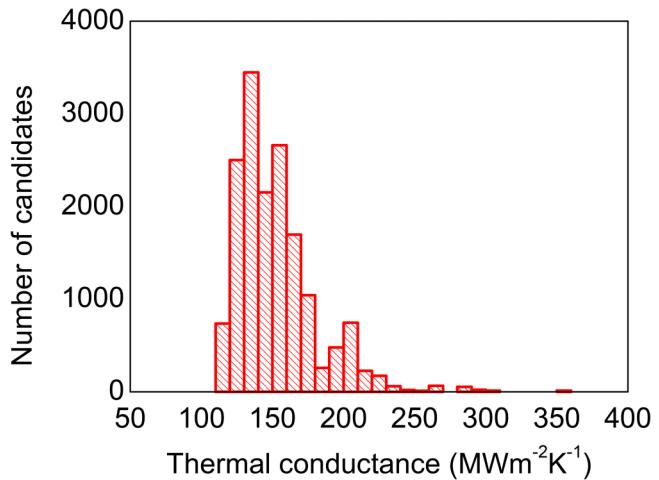


FIG. 6. Histogram of the thermal conductance of all candidate structures for the 14-UL-layer Si/Ge superlattices.

much higher than that of random search. As for the Monte Carlo tree search, the efficiency is better than the random search but is lower than that of Bayesian optimization.

It is noteworthy that the Bayesian optimization is extremely effective when the total number of candidates is around several hundred thousand. However, for cases with a much larger or even unlimited number of candidates, the required computer memory becomes massive because the algorithm simultaneously handles and scores the entire candidate structures through the built model. To overcome this limitation, two options are offered. One is to divide the candidate pool into subgroups and perform the Bayesian optimization for each subgroup. Section IV shows recent applications of this option to design a thermal radiation surface. The other option

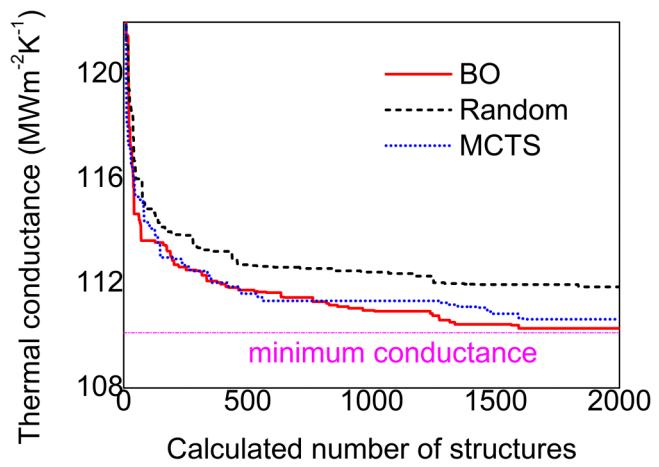


FIG. 7. Comparison of the efficiencies between Bayesian optimization (BO), Monte Carlo tree search (MCTS), and random search.

is to adopt the Monte Carlo tree search algorithm, which has superior scalability for targeting candidates that are approaching the global optimal ones when the total number of candidates is large or infinite. As shown in Fig. 7, the optimization efficiency is reduced.

B. Comparison of different descriptors

The optimization efficiencies for different descriptors described in Sec. II B are compared. Figure 8 shows the average number of calculated structures needed to find the superlattice with the minimum thermal conductance. To compare the efficiency of different descriptors, here we define the average efficiency by averaging the percentage of structures calculated until we find the optimal structure for each optimization experiment with random seed. The average efficiency is 6.59%, 4.25%, 3.10%, 6.31%, and 3.69% for descriptors using binary flag values, CM, EVMC, MM, and EVMM, respectively. The efficiencies for binary flag and MM descriptors are comparable, whereas CM improves the search efficiency. When reducing the dimension size of the descriptors using eigenvalues, the optimization efficiency further improves for both the CM and the MM.

The search efficiencies of Bayesian optimization and a random search are compared for different descriptors. For each type of descriptor, N_{Bayes} and N_{rand} are evaluated, which denote the average number of calculations needed until the first structure belonging to top k ($k = 1\text{--}20$) of all the candidates is found by Bayesian optimization and random search. Here, we compare the optimization efficiency using the ratio of N_{Bayes} to N_{rand} ($N_{\text{Bayes}}/N_{\text{rand}}$). For all descriptor types, the Bayesian optimization has a higher efficiency than the random search (Fig. 9). For $k \geq 6$, the search efficiencies of EVCM and EVMM are worse than that of CM, MM, and binary flag values. Binary flag values, CM, and MM descriptors are more efficient than the eigenvalue descriptors to find superlattice structures approaching the global minimum thermal conductance.

C. Number of candidates calculated and added in each round

In the Bayesian optimization, one adjustable parameter is the number of candidates added in each round. Figure 10 compares the

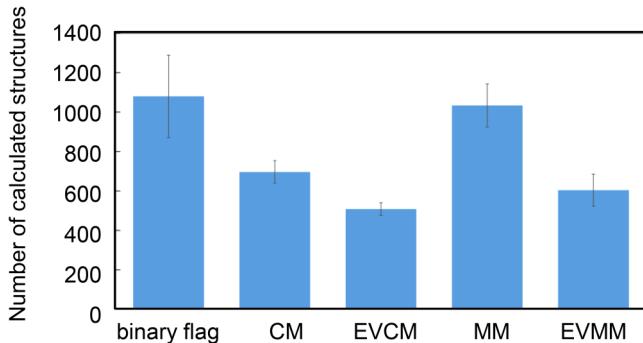


FIG. 8. Comparison of the efficiencies for different descriptors using binary flag values, Coulomb matrix (CM), eigenvalues of the Coulomb matrix (EVCM), mass matrix (MM), and eigenvalues of the mass matrix (EVMM).

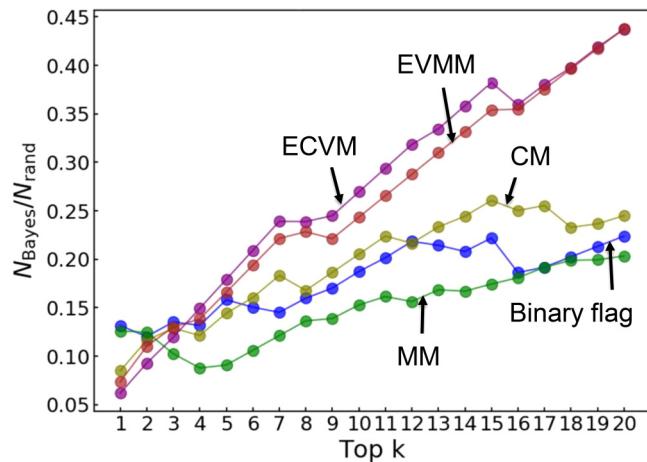


FIG. 9. Efficiency comparison between the Bayesian optimization and a random search for different descriptors.

consumed optimization time and the efficiency for different numbers of candidates added in each round. The number added should not be too small or too large. If it is too small, modeling and selection must be performed frequently, resulting in a longer optimization time. If it is too large, the total number of calculations required to obtain the optimal structure becomes large. By balancing these two factors, the optimal number is 50 in the current superlattice structure optimization. The optimal number depends on the characteristics of each problem. However, our experience suggests that the optimal number typically ranges between 20 and 50.

D. New knowledge gained via MI

MI not only finds the optimal structure with a high efficiency but also helps find the physical mechanisms behind designed novel structures. The aforementioned optimized aperiodic superlattice structure (10010101101111) is one such example. In fact, we have optimized superlattices with different layer thicknesses. All the results show that the aperiodic superlattice has a relatively lower thermal conductivity than its corresponding periodic one [Fig. 11(a)].

Knowing the optimal structures or classes, we can “reverse-study” the mechanism that gives rise to minimum thermal conductance. Using two separate calculations, one where the thickness is changed for a constant number of layers and the other where the number of layers is changed at a constant thickness, the optimal periodic superlattice is evaluated. Both reduce the thermal conductance but for different reasons. The change when altering the thickness is due to Fabry-Pérot wave interference, while that when altering the number of layers is due to interfacial scattering. When the total thickness of the superlattice is fixed, the thickness and the number of layers cannot be freely varied, imposing a trade-off to balance the two. Here, an aperiodic structure has more degrees of freedom to optimize the balance. The significance of the added degrees of freedom in an aperiodic superlattice can be seen in Fig. 11(b) as the spread of the data for a certain number of layers (i.e., the number of interfaces). Hence, the

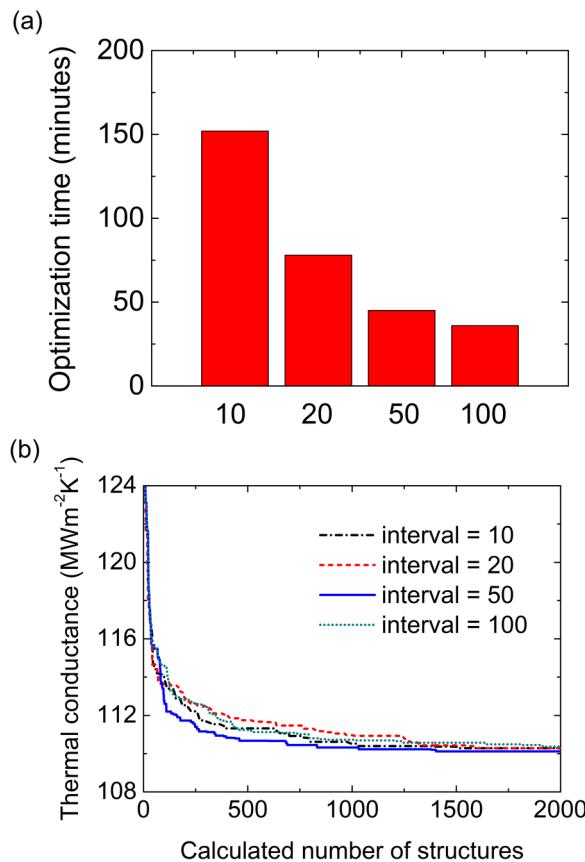


FIG. 10. Comparison of the (a) consumed optimization time and (b) convergence for different numbers of candidates calculated and added in each round.

minimum value is the one with the best balance between wave interference and interfacial scattering effects. Recent progress on the aperiodic GaAs/AlAs superlattice design via Bayesian optimization has been demonstrated experimentally in Ref. 98. The agreement between the fully coherent calculation and the experiments assures that the minimization of thermal conductivity is due to the optimal control of the coherent phonon thermal transport. The significant reduction in the optimal aperiodic superlattice compared with the optimal periodic superlattice is attributed to the stronger and more extensive localization of the acoustic phonons.

IV. RECENT APPLICATIONS OF NANOSTRUCTURE OPTIMIZATION BY MI

Here, we briefly introduced recent MI applications on thermoelectrics⁴² and thermal radiation^{44,45} performed using Bayesian optimization.

A. Thermoelectric optimization

The Bayesian optimization is also useful to optimize the figure-of-merit, which is comprised of multiple properties in trade-off relations such as thermal and electric transport in thermoelectric materials. Here, we introduce graphene nanoribbons (GNRs) structural optimization,⁴² which alternated the phonon/electron transport calculation and the Bayesian optimization to obtain a high thermoelectric figure of merit.

1. Problem description and objective function

Figures 12(a) and 12(b) show the two types of defective GNR structures: the periodically nanostructured GNR and the antidot GNR. For the periodic GNR, the optimization problem removes m atoms from the system to achieve the highest figure of merit. To ensure structural stability during optimization, it is constrained

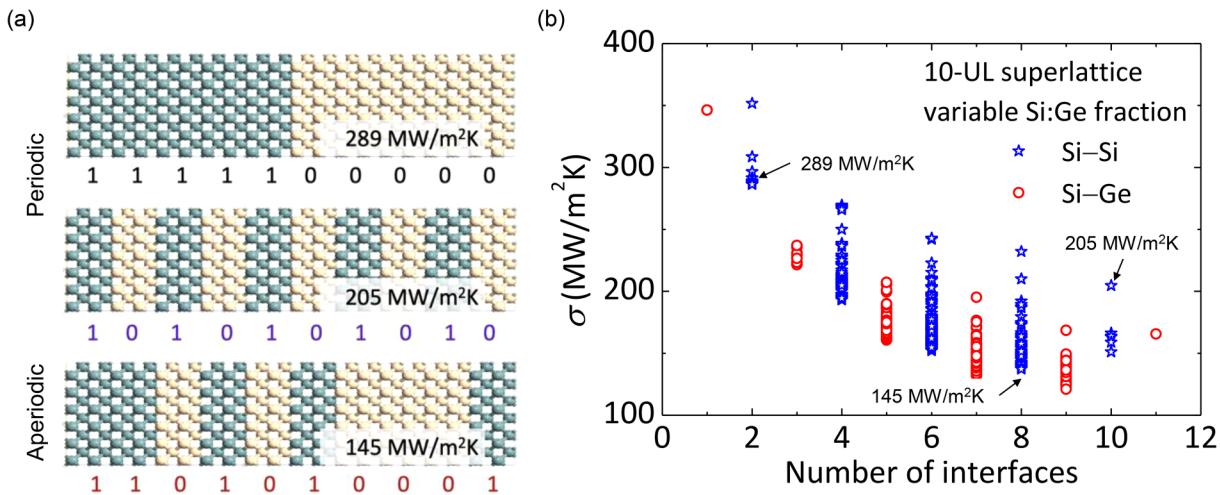


FIG. 11. (a) Comparison of the thermal conductance for periodic and aperiodic superlattices. (b) Thermal conductance vs the number of interfaces for various superlattices with the same total thickness.⁶⁰

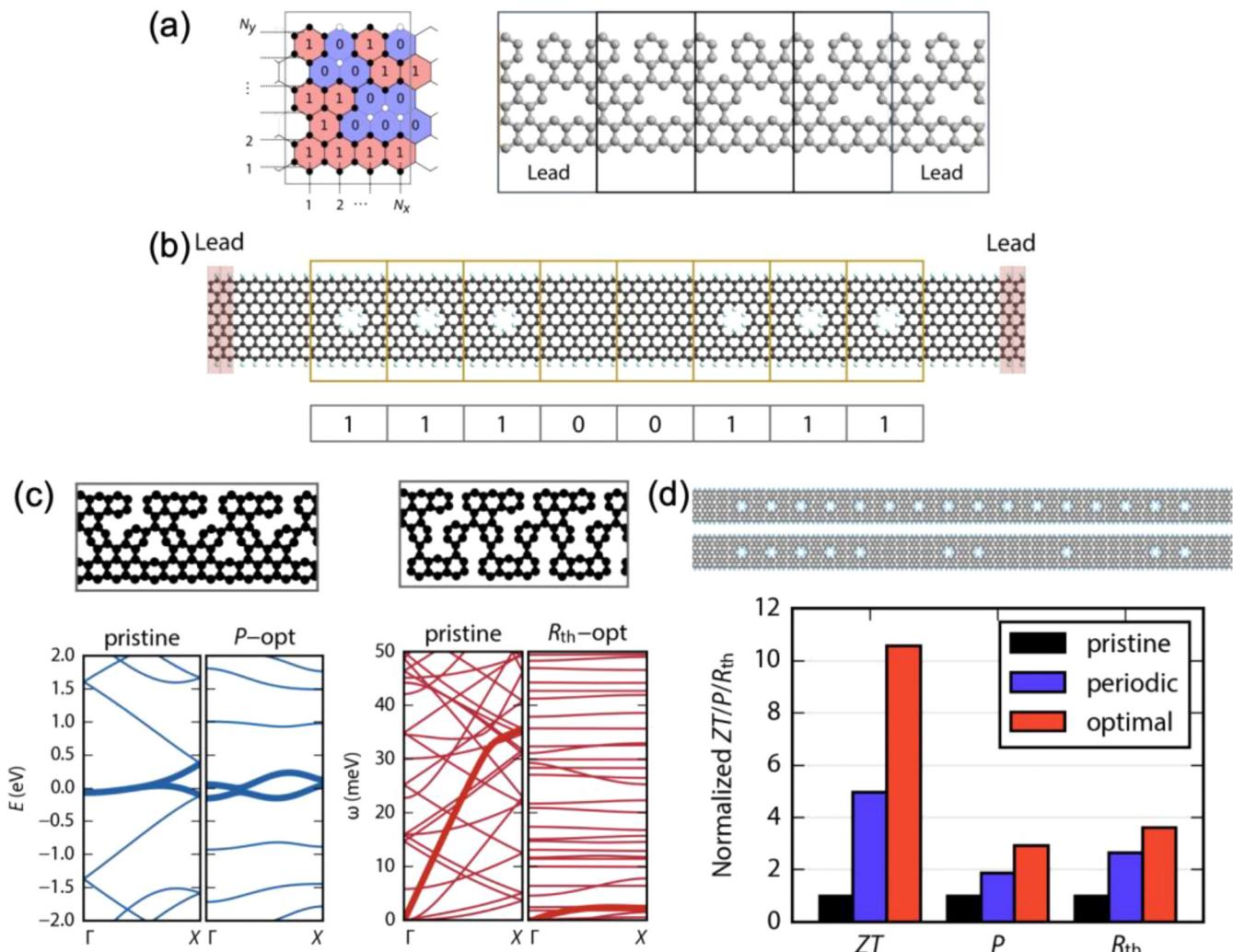


FIG. 12. Designing defective graphene nanoribbons for a thermoelectric application using Bayesian optimization. (a) Periodically nanostructured graphene nanoribbon. (b) Antidot graphene nanoribbon. (c) Optimal geometrical structures for a high power factor and low thermal resistance, corresponding to electron/phonon band structures with that of pristine graphene nanoribbon. (d) Comparison of the normalized figure of merit (ZT), power factor (P), and thermal resistance (R_{th}) for the pristine, periodic, and optimal structures.⁴²

such that every carbon atom in each candidate structure belongs to at least one complete hexagonal lattice, and no hexagonal lattice is isolated. As for the antidot GNR, it is composed of multiple sections, and each section is either a pristine structure or an antidot structure with a pore possessing a diameter of $d=2.9$ Å in the center. The optimization problem is whether to dig a hole in each section to achieve the highest figure of merit.

2. Descriptors and transport property evaluation

Similar to the phonon transport optimization, binary flag values are employed as descriptors, where 0 and 1, respectively, represent the defective and complete hexagonal lattices in periodic GNR

optimization [Fig. 12(a)] and the pristine structure and the antidot structure in antidot GNR optimization [Fig. 12(b)]. In this way, the vector value $\mathbf{g}=(1, 0, \dots)$ can describe each candidate structure in optimization. Green's function with the Landauer formula is employed to evaluate the thermoelectric properties, which includes the phonon/electron thermal resistance, the Seebeck coefficient, and the power factor calculated using the Atomistix ToolKit package.⁸⁶

3. Results and discussions

The efficiencies of Bayesian optimization and a random search for thermoelectric optimization are first compared for periodic nanostructured GNR optimization. In most cases, a random search

requires double the number of candidate structure calculations compared to the Bayesian optimization to find top 0.5% of all candidates with the highest figure of merit. The absence of m -dependence indicates that the optimization efficiency is independent of the total number of candidates. Hence, a comparable efficiency is expected in cases for larger GNR systems.

Figure 12(c) shows GNR structures with $m = 4$ and 10 with the largest power factor and lowest thermal resistance, respectively. For the highest power factor structure, vacancies are introduced over the entire area of GNR, except for the hexagonal lattices along the edge. The structural optimization leads to a strong flattening of the electronic bands around the energy levels of the edge state, and eventually, bandgaps are generated. Nanostructuring in the middle areas of zigzag GNRs leads to phonon scattering without a significant change in the edge state and enhances the thermoelectric performance. For the structure with the lowest thermal resistance, GNR has a complicated labyrinthian shape, where the phonons are significantly localized with reduced group velocities.

For the antidot GNR optimization, 16 sections, which give a total of 32 896 candidates, are considered. The final optimal structure has an aperiodic array of antidots. Figure 12(d) compares the figure of merit, power factor, and thermal resistance of the pristine, periodic antidot, and optimal structures. The optimal arrangement of antidots increases the figure of merit 11 times with an enhanced power factor and thermal resistance of 2.9 and 3.6 times, respectively. It is interesting to note that simply arranging the antidot periodically increases the figure of merit 5.0 times compared with the pristine structure. However, the remaining 2.1-time enhancement requires optimization. These results show that the optimization of the arrangement of antidots can simultaneously improve thermal and electronic properties. The successful thermoelectric optimization for defective GNR structures demonstrates the versatility of the Bayesian optimization in a robust and highly efficient multi-variable optimal search.

B. Thermal radiation

Tailoring thermal radiation plays an important role in designing ultranarrow-band, polarization-dependent, directional emission nanophotonics, and metamaterials. Such materials have a wide range of applications in solar emitters/absorbers, thermal functional textiles, light incandescent sources, and radiative cooling. The Bayesian optimization has been successfully applied to an ultranarrow-band wavelength-selective thermal radiator and thermal radiative cooling.⁴⁴ Here, we introduce the work by Guo *et al.*⁴⁴ They employed the Bayesian optimization to design structures consisting of a one-dimensional grating and multilayer structures for a spectral-selective radiative cooling application.

1. Problem description and objective function

Thermal radiative cooling takes advantage of the atmospheric transparent window in the wavelength range of 8–13 μm . This window directly transmits the thermal radiation energy to 3 K outer space without any energy consumption. A high radiative emissivity is essential to maximize the heat flux radiated from this atmospheric transparent window. The structure to be optimized is composed of a top dielectric grating layer, a middle multilayer, and

a thin silver layer at the bottom [Fig. 13(a)]. The grating material is silica because it can excite the surface phonon polaritons around 10 μm to enhance the absorption in the atmospheric window. Candidate materials for the multilayer are selected based on the criteria that the absorption should be high in the atmospheric window while the solar energy wavelength range should be as low as possible. Silver is selected as the substrate material due to its high plasmon frequency to reflect the solar energy.

To evaluate the thermal radiative property within the atmospheric window for each candidate structure, the figure of merit (FOM), which compares the emission power with blackbody case, is defined as

$$\text{FOM} = \frac{\int_{\lambda_1}^{\lambda_2} \varepsilon E_b d\lambda}{\int_{\lambda_1}^{\lambda_2} E_b d\lambda} - \frac{\int_{\lambda_{\min}}^{\lambda_1} \varepsilon E_b d\lambda}{\int_{\lambda_{\min}}^{\lambda_1} E_b d\lambda} - \frac{\int_{\lambda_2}^{\lambda_{\max}} \varepsilon E_b d\lambda}{\int_{\lambda_2}^{\lambda_{\max}} E_b d\lambda}, \quad (14)$$

where ε is the emissivity of each structure and E_b is the spectral blackbody emissive. λ_1 , λ_2 , λ_{\min} , and λ_{\max} are 8, 13, 1, and 20 μm , respectively.

2. Descriptors and transport property evaluation

The parameters involved in optimization are the grating period length P ; filling ratio f ; height h ; the material type for three layers m_1 , m_2 , and m_3 ; and thickness d_1 , d_2 , and d_3 for three layers in the multilayer region. Each possible candidate is represented by descriptors [P , f , h , m_1 , m_2 , m_3 , d_1 , d_2 , and d_3]. Prior to optimization, different length scales and material types are roughly estimated to determine the variation range of these parameters as well as to fix increments to control the searching space. Here, the grating period length P ranges from 1 μm to 10 μm in 0.5- μm steps, the filling ratio f ranges from 0.1 to 1 in 0.05 steps, and the height h ranges from 0.5 μm to 10 μm in 1- μm steps. The material types are 0, 1, and 2 denoting Si, SiO_2 , and Al_2O_3 , respectively. The multilayer thicknesses d_1 , d_2 , and d_3 range from 1 μm to 10 μm in 0.5- μm steps. Hence, the total number of candidate structures is 510,183,360.

The property of the thermal photonic structure is evaluated by the rigorous coupled-wave analysis method. The structures are sliced into several layers along the z -direction. The electromagnetic field and dielectric distribution of the structure are decomposed by summation of the finite Fourier harmonic terms in each sliced layer. By imposing the continuous boundary condition between the adjacent sliced layers of the electric and magnetic fields, the governing Maxwell equation of the electric or the magnetic field can be solved fast and accurately.

3. Results and discussion

Due to memory-size constraints in Bayesian optimization, the candidates are divided into 320×81 subgroups. Then, the Bayesian optimization is performed for each group. Finally, the best structure is selected from each subgroup. Less than 50 candidates in each subgroup are used to identify the optimal structure in Bayesian optimization. By contrast, a random search requires 4500 structures to find the optimal candidate in each subgroup.

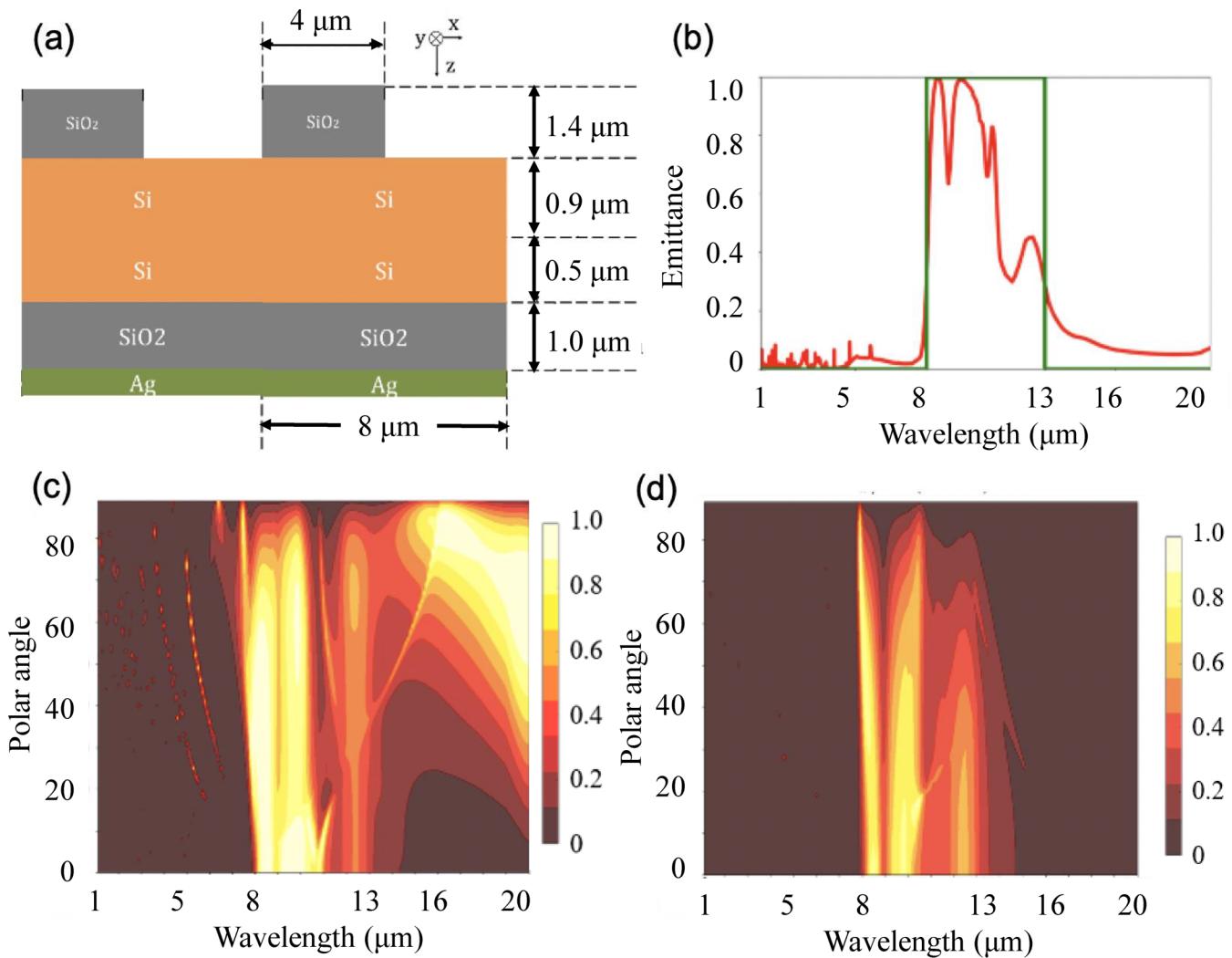


FIG. 13. Application of Bayesian optimization on thermal radiative cooling. (a) Optimized hybrid multilayer and grating structure. (b) Emittance of the optimized structure. (c) and (d) Polar angle dependence of the emittance for *p*-polarized and *s*-polarized waves, respectively.⁴⁴

The descriptor of the final targeted optimal structure is [8, 0.5, 1.4, 0, 0, 1, 0.9, 0.5, and 1] [Fig. 13(a)]. Its spectral selectivity exactly matches the atmospheric window [Fig. 13(b)]. It is interesting to see that, despite including Al₂O₃ in the optimization, the optimal structure does not contain any Al₂O₃. Figures 13(c) and 13(d) give the polar angle dependence of emittance for the designed optimal structure under *p*-polarized and *s*-polarized incident waves, respectively. For *s*-polarized wave incidence, the phonon polariton cannot be excited due to the absence of a material with negative permeability. The absorption is from the intrinsic absorption band of the materials. Comparing the *s*- and *p*-polarized incident results, the magnitude of *p*-polarized incident emittance increases greatly within 8–11 μm, indicating the necessity of a grating structure on the top surface to excite the phonon polaritons.

V. SUMMARY AND PERSPECTIVES

In the past few years, MI has evolved into a powerful and highly efficient tool for thermal functional materials design. In this Tutorial, we discuss the basic principles and implementation of MI involving descriptors, objective functions, property calculators, and informatics algorithms. The Bayesian optimization has been applied to find the superlattice structure with a minimized thermal conductance. The optimal aperiodic superlattice can be designed by calculating only a few percent of the total number of candidates. The low thermal conductivity of the aperiodic superlattice is attributed to the balance of the Fabry-Pérot wave interference and interfacial particle scattering. These results show that MI not only supports finding the optimal structure, but also promotes the discovery of

physical mechanisms. The thermoelectric optimization for defective GNR structures demonstrates the versatility, efficiency, and robustness of the Bayesian optimization in a multi-variable optimal search. For thermal radiation, a hybrid multilayer and grating structure realizes radiative cooling tailored to the thermal emittance within the atmospheric window. This material has the highest cooling performance to date. These successful applications demonstrate the advantage of MI in designing thermal functional materials with a desired property.

Despite the successful applications, MI for thermal transport is still in its infancy, which offers great opportunities for future development.

- (1) Most current MI work focuses on computational thermal transport due to ease and speed of investigation. In the future, experimental MI research should be expanded. In materials science, high-throughput materials synthesis and composition/structure characterization on combinatorial chips have been already relatively well developed. Combining MI with thermal property mapping measurements on a micrometer-scale resolution using methods such as time-domain thermoreflectance⁹⁹ may accelerate this process.
- (2) Although various elemental, structural, digital, and physical property descriptors have been developed for MI, novel and effective descriptors depend on the application. How to optimize the selection and combination of different types of descriptors is also important. One challenge is that expert experience or knowledge about descriptor selection is unknown prior to machine learning. One way to overcome this is to use machine learning to select the descriptors.¹⁰⁰
- (3) When using MI, there is a large gap between the “big data” required for credible machine learning and the “small data” currently available. The technique of transfer learning⁵⁶ has considerable potential to address this issue. For the target property to be predicted from limited training data, an intermediate property, which holds physical connections with the target property, is pretrained using available big data. The final prediction model can then be obtained by adjusting minor parameters in the pretrained model using small data.
- (4) As the design degrees of freedom increase, optimizations tend to exceed the ability of classical algorithms since optimization problems grow exponentially more difficult with the number of variables. Although the Monte Carlo tree search and genetic algorithms can deal with such optimization problems, their efficiency is not high. Recently, quantum annealing⁶⁵ has been incorporated into automated materials’ discovery to circumvent the difficulty with global optimization. Its application to design complex metamaterials for thermal radiative cooling using quantum annealing has shown its amazing efficiency and searching ability.

The concept of MI should be widely applied in more advanced materials’ design related to phonons, electrons, photons, and magnons.

SUPPLEMENTARY MATERIAL

Information concerning installation of COMBO, a sample script for the Si/Ge superlattice optimization, and data file description are provided.

ACKNOWLEDGMENTS

We thank Koji Tsuda, Thaer M Dieb, Atsushi Sakurai, Jiang Guo, and Masaki Yamawaki for their significant contributions to some of the works reviewed here. This work was supported by the Materials Research by Information Integration Initiative (MI²I) project with the Support Program for Starting Up Innovation Hub and CREST (Grant Nos. JPMJCR16Q5 and JPMJCR20Q3) from the Japan Science and Technology Agency (JST), and KAKENHI (Grant No. 19K14902) from the Japan Society for the Promotion of Science (JSPS). This work was also partially supported by the National Natural Science Foundation of China (NNSFC; Grant No. 52006134), the Materials Genome Initiative Center of Shanghai Jiao Tong University, and the Materials Genetic Engineering Project for Rare and Precious Metals by Yunnan Province (Grant No. 202002AB080001-1).

DATA AVAILABILITY

The data and unavailable thesis reference that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹S. M. Lee, D. G. Cahill, and R. Venkatasubramanian, “Thermal conductivity of Si–Ge superlattices,” *Appl. Phys. Lett.* **70**(22), 2957–2959 (1997).
- ²M. N. Luckyanova, J. Garg, K. Esfarjani, A. Jandl, M. T. Balsara, A. J. Schmidt, A. J. Minnich, S. Chen, M. S. Dresselhaus, Z. Ren, E. A. Fitzgerald, and G. Chen, “Coherent phonon heat conduction in superlattices,” *Science* **338**(6109), 936–939 (2012).
- ³Y. Liao, T. Shiga, M. Kashiwagi, and J. Shiomi, “Akhiezer mechanism limits coherent heat conduction in phononic crystals,” *Phys. Rev. B* **98**(13), 134307 (2018).
- ⁴M. Nomura, “Control of phonon transport by phononic crystals and application to thermoelectric materials,” *Mater. Trans.* **57**, 1022 (2016).
- ⁵P. E. Hopkins, C. M. Reinke, M. F. Su, R. H. Olsson, E. A. Shaner, Z. C. Leseman, J. R. Serrano, L. M. Phinney, and I. El-Kady, “Reduction in the thermal conductivity of single crystalline silicon by phononic crystal patterning,” *Nano Lett.* **11**(1), 107–112 (2011).
- ⁶T. Oyake, L. Feng, T. Shiga, M. Isogawa, Y. Nakamura, and J. Shiomi, “Ultimate confinement of phonon propagation in silicon nanocrystalline structure,” *Phys. Rev. Lett.* **120**(4), 045901 (2018).
- ⁷J. Shiomi, “Research update: Phonon engineering of nanocrystalline silicon thermoelectrics,” *APL Mater.* **4**(10), 104504 (2016).
- ⁸S. H. Ju and X. G. Liang, “Thermal conductivity of nanocrystalline silicon by direct molecular dynamics simulation,” *J. Appl. Phys.* **112**(6), 064305 (2012).
- ⁹A. Miura, S. Zhou, T. Nozaki, and J. Shiomi, “Crystalline-Amorphous silicon nanocomposites with reduced thermal conductivity for bulk thermoelectrics,” *ACS Appl. Mater. Interfaces* **7**(24), 13484–13489 (2015).
- ¹⁰X. Xu, J. Chen, J. Zhou, and B. Li, “Thermal conductivity of polymers and their nanocomposites,” *Adv. Mater.* **30**(17), 1705544 (2018).
- ¹¹G. Xie, D. Ding, and G. Zhang, “Phonon coherence and its effect on thermal conductivity of nanostructures,” *Adv. Phys. X* **3**(1), 1480417 (2018).
- ¹²M. Maldovan, “Sound and heat revolutions in phononics,” *Nature* **503**(7475), 209–217 (2013).
- ¹³J. Maire, R. Anufriev, R. Yanagisawa, A. Ramiere, S. Volz, and M. Nomura, “Heat conduction tuning by wave nature of phonons,” *Sci. Adv.* **3**(8), e1700027 (2017).
- ¹⁴S. Ju and X. Liang, “Detecting the phonon interference effect in Si/Ge nanocomposite by wave packets,” *Appl. Phys. Lett.* **106**(20), 203107 (2015).

- ¹⁵M. Maldovan, "Phonon wave interference and thermal bandgap materials," *Nat. Mater.* **14**(7), 667–674 (2015).
- ¹⁶L. Feng, T. Shiga, H. Han, S. Ju, Y. A. Kosevich, and J. Shiomi, "Phonon-interference resonance effects by nanoparticles embedded in a matrix," *Phys. Rev. B* **96**(22), 220301(R) (2017).
- ¹⁷S. Hu, Z. Zhang, P. Jiang, J. Chen, S. Volz, M. Nomura, and B. Li, "Randomness-induced phonon localization in graphene heat conduction," *J. Phys. Chem. Lett.* **9**(14), 3959–3968 (2018).
- ¹⁸Z. Lazcano, O. Meza, and J. Arriaga, "Localization of acoustic modes in periodic porous silicon structures," *Nanoscale Res. Lett.* **9**(1), 419 (2014).
- ¹⁹S. Merabia and K. Termentzidis, "Thermal boundary conductance across rough interfaces probed by molecular dynamics," *Phys. Rev. B* **89**(5), 054309 (2014).
- ²⁰Z. Tian, K. Esfarjani, and G. Chen, "Enhancing phonon transmission across a Si/Ge interface by atomic roughness: First-principles study with the green's function method," *Phys. Rev. B* **86**(23), 235304 (2012).
- ²¹Y. Liu, C. Hu, J. Huang, B. G. Sumpter, and R. Qiao, "Tuning interfacial thermal conductance of graphene embedded in soft materials by vacancy defects," *J. Chem. Phys.* **142**(24), 244703 (2015).
- ²²P. K. Schelling, S. R. Phillpot, and P. Kebinski, "Kapitza conductance and phonon scattering at grain boundaries by simulation," *J. Appl. Phys.* **95**(11), 6082–6091 (2004).
- ²³S.-H. Ju and X.-G. Liang, "Investigation on interfacial thermal resistance and phonon scattering at twist boundary of silicon," *J. Appl. Phys.* **113**(5), 053513 (2013).
- ²⁴M. Sakata, T. Hori, T. Oyake, J. Maire, M. Nomura, and J. Shiomi, "Tuning thermal conductance across sintered silicon interface by local nanostructures," *Nano Energy* **13**, 601–608 (2015).
- ²⁵M. Sakata, T. Oyake, J. Maire, M. Nomura, E. Higurashi, and J. Shiomi, "Thermal conductance of silicon interfaces directly bonded by room-temperature surface activation," *Appl. Phys. Lett.* **106**(8), 081603 (2015).
- ²⁶K. Zheng, F. Sun, X. Tian, J. Zhu, Y. Ma, D. Tang, and F. Wang, "Tuning the interfacial thermal conductance between polystyrene and sapphire by controlling the interfacial adhesion," *ACS Appl. Mater. Interfaces* **7**(42), 23644–23649 (2015).
- ²⁷T. Hori and J. Shiomi, "Tuning phonon transport spectrum for better thermoelectric materials," *Sci. Technol. Adv. Mater.* **20**(1), 10–25 (2019).
- ²⁸K. Cui, P. Lemaire, H. Zhao, T. Savas, G. Parsons, and A. John Hart, "Tungsten-Carbon nanotube composite photonic crystals as thermally stable spectral-selective absorbers and emitters for thermophotovoltaics," *Adv. Energy Mater.* **8**(27), 1801471 (2018).
- ²⁹A. Lenert, Y. Nam, D. M. Bierman, and E. N. Wang, "Role of spectral non-idealities in the design of solar thermophotovoltaics," *Opt. Express* **22**(Suppl 6), A1604–18 (2014).
- ³⁰K. Krishnamurthy, H. K. Khurana, J. Soojin, J. Irudayaraj, and A. Demirci, "Infrared heating in food processing: An overview," *Compr. Rev. Food Sci. Food Saf.* **7**(1), 2–13 (2008).
- ³¹B. Bhatia, A. Leroy, Y. Shen, L. Zhao, M. Gianello, D. Li, T. Gu, J. Hu, M. Soljačić, and E. N. Wang, "Passive directional sub-ambient daytime radiative cooling," *Nat. Commun.* **9**(1), 5001 (2018).
- ³²Y. Zhai, Y. Ma, S. N. David, D. Zhao, R. Lou, G. Tan, R. Yang, and X. Yin, "Scalable-manufactured randomized glass-polymer hybrid metamaterial for daytime radiative cooling," *Science* **355**(6329), 1062–1066 (2017).
- ³³A. P. Raman, M. A. Anoma, L. Zhu, E. Rephaeli, and S. Fan, "Passive radiative cooling below ambient air temperature under direct sunlight," *Nature* **515**(7528), 540–544 (2014).
- ³⁴A. Agrawal and A. Choudhary, "Perspective materials informatics and big data realization of the "fourth paradigm" of science in materials science," *APL Mater.* **4**(5), 053208 (2016).
- ³⁵K. Rajan, "Materials informatics: The materials "gene" and big data," *Annu. Rev. Mater. Res.* **45**(1), 153–169 (2015).
- ³⁶S. Ju and J. Shiomi, "Materials informatics for heat transfer: Recent progresses and perspectives," *Nanoscale Microscale Thermophys. Eng.* **23**(2), 157–172 (2019).
- ³⁷P. B. Wigley, P. J. Everitt, A. van den Hengel, J. W. Bastian, M. A. Sooriyanbandara, G. D. McDonald, K. S. Hardman, C. D. Quinlivan, P. Manju, C. C. Kuhn, I. R. Petersen, A. N. Luiten, J. J. Hope, N. P. Robins, and M. R. Hush, "Fast machine-learning online optimization of ultra-cold-atom experiments," *Sci. Rep.* **6**, 25890 (2016).
- ³⁸D. S. Wishart, "Bioinformatics in drug development and assessment," *Drug Metab. Rev.* **37**(2), 279–310 (2005).
- ³⁹T. L. Blundell, B. L. Sibanda, R. W. Montalvão, S. Brewerton, V. Chelliah, C. L. Worth, N. J. Harmer, O. Davies, and D. Burke, "Structural biology and bioinformatics in drug design: Opportunities and challenges for target identification and lead discovery," *Philos. Trans. R. Soc. B* **361**(1467), 413–423 (2006).
- ⁴⁰N. Adams and P. Murray-Rust, "Engineering polymer informatics: Towards the computer-aided design of polymers," *Macromol. Rapid Commun.* **29**(8), 615–632 (2008).
- ⁴¹S. Kiyohara, H. Oda, K. Tsuda, and T. Mizoguchi, "Acceleration of stable interface structure searching using a kriging approach," *Jpn. J. Appl. Phys.* **55**(4), 045502 (2016).
- ⁴²M. Yamawaki, M. Ohnishi, S. Ju, and J. Shiomi, "Multifunctional structural design of graphene thermoelectrics by Bayesian optimization," *Sci. Adv.* **4**(6), eaar4192 (2018).
- ⁴³S. Ju, Y. Miura, K. Yamamoto, K. Masuda, K.-i. Uchida, and J. Shiomi, "Machine learning analysis of tunnel magnetoresistance of magnetic tunnel junctions with disordered MgAl₂O₄," *Phys. Rev. Res.* **2**(2), 023187 (2020).
- ⁴⁴J. Guo, S. Ju, and J. Shiomi, "Design of a highly selective radiative cooling structure accelerated by materials informatics," *Opt. Lett.* **45**(2), 343 (2020).
- ⁴⁵A. Sakurai, K. Yada, T. Simomura, S. Ju, M. Kashiwagi, H. Okada, T. Nagao, K. Tsuda, and J. Shiomi, "Ultranarrow-band wavelength-selective thermal emission with aperiodic multilayered metamaterials designed by Bayesian optimization," *ACS Cent. Sci.* **5**(2), 319–326 (2019).
- ⁴⁶A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The materials project A materials genome approach to accelerating materials innovation," *APL Mater.* **1**(1), 011002 (2013).
- ⁴⁷S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. W. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, "AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations," *Comput. Mater. Sci.* **58**, 227–235 (2012).
- ⁴⁸S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, "The open quantum materials database (OQMD): Assessing the accuracy of DFT formation energies," *npj Comput. Mater.* **1**(1), 15010 (2015).
- ⁴⁹J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, "Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD)," *JOM* **65**(11), 1501–1509 (2013).
- ⁵⁰See <https://nomad-coe.eu/> for NOMAD.
- ⁵¹A. van Roeghem, J. Carrete, C. Osés, S. Curtarolo, and N. Mingo, "High-throughput computation of thermal conductivity of high-temperature solid phases: The case of oxide and fluoride perovskites," *Phys. Rev. X* **6**(4), 041061 (2016).
- ⁵²K. Ghosh, A. Stuke, M. Todorović, P. B. Jørgensen, M. N. Schmidt, A. Vehtari, and P. Rinke, "Deep learning spectroscopy: Neural networks for molecular excitation spectra," *Adv. Sci.* **6**, 1801367 (2019).
- ⁵³D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics* **5**(4), 1365–1369 (2018).
- ⁵⁴A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, "Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization," *Phys. Rev. Lett.* **115**(20), 205901 (2015).
- ⁵⁵J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, "Finding unprecedentedly low-thermal-conductivity half-heusler semiconductors via high-throughput materials modeling," *Phys. Rev. X* **4**(1), 011019 (2014).

- ⁵⁶H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, and R. Yoshida, "Predicting materials properties with little data using shotgun transfer learning," *ACS Cent. Sci.* **5**(10), 1717–1730 (2019).
- ⁵⁷L. Gosselin, M. Tye-Gingras, and F. Mathieu-Potvin, "Review of utilization of genetic algorithms in heat transfer problems," *Int. J. Heat Mass Transfer* **52**(9–10), 2169–2188 (2009).
- ⁵⁸H. Wei, H. Bao, and X. Ruan, "Genetic algorithm-driven discovery of unexpected thermal conductivity enhancement by disorder," *Nano Energy* **71**, 104619 (2020).
- ⁵⁹E. Briones, R. Ruiz-Cruz, J. Briones, N. Gonzalez, J. Simon, M. Arreola, and G. Alvarez-Alvarez, "Particle swarm optimization of nanoantenna-based infrared detectors," *Opt Express* **26**(22), 28484–28496 (2018).
- ⁶⁰S. Ju, T. Shiga, L. Feng, Z. Hou, K. Tsuda, and J. Shiomi, "Designing nanostructures for phonon transport via Bayesian optimization," *Phys. Rev. X* **7**(2), 021024 (2017).
- ⁶¹T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, and K. Tsuda, "COMBO: An efficient Bayesian optimization library for materials science," *Mater. Discovery* **4**, 18–21 (2016).
- ⁶²J. Guo, S. Ju, and J. Shiomi, "Application of Bayesian optimization to thermal science," in *Nanoscale Energy Transport* (IOP Publishing, 2020), pp. 5–1–5–17.
- ⁶³T. M. Dieb, S. Ju, K. Yoshizoe, Z. Hou, J. Shiomi, and K. Tsuda, "MDTS: Automatic complex materials design using Monte Carlo tree search," *Sci. Technol. Adv. Mater.* **18**(1), 498–503 (2017).
- ⁶⁴X. Yang, J. Zhang, K. Yoshizoe, K. Terayama, and K. Tsuda, "ChemTS: An efficient python library for de novo molecular generation," *Sci. Technol. Adv. Mater.* **18**(1), 972–976 (2017).
- ⁶⁵K. Kitai, J. Guo, S. Ju, S. Tanaka, K. Tsuda, J. Shiomi, and R. Tamura, "Designing metamaterials with quantum annealing and factorization machines," *Phys. Rev. Res.* **2**(1), 013319 (2020).
- ⁶⁶P. K. Schelling and S. R. Phillpot, "Multiscale simulation of phonon transport in superlattices," *J. Appl. Phys.* **93**(9), 5377–5387 (2003).
- ⁶⁷B. Yang and G. Chen, "Partially coherent phonon heat conduction in superlattices," *Phys. Rev. B* **67**(19), 195311 (2003).
- ⁶⁸A. R. Abramson, C. L. Tien, and A. Majumdar, "Interface and strain effects on the thermal conductivity of heterostructures: A molecular dynamics study," *J. Heat Transfer-Trans.* **124**(5), 963–970 (2002).
- ⁶⁹C. Colvard, R. Merlin, M. V. Klein, and A. C. Gossard, "Observation of folded acoustic phonons in a semiconductor superlattice," *Phys. Rev. Lett.* **45**, 298 (1980).
- ⁷⁰J. Garg and G. Chen, "Minimum thermal conductivity in superlattices: A first-principles formalism," *Phys. Rev. B* **87**(14), 140302(R) (2013).
- ⁷¹B. Qiu, G. Chen, and Z. Tian, "Effects of aperiodicity and roughness on coherent heat conduction in superlattices," *Nanoscale Microscale Thermophys. Eng.* **19**(4), 272–278 (2015).
- ⁷²S. Wang, Z. Wang, W. Setyawan, N. Mingo, and S. Curtarolo, "Assessing the thermoelectric properties of sintered compounds via high-throughput ab-initio calculations," *Phys. Rev. X* **1**(2), 021012 (2011).
- ⁷³A. Seko, A. Togo, and I. Tanaka, "Descriptors for machine learning of materials data," in *Nanoinformatics*, edited by I. Tanaka (Springer Singapore, Singapore, 2018), pp. 3–23.
- ⁷⁴S. Shimizu, "Development of descriptors for nanostructure optimization in thermal functional materials," in *Department of Mechanical Engineering* (The University of Tokyo, 2019).
- ⁷⁵M. Rupp, A. Tkatchenko, K.-R. Müller, and O. Anatole von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning," *Phys. Rev. Lett.* **108**(5), 058301 (2012).
- ⁷⁶K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. von Lilienfeld, A. Tkatchenko, and K. R. Müller, "Assessment and validation of machine learning methods for predicting molecular atomization energies," *J. Chem. Theory Comput.* **9**(8), 3404–3419 (2013).
- ⁷⁷K. Esfarjani, G. Chen, and H. T. Stokes, "Heat transport in silicon from first-principles calculations," *Phys. Rev. B* **84**(8), 085204 (2011).
- ⁷⁸J. R. Lukes, D. Y. Li, X. G. Liang, and C. L. Tien, "Molecular dynamics study of solid thin-film thermal conductivity," *J. Heat Transfer-Trans.* **122**(3), 536–543 (2000).
- ⁷⁹T. Hori, J. Shiomi, and C. Dames, "Effective phonon mean free path in polycrystalline nanostructures," *Appl. Phys. Lett.* **106**(17), 171901 (2015).
- ⁸⁰X. Liu, T. Tyler, T. Starr, A. F. Starr, N. M. Jokerst, and W. J. Padilla, "Taming the blackbody with infrared metamaterials as selective thermal emitters," *Phys. Rev. Lett.* **107**(4), 045901 (2011).
- ⁸¹S. Basu and Z. M. Zhang, "Maximum energy transfer in near-field thermal radiation at nanometer distances," *J. Appl. Phys.* **105**(9), 093535 (2009).
- ⁸²J. Dai and Z. Tian, "Rigorous formalism of anharmonic atomistic green's function for three-dimensional interfaces," *Phys. Rev. B* **101**(4), 041301(R) (2020).
- ⁸³W. Zhang, T. S. Fisher, and N. Mingo, "Simulation of interfacial phonon transport in Si–Ge heterostructures using an atomistic green's function method," *J. Heat Transfer* **129**(4), 483 (2007).
- ⁸⁴J. S. Wang, J. Wang, and J. T. Lü, "Quantum thermal transport in nanostructures," *Eur. Phys. J. B* **62**(4), 381–404 (2008).
- ⁸⁵R. Landauer, "Electrical resistance of disordered one-dimensional lattices," *Philos. Mag.* **21**(172), 863–867 (1970).
- ⁸⁶Q. Wu, B. He, T. Song, J. Gao, and S. Shi, "Cluster expansion method and its application in computational materials science," *Comput. Mater. Sci.* **125**, 243–254 (2016).
- ⁸⁷J. Tersoff, "Modeling solid-state chemistry—Interatomic potentials for multicomponent systems," *Phys. Rev. B* **39**(8), 5566–5568 (1989).
- ⁸⁸J. Tersoff, "Erratum modeling solid-state chemistry interatomic potentials for multicomponent systems," *Phys. Rev. B* **41**(5), 3248–3248 (1990).
- ⁸⁹D. R. Jones, "A taxonomy of global optimization methods based on response surfaces," *J. Global Optim.* **21**(4), 345–383 (2001).
- ⁹⁰See <http://github.com/SheffieldML/GPyOpt> for GPyOpt.
- ⁹¹See <https://CRAN.R-project.org/package=DiceOptim> for DiceOptim.
- ⁹²See <https://github.com/HIPS/Spearmint> for Spearmint.
- ⁹³See <https://www.gpflow.org/> for GPFlow.
- ⁹⁴See <https://github.com/Yelp/MOE> for Metrics Optimization Engine.
- ⁹⁵A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *Advances in Neural Information Processing Systems 20* (NIPS, 2007).
- ⁹⁶Z. Yang, A. J. Smola, L. Song, and A. G. Wilson, "A la Carte—Learning fast kernels," in *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, JMLR: Workshop and Conference Proceedings, San Diego, CA (ML Research Press, 2015).
- ⁹⁷D. P. Kingma and J. L. Ba, "Adam A method for stochastic optimization," [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- ⁹⁸R. Hu, S. Iwamoto, L. Feng, S. Ju, S. Hu, M. Ohnishi, N. Nagai, K. Hirakawa, and J. Shiomi, "Machine-learning-optimized aperiodic superlattice minimizes coherent phonon heat conduction," *Phys. Rev. X* **10**(2), 021050 (2020).
- ⁹⁹S. Huxtable, D. G. Cahill, V. Fauconnier, J. O. White, and J. C. Zhao, "Thermal conductivity imaging at micrometre-scale resolution for combinatorial studies of materials," *Nat. Mater.* **3**(5), 298–301 (2004).
- ¹⁰⁰Y. Zhang, C. Wen, C. Wang, S. Antonov, D. Xue, Y. Bai, and Y. Su, "Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models," *Acta Mater.* **185**, 528–539 (2020).