Energy and AI 8 (2022) 100153

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.sciencedirect.com/journal/energy-and-ai

Perspective

Perspective: Predicting and optimizing thermal transport properties with machine learning methods

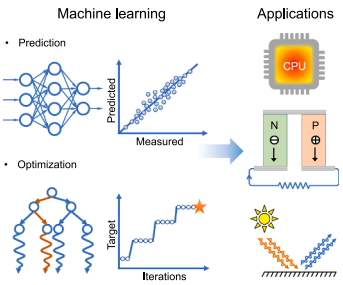
Han Wei [a], Hua Bao [a,*], Xiulin Ruan [b,*]

[a] University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China
[b] School of Mechanical Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907, United States

HIGHLIGHTS

- Machine learning prediction methods can serve as surrogate models and perform better than physics-based analytical models for predicting thermal transport properties.
- Machine learning optimization algorithms can serve as search tools for optimal structures of desired thermal transport properties, with superior effectiveness compared to random search or tuning based on intuition.
- Outstanding challenges and opportunities for future developments are identified, including developing machine learning methods suitable for small datasets, discovering effective physics-based descriptors, generating dataset from experiments and validating machine learning results with experiments, and making breakthroughs via discovering new physics.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Thermal transport properties
Machine learning
Prediction
Optimization

ABSTRACT

In recent years, (big) data science has emerged as the "fourth paradigm" in physical science research. Data-driven techniques, e.g. machine learning, are advantageous in dealing with problems of high-dimensional features and complex mappings between quantities, which are otherwise of great difficulty or huge cost with other scientific paradigms. In the past five years or so, there has been a rapid growth of machine learning-assisted research on thermal transport. In this perspective, we review the recent progress in the intersection between machine learning and thermal transport, where machine learning methods generally serve as surrogate models for predicting the thermal transport properties, or as tools for designing structures for the desired thermal properties and exploring thermal transport mechanisms. We provide perspectives about the advantages of machine learning methods in comparison to the physics-based methods for studying thermal transport properties. We also discuss how to improve the accuracy of predictive analytics and efficiency of structural optimization, to provide guidance for better utilizing machine learning-based methods to advance thermal transport research. Finally, we identify several outstanding challenges in this active area as well as opportunities for future developments,

* Corresponding authors.
E-mail addresses: hua.bao@sjtu.edu.cn (H. Bao), ruan@purdue.edu (X. Ruan).

https://doi.org/10.1016/j.egyai.2022.100153

Available online 12 March 2022

including developing machine learning methods suitable for small datasets, discovering effective physics-based descriptors, generating dataset from experiments and validating machine learning results with experiments, and making breakthroughs via discovering new physics.

1. Introduction

Thermal transport plays an important role in various applications, including thermal management in electronics, waste heat harvesting, and thermal insulation [1–4]. Understanding and tailoring the thermal transport properties is significant for these applications. However, it is a challenging task for two main reasons. First, the physical laws that govern thermal transport are quite distinct in different material types and at different scales [5,6]. Second, the detailed structures of material, from atomic to macroscopic scales, have a significant effect on the thermal transport properties, and therefore make the structure-property relationship quite complicated [7–9].

In the past few decades, significant advances have been made in this dynamic research field, mostly driven by physics-based methods [5,6,10,11]. For example, advanced simulation methods and computational tools have been developed, including first-principles density functional theory (DFT) calculation [4,11,12], molecular dynamics (MD) simulation [13–15], atomic Green's function (AGF) [16–19], phonon Boltzmann transport equation (BTE) [20–23], and the heat diffusion equation [24–26]. Various experimental thermal characterization methods have also been developed [3,27,28], such as laser flash method [29–31], electrical heating methods [32–34], 3ω methods [35–37], Raman microscopy [38–40], the transient thermoreflectance technique in time-domain [41,42] and frequency-domain [42,43] analysis. Recently, the (big) data science started to attract attention from the physical science area and emerged as the "fourth paradigm for scientific research" [44]. The data-driven scheme bypasses the physics and thus has no special preference for the specific physical problems and source of the data [44]. Therefore, it can possibly complement the first three paradigms, i.e., theory, experiment, and simulation [44]. Moreover, it is quite suitable to solve problems of high-dimensional features and complex mappings between quantities, which is of great difficulty or huge cost for other paradigms [44]. Owing to this advantage, data-driven techniques, e.g. machine learning methods, have been applied to materials science and expedite the understanding, selection, and design of materials [45–49]. For the thermal science community, with the developments of computational and experimental methods, a large amount of data has gradually accumulated and can also facilitate the application of data-driven methods to assist the study of thermal transport properties. In the past five years or so, there has been an outbreak of machine learning-assisted research in thermal science. For example, machine learning methods can serve as surrogate models for predicting the thermal transport properties [50–52], or as tools for designing structures for the desired thermal properties and exploring thermal transport mechanisms [53–56].

In this perspective, we draw attention to the recent progress on machine learning-assisted studies of thermal transport properties. We note that there are already some existing review papers in this area, focusing on different aspects of machine-learning applications in the thermal science area. For example, Ju et al. surveyed the studies on exploration and nanostructure design of materials with high/low thermal conductivity/conductance with machine learning-based methods [57]. Wan et al. covered the latest machine learning applications for properties prediction and materials discovery in thermal transport [58]. Zhang et al. reviewed the use of machine learning for screening of thermal conductivity and interfacial thermal conductance [59]. Herein, we provide our perspectives about the advantages of machine learning methods in comparison to the physics-based methods and how to improve the accuracy of prediction models and the efficiency of structural optimization, with the purpose of better utilizing machine learning-based methods to advance the thermal transport research. We only elaborate on some illustrative examples rather than covering the entire scope. The remainder of this paper is structured as follows. In Section 2, we begin by introducing the principles of machine learning for the prediction and optimization of thermal transport properties. In Section 3, we discuss the key applications of predictive analytics and structural optimization, separately. In Section 4, we comment on the challenges and opportunities in this emerging interdisciplinary research area.

2. Principles of machine learning

Prediction is defined as establishing a predictive model between the output and input, which usually involves three steps [53,60–63]. The first step is the acquisition of the dataset, which consists of the target property of interest. The dataset can be obtained from experiments, numerical simulations, published data. There are also existing material databases [64], such as Materials Project [65], AFLOW [66], Inorganic Crystal Structure Database [67], Open Quantum Materials Database [68], and Polymer database [69]. In these databases, various material properties are collected, such as formation enthalpy, bandgaps, and modulus. However, the data of thermal transport properties is limited. The second step is identifying a set of quantities describing the characteristics of the target system that are strongly correlated with the property of interest. These quantities are often called descriptors [70]. Selected descriptors preferably fulfill several requirements [45,70]: (a) unique: all the systems should have exactly one representation; (b) descriptive: systems that are very different (similar) should be characterized by very different (similar) descriptor values; (c) complete: a "complete" representation should provide enough information to sufficiently differentiate systems; (d) simple: the determination of the descriptor should be fast and must not involve calculations as intensive as those needed for the evaluation of the property to be predicted [45,70]. The third step is applying the machine learning methods to discover the underlying relationship between descriptors and the target property. Two types of machine learning methods are usually adopted [71]. One is the conventional machine learning method that is required to manually select the descriptors based on understanding and expertise of the problem [71]. The other is the deep learning method that can automatically extract features of the system, and analyze the information from the raw data. This type of method is advantageous for learning problems in which the input is a structural image or matrix [71]. Both conventional machine learning and deep learning have two categories of learning methodology, which are supervised and unsupervised learning [72]. The supervised learning approach is used to solve regression and classification problems, while the unsupervised learning approach is used for deciding association and clustering problems [72,73]. Most prediction problems of thermal transport study are regression and target thermal transport properties are continuous. Thus, supervised regression methods are commonly used. Typical conventional machine learning methods include linear regression [62,74], support vector regression [51,61], Gaussian process regression [52,61,75,76], random forest regression [62,74,77,78], kernel ridge regression [51], eXtreme Gradient Boosting [51], artificial neutral network [53,60,62,77]. Popular deep learning methods include convolutional neural networks [75,79], and generative adversarial networks [80,81]. The schematic of the prediction process is shown in Fig. 1a. Usually, the dataset is divided into training, validation, and test sets [82]. The model is initially learned on a training set and further tuned by examining the validation set. For simplification, the training set and validation set are collectively

referred to as training data in Fig. 1a. The test set that is unseen in the training process is to test the prediction accuracy of the trained model. The prediction errors can be evaluated by some metrics, such as mean absolute error, root mean squared error, and determination coefficient (R^2) [58].

Optimization means selecting the desired element from a set of available alternatives under a specified criterion [83]. There are four essential ingredients in optimization: objective function, descriptor, evaluator of the objective function, and an optimization algorithm [64,84]. Firstly, the optimization problem needs to be clearly defined with a specific objective function under certain constraints [84]. Generally, the objective function is to minimize or maximize the target property [58]. Then, the descriptors that are appropriate to describe the design system should be determined. For example, in structural optimization, digital representations of the structures are mostly adopted as descriptors that are often encoded as an N-bit array, where each bit represents the value of a degree of freedom [56,85]. Next, the optimization process is carried out with certain optimization algorithms to find out the optimal solution. During optimization, the properties of selected candidates are evaluated to formulate the objective function, usually by the corresponding physics-based methods. For structural optimization, the degree of freedom of structure is usually large, for which the iterative optimization methods are more useful than direct methods [83]. In addition, randomness often presents in the iterative search procedure and thus stochastic optimization algorithms are more preferable [83], such as simulated annealing [86], particle swarm optimization [87], evolutionary algorithms [88], genetic algorithms [89], Bayesian optimization [90], Monte Carlo tree search [91] Fig. 1.b shows the schematic of the iterative optimization process. Firstly, the initial generation is produced by random selection from the design space. The target properties are evaluated and the values of the objective function are obtained. Then the optimization algorithm is applied to identify the new generation that may meet the expectations of optimization. This step will carry out iteratively until the optimal element is found or a certain criterion is satisfied, e.g. the best individual in each generation does not change over serval generations. With an optimization algorithm, one can efficiently search the desired properties instead of tedious trial and error.

3. Recent progress

Next, we introduce the recent progress of applying machine learning methods to study the thermal transport properties from the aspects of prediction and optimization.

3.1. Predictive analytics

In this section, we first introduce the recent advances in applications of machine learning to predict thermal properties. In particular, we demonstrate the advantage of machine learning in building surrogate models compared to physics-based analytical models. Then, we discuss the determining factors on the prediction accuracy of machine learning models. After that, two applications of machine learning surrogate models—high-throughput screening and machine learning potential, are covered.

3.1.1. Prediction as a surrogate model

Machine learning methods extract the relationships between the thermal transport properties and materials characteristics, such as elemental information, structural features, or other material properties [57,58,82]. Properly-trained machine learning models should possess comparable predicting accuracy of the methods used to generate data. For instance, if the training data is from numerical simulations, the machine learning prediction should be as accurate as the corresponding numerical results. Moreover, the machine learning prediction models can be regarded as semi-analytical models, which is advantageous over numerical simulation (solving physics-based equations) and experiment in that they can provide a fast prediction [75]. A properly-trained machine learning model should achieve higher predicting accuracy compared to the existing physics-based analytical model. In other words, if a machine learning model is developed, it is necessary to show this model has better accuracy than existing physics-based analytical models, as has been done in literature [50,61,75,92].

There have been quite some recent works that use machine learning methods to predict thermal conductivity, including lattice thermal conductivity and effective thermal conductivity. For example, Juneja et al. built a Gaussian process regression model to predict the lattice thermal conductivity of compounds based on a dataset of 120 pieces of data collected from Materials Project [66], of which the lattice thermal conductivity was obtained from first-principles DFT calculations [50] Fig. 2.a and Fig. 2b show the comparison between DFT calculations and predictions from the physics-based Slack model and machine learning-based Gaussian process regression model, respectively. The Slack model severely overestimates for all the compounds, while the predicted lattice thermal conductivities by Gaussian process regression agree very well with the calculated first-principles values. The Gaussian process regression model outperforms the physics-based Slack model by 1 order of magnitude in terms of accuracy [50]. Zhu et al. applied kernel ridge regression, artificial neutral network, and convolutional neural network to predict the lattice thermal conductivity of the single-chain polymer directly from the molecular structures based on around 300 pieces of data established by non-equilibrium molecular dynamics (NEMD) simulation [79]. The trained models present high prediction accuracy, among which the convolutional neural network shows the best performance with a mean absolute error of 5.20 W/mK, RMSE of 6.83 W/mK. Chen et al. built a Gaussian process regression model to predict the lattice thermal conductivity of inorganic materials using a dataset of 100 experimental data [52]. The accuracy of the Gaussian process regression model is comparable to previous semi-empirical models using descriptors that are more easily and rapidly accessible

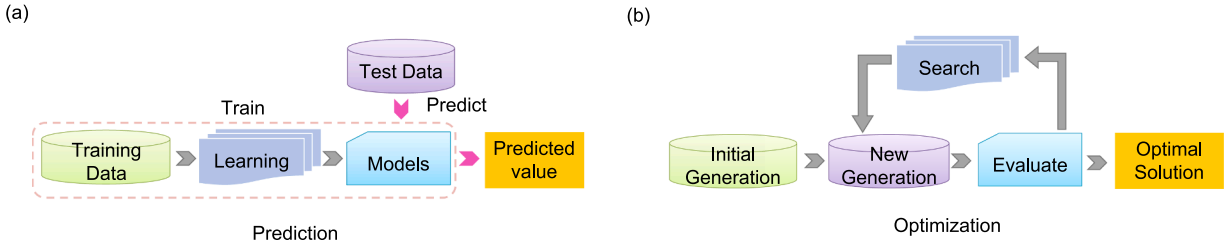


Fig. 1. The framework of applying machine learning approaches for (a) prediction and (b) optimization purposes. In the prediction process, the training data is first fed into machine learning algorithms to learn the relationship between descriptors and target properties. Then the model can be used to predict the test data that are unseen in the training process. In the optimization process, initial generation is firstly produced by certain criteria, e.g. random selection, and the corresponding properties are evaluated by physics-based methods. The machine learning optimization algorithm is then implemented to select the best individuals that are closest to the expectations of optimization to form the next generation. The searching process will carry out iteratively until the optimal solution is obtained.

[52]. These works generally develop machine learning models to establish the linkage between atomic features and lattice thermal conductivity. On the other hand, there are also efforts to link microscopic structure features with the effective thermal conductivity of heterogeneous materials. Wei et al. trained support vector regression, Gaussian process regression, convolutional neural network models to predict the effective thermal conductivity of two-dimensional composite materials and porous media based on around 1000 training data by lattice Boltzmann method [75]. All three machine learning models are more accurate than the physics-based effective medium theory models Fig. 2.c presents the prediction errors of convolutional neural network and physics-based effective medium theory models. The root mean absolute error of convolutional neural network is only 1.9%, which is smaller than that of effective medium theory models (7.3% for the Maxwell-Eucken model and 18.3% for the Bruggeman model). Rong et al. used a two-dimensional convolutional neural network to predict the effective thermal conductivity of three-dimensional composites with the input of two-dimensional cross-section images based on a training dataset of 2000 data obtained by the finite element method [92]. The prediction results of the convolutional neural network model agree quite well with the simulated values, and the accuracy is superior to three commonly used effective medium theory models.

Machine learning methods have also been used to predict other thermal transport properties, such as interfacial thermal resistance [61, 62]. For instance, Yang et al. tried five machine learning models, including linear regression, polynomial regression, decision tree,

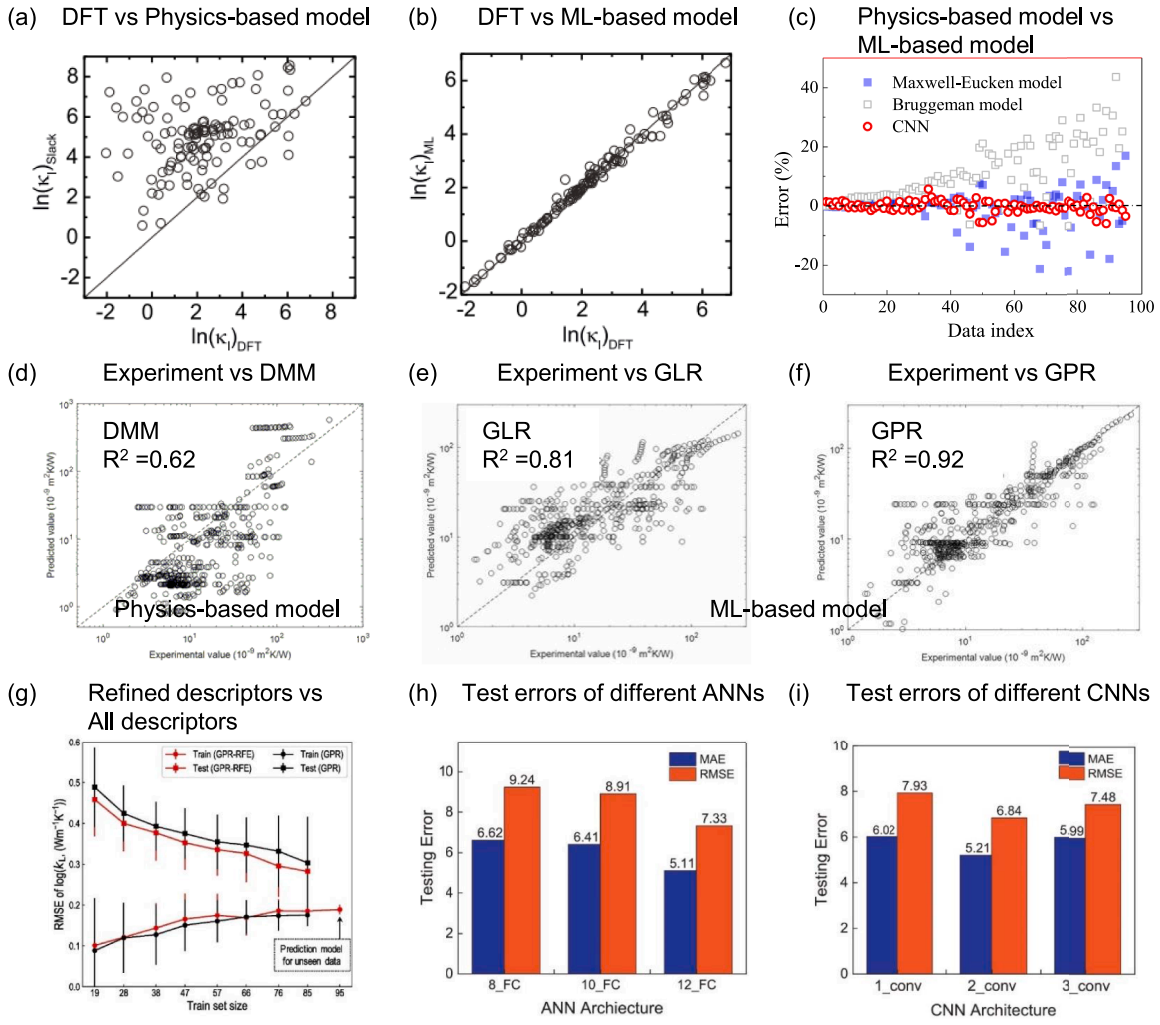


Fig. 2.. Application of machine learning methods to predict the thermal transport properties. (a) The comparison between DFT-calculated log-scaled lattice thermal conductivity(κ) vs Slack model-predicted ones [50]. The Slack model severely overestimates all the compounds. (b) The comparison between DFT-calculated log-scaled lattice thermal conductivity vs machine learning (ML) model-predicted ones [50]. The predicted lattice thermal conductivities by Gaussian process regression (GPR) agree very well with the calculated first-principles values. (c) The prediction errors of machine learning-based convolutional neural network (CNN) model and physics-based effective medium theory models [75]. The prediction errors of the CNN model are much smaller than the physics-based effective medium theory models. (d) The comparison between benchmark experimental data and predictions from the physics-based diffuse mismatch (DMM) model [61]. (e) The comparison between benchmark experimental data and predictions from machine learning-based generalized linear regression (GLR) model [61]. (f) The comparison between benchmark experimental data and predictions from machine learning-based GPR model [61]. (g) Prediction accuracy for GPR and GPR-RFE models trained using different train set sizes. RFE means recursive feature elimination, which is a descriptor dimension reduction technique, used to eliminate the irrelevant features from the total descriptor list [52]. The errors of GPR-RFE are lower than the GPR models, indicating that the dimensionality reduction step can improve model performance with lower prediction errors. (h) The test errors of artificial neural networks (ANN) with different architectures [79]. The performance of ANN of 12 fully connected layers is better than the network of 8 and 10 fully connected layers. (i) The test errors of CNN with different architectures [79]. The CNN model of 2 convolutional layers is better than that of 1 and 3 convolutional layers.

random forest, artificial neutral network, to predict the interfacial thermal resistance between graphene and hexagonal boron–nitride with the total data size of 1650 calculated by MD simulations [62]. A small mean square error of $0.045 \times 10^{-7} Km^2W^{-1}$ is obtained compared to the true values. Zhan et al. trained generalized linear regression, support vector regression, and Gaussian process regression to predict the thermal boundary resistance of interfaces based on 876 experimental data collected from 62 published papers [61]. Fig. 2d, e, and f show the comparison between benchmark experimental data and physics-based diffuse mismatch model, machine learning-based generalized linear regression, and Gaussian process regression models, respectively. The prediction coefficient R^2 of diffuse mismatch model, generalized linear regression, and Gaussian process regression are 0.62, 0.81, and 0.92, respectively, which demonstrates the higher predictive accuracy of machine learning models than the physics-based models.

3.1.2. Prediction accuracy

For predictive analytics, achieving the high accuracy of a prediction model with limited training data is highly desired. The prediction accuracy relies on the three key aspects: dataset, descriptors, and machine learning algorithms [57,58,82]. To achieve high accuracy, researchers should pay attention to the collection of the dataset, the selection of representative descriptors, and the machine learning algorithms that are suitable for the size of the available data and the dimensionality of descriptors.

The dataset is the cornerstone of data-driven study. The size of the dataset plays an important role in the success of machine learning [82,93]. Using more training data generally indicates that more information is included, and therefore results in improved accuracy and the generalization performance of the machine learning prediction [52,75,92,94]. For example, in Ref. [92], the authors used a convolutional neural network to predict the effective thermal conductivity of composites based on a training dataset of 2000 data obtained by numerical simulations. The mean absolute error and root mean square error of the test set decreased from 5.3% to 3.5% and 6.5% to 4.2%, respectively when the training data size increased from 200 to 2000. Although it is desirable to collect more training data, it is at an expense of large-scale experiments or numerical simulations [95–97]. In the thermal transport property area, the size of available data is still much smaller than in some other fields, such as material science. Therefore, it is more important to choose appropriate descriptors and machine learning algorithms to avoid large errors and overfitting.

In most studies, depending on the target application, the possible descriptors are firstly gathered based on physical understanding in literature. Then, the descriptors that are strongly correlated with the target properties are selected. Furthermore, the descriptors may be interrelated, and the redundant ones can be removed from the list. For example, in Ref. [61], twelve descriptors of determining the thermal boundary resistance are collected in total. Six of them were selected as "reliable" descriptors because they are easily collected from references and the provided values are more reliable. The predictions using all collected descriptors and the "reliable" descriptors have similar predictive accuracy, which is attributed to the physical correlation of the descriptors in the complete set. In Ref. [52], the authors applied the Gaussian process regression model to predict the lattice thermal conductivity of inorganic materials using a dataset of 100 experimental data. The strategy of recursive feature elimination was performed to eliminate the irrelevant features in the initial 63-dimensional feature vector and refine the descriptor list Fig. 2.g presents the prediction errors of Gaussian process regression models with the complete set of descriptors and refined set of descriptors by recursive feature elimination on the training and test set. The results show that the dimensionality reduction step improves model performance with lower test errors. In principle, the descriptor list should follow the requirements of being unique, descriptive, complete, and simple [45,70], as mentioned in Section 2.

With the same descriptors, the prediction accuracies of different machine learning algorithms could be different. For example, in Ref [62]., five machine learning models: linear regression, polynomial regression, decision tree, random forest, and artificial neutral network have been tried. The decision tree, random forest, and artificial neutral network are more accurate than the other two methods. In Ref. [61], the authors tried three machine learning methods and discovered that the accuracy of the Gaussian process regression model is higher than the generalized linear regression model (as seen in Fig. 2e and f). The performance of the algorithms heavily depends on the nature of the problems, including the type and dimensionality of descriptors, the size of the dataset, and the setup of hyper-parameters [98]. Therefore, determining which algorithm to use is case-dependent. It is difficult or even impossible to tell which algorithm will perform the best a priori. Researchers usually try several different machine learning algorithms and then select the best one as the final predictive model. Generally, for small datasets, simple models have fair performances. For big datasets with the high dimensionality of descriptors, complex models perform better [82]. However, it is still recommended to start with simple models before trying complex models. Complex models (i.e. with many parameters) can achieve high prediction accuracy because they can identify more hidden patterns of the input. However, too many parameters may lead to overfitting, which is detrimental to the prediction performance [82,99,100]. For example, in Ref. [79], the authors used kernel ridge regression, artificial neutral network, and convolutional neural network to predict the lattice thermal conductivity of the single-chain polymer Fig. 2.h and i present the test errors of different architectures of artificial neutral networks, and convolutional neural network, respectively. The number before "FC" and "conv" indicates the number of layers in the models. It shows that the performance of artificial neutral network of 12 fully connected layers is better than the network of 8 and 10 fully connected layers (Fig. 2h). Whereas, the convolutional neural network model of 2 convolutional layers is better than that of 1 and 3 convolutional layers (Fig. 2i). We suggest starting with basic algorithms to build a baseline model like linear regression [101], logistic regression [102], and linear support vector regression [103]. If large errors are obtained, more complex algorithms could be tried, such as kernel support vector regression [103], and neural networks [104]. Besides the accuracy, it is also necessary to consider the ease of implementation, interpretability of results, available time and cost, etc. In a word, selecting which algorithm depends on the objective of the problem. There is no single algorithm that works well for all problems.

3.1.3. Applications of the machine learning surrogate model

Many existing works demonstrate that machine learning predictive models can have a fast and accurate prediction of thermal transport properties. However, it is necessary to further explore how to better use the surrogate models to address major challenges that are inaccessible by previous physics-based approaches. To date, we believe there are two applications, in which machine learning surrogate models significantly outperform traditional methods. One is to couple with high-throughput techniques for screening the desired thermal transport properties [105–111]. The other is to establish machine learning interatomic potentials, which fills the gap between the first-principles MD and empirical-potential-based MD simulations [112–118].

High-throughput screening is to exhaust all possible material candidates in the design space and select the best one [111]. The pool of materials is usually too many for the first-principle prediction of thermal transport property due to the computational cost. Previously, analytical physics-based models like the Slack model are applied for accelerating the screening, but the prediction accuracy is usually low [78,119]. Machine learning models can be constructed as the surrogate of first-principle calculations to aid high-throughput prediction and screening of thermal transport properties with good accuracy. For example, Wang et al. trained an XGBoost model with structural and compositional descriptors to predict the lattice thermal conductivity of

crystalline materials [51]. They employed the learned surrogate machine learning model on the materials in the entire Inorganic Crystallographic Structure Database and identified that the heavy elements like Cs, Au, Hg, Tl, and Pb are helpful to reduce thermal conductivity. They further screened crystalline materials with low thermal conductivity such as $BiTe_2Tl$ and Cl_2CsI and validated using first-principles calculations [51]. More illustrative examples of high-throughput screening combined with machine learning to search for extreme thermal conductivity are summarized in a previous review paper [57]. With the adaptation of the machine learning algorithms, more new materials with excellent thermal transport properties are identified. They are generally based on numerical calculations. Material fabrication and experimental measurement are further needed to validate the results of high-throughput screening.

Another application is the machine learning potential. As MD simulation is probably the only effective approach to modeling complex crystals, disordered solids, it has been widely used for modeling thermal transport processes [11]. First-principles-based MD simulations have high accuracy but the computational cost is extremely large, limiting the simulation to ~100 atoms and ~10 ps [120]. In contrast, empirical potential-based MD simulations can simulate up to $\sim 10^6$ atoms and ~1000 ns [121]. However, the accuracy is limited by empirical interatomic potentials. Because the ab-initio potential energy surface in the atomic configurational space is high-dimensional, it cannot be accurately fit by simple functional forms that are manually assigned based on the pre-knowledge of the interatomic bonding nature. Hence, improving the accuracy of empirical interatomic potential is of great difficulty [82,122,123]. On the contrary, machine learning potential can flexibly fit the ab-initio potential energy surface in a data-driven manner, without the need to assign specific functional forms beforehand [82,122,123]. Therefore, the machine learning potential is an appropriate choice to fill the gap between the first-principles calculations and MD simulations for modeling thermal transport. In the past five years, machine learning potentials have been successfully developed and employed to model the thermal properties of crystalline [124], amorphous materials [112,116], alloys [125,126], heterogeneous structures [127], and various 2D materials, such as graphene [117], silicene [114]. Widely-used machine learning potential frameworks include neural network potentials [128], Gaussian approximation potentials [129], and spectral neighbor analysis potential [130]. The training data come from first-principles DFT calculations. It has been shown the thermal properties, such as phonon dispersion relation, scattering rate, Grüneisen parameter, thermal expansion coefficient, and total thermal conductivity that are calculated from the machine learning potential are very close to the first-principle calculations [114,115,131,132]. We hope that in the future, more machine learning potentials can be integrated into popular MD simulation packages (such as LAMMPS [133] and GPUMD [134]). This will greatly promote thermal transport research.

3.2. Optimization and design

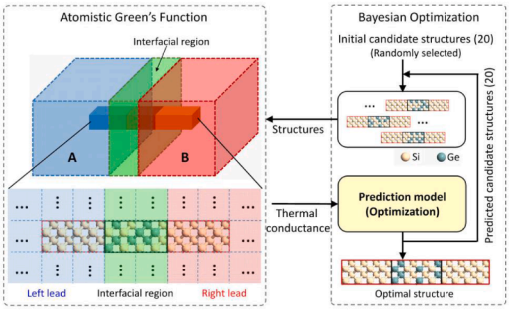
In addition to prediction analytics, another important category of machine learning application in thermal transport study is structural optimization. In this section, we first introduce applications of optimization methods to search for the optimal structure with the desirable thermal transport property. In particular, we present the superior effectiveness of optimization algorithms compared to the previously used random search or intuitive designs. Then, we discuss the strategies for improving optimization efficiency. Finally, we illustrate the important role of optimization in discovering new patterns and the underlying physics.

3.2.1. Optimization as a search tool

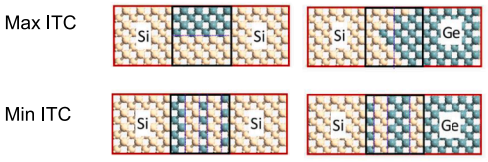
In order to obtain desired thermal transport properties, the design and optimization of material structures are often carried out [135–138]. In structural optimization, the studied system usually has a large amount of degree of freedom, which results in an extremely large design space [135–138]. A thorough search in the space is usually prohibitive and inefficient. In the contrast, machine learning optimization algorithms are advantageous at learning the complex functions between quantities and can thus efficiently search the design space under the guidance of the learned functions [139,140].

Nanostructures possess large degrees of freedom for the structural design [141–143] and new physical mechanisms behind, especially the ballistic transport effect and wave effect, compared to the macroscopic thermal transport [11,27,144]. Therefore, structural design of nanoscale thermal transport has attracted considerable interest, such as atomic-interface [55], alloys [54,145], and thermal metamaterials [56,85,146–153], e.g., superlattice [56,85,150], pillared nanoribbon [151], and nanomesh [152,153]. In 2017, Ju et al. developed a method combining AGF simulations and Bayesian optimization for optimization of thermal conductance across atomic interfaces [55]. It is an early demonstration of machine learning-assisted design of nanoscale thermal transport Fig. 3.a shows the workflow of the optimization scheme combing the AGF simulations and Bayesian optimization. This optimization scheme can identify the optimal structures from calculations of only a few percent of all candidates [55]. They optimized the interfaces for the maximum thermal conductance between Si and Ge. The optimized structures are shown in Fig. 3b. For the maximum thermal conductance, the optimal structure is composed of a continuous path of Si or Ge atoms, which is consistent with the physical intuition. For the minimum thermal conductance, the structure is composed of a random arrangement of layers, which results from the best balance between wave interference and interfacial scattering effects. Interestingly, the resulted random structure from machine learning is consistent with earlier intuition-driven proposal of using random layers to localize phonons and reduce thermal conductivity [154]. Dieb et al. also adopted the Monte Carlo tree to optimize the conductance of interfaces with more atoms [155]. Subsequently, the optimization of thermal conductivity (conductance) by designing the atomic configuration of alloys, superlattice, and nanoribbons with nanopillars has been demonstrated. For example, Yan et al. combined the Bayesian optimization with a high-throughput thermal conductivity calculation to search for the lowest thermal conductivity atomic configuration of SiGe alloy [145]. It was found that layered structures are most beneficial for reducing the thermal conductivity among all atomic configurations, caused by the flattening of the phonon dispersion curve and the filtering effect of alternate Si/Ge layers. Chowdhury et al. cooperated genetic algorithm with MD simulations to identify the structure of random multiple layers composed of Si and Ge with the minimum thermal conductivity [56] Fig. 3.c presents the schematic of the genetic algorithm-based optimization method for optimizing the thermal conductivity of Si/Ge superlattice and the optimization results are shown in Fig. 3d. The intuition-based manual search yields only a local minimum, while a genetic algorithm-based approach can efficiently identify the globally minimum thermal conductivity by only exploring a small fraction of the design space. Hu et al. performed Bayesian optimization to minimize coherent phonon heat conduction in aperiodic GaAs/AlAs superlattices quantified by AGF simulations [85]. They further fabricated the optimal superlattice structure according to the optimization results. The experimentally measured thermal conductivity agrees well with the full coherent calculations, which are significantly smaller than the conventional periodic superlattice [85]. In their later work, they proposed a pattern analysis method, through which the optimal structure of multivariable optimization can be effectively obtained [156]. They first demonstrated the effectiveness of this method with 1D atomic chains and then validated it with 3D superlattices. It was found that the thermal conductance of optimal aperiodic structure obtained by the pattern analysis method is close to the optimization result of Monte Carlo tree search [156]. Wan et al. also adopted the strategy of AGF simulation combining Bayesian optimization to minimize the thermal conductance across graphene nanoribbon by designing nanopillared nanostructures

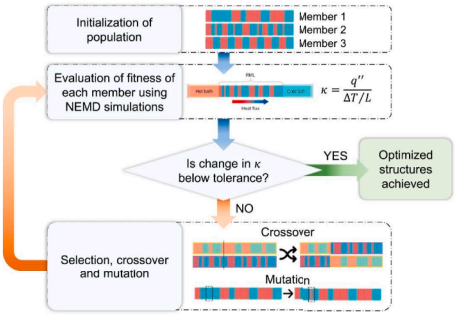
(a) Atomistic Green function combined with Bayesian optimization



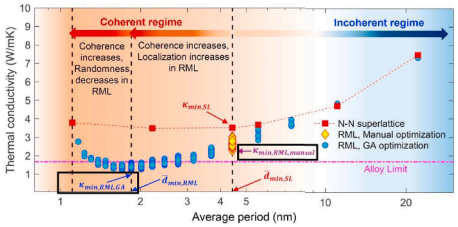
(b) Optimal atomic configurations



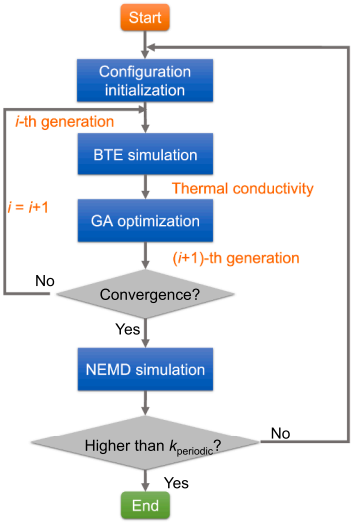
(c) Non-equilibrium molecular dynamics simulation combined with genetic algorithm



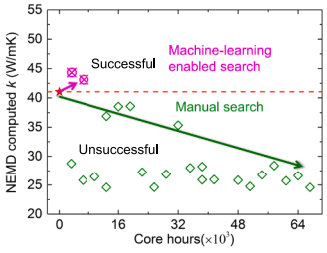
(d) Manual search vs genetic algorithm-based search



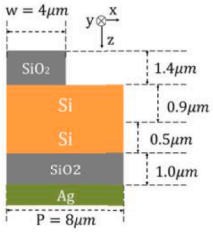
(e) Genetic algorithm-enabled two step search



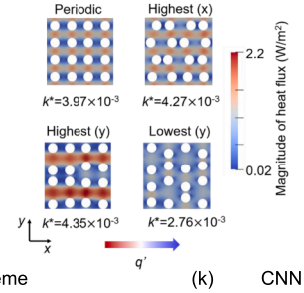
(f) Manual search vs genetic algorithm-enabled search



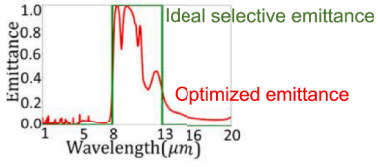
(h) Optimal thermal photonic structure



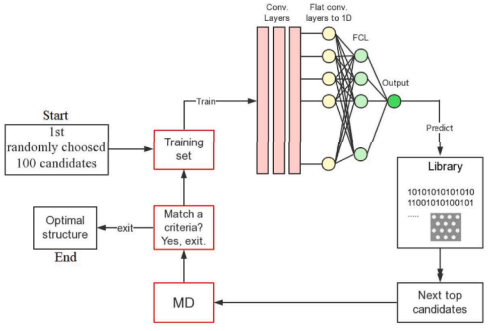
(g) Representative nanoporous structure



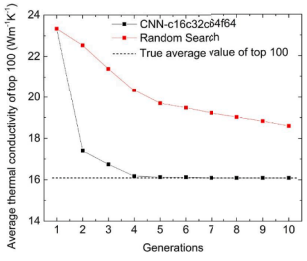
(i) Emittance of the optimal structure



(j) CNN prediction model based search scheme



(k) CNN prediction model based search vs random search



(caption on next page)

Fig. 3. Application of machine learning methods to optimize the structures for desired thermal transport studies. (a) Schematics of the optimization scheme combing the atomistic Green function and Bayesian optimization [55]. (b) Optimal structures with the maximum and minimum interfacial thermal conductance (ITC) for Si-Si and Si-Ge interface [55]. (c) Schematic of the genetic algorithm-based optimization method for optimizing the thermal conductivity of Si/Ge superlattice [56]. (d) Variation of thermal conductivity with average period length [56]. Random multiple layer structures obtained using manual intuition-based optimization are denoted by diamonds and random multiple layer structures obtained using machine-learning-based optimization are denoted by filled circles. The dashed purple line marks the random alloy limit. (e) The workflow of the two-step search process [153]. BTE simulation is demonstrated for pre-screening and then NEMD simulation is performed to validate the optimal configuration discovered by genetic algorithm. (f) The comparison of optimization output and the overall time cost between the genetic algorithm enabled search method and manual search method. 2 successful and 21 unsuccessful cases are distinguished using magenta and green colors [153]. (g) The magnitude of BTE computed heat flux in four typical nanoporous configurations: periodic configuration, two configurations with the highest thermal conductivity obtained from genetic algorithm enabled two-step search, and also the configuration with lowest thermal conductivity. The corresponding BTE simulated thermal conductivities are shown below the structures [153]. (h) Optimized structure of the thermal photonic emitter for radiative cooling with three material candidates (Al_2O_3, Si, and SiO_2) [164]. The optimal structure turned out to consist of only Si and SiO_2. (i) Emittance of the optimal structure. The green rectangular line denotes the ideal selective emittance [164]. (j) Schematic of the search algorithm based on the CNN model. 100 structures are randomly selected from the pool of all candidate structures and their thermal conductivity is computed by MD simulations. The CNN model is trained as the first generation surrogate prediction model. Then the CNN model is used to predict the thermal conductivity of all remaining structures and select new 100 structures with the lowest thermal conductivity. The true thermal conductivities for these 100 new structures are calculated from MD simulations and this data is added into the training set, which is used to train the next generation CNN model [152]. The CNN surrogate prediction model is updated iteratively like this until the optimal solution is obtained. (k) The average thermal conductivity of top 100 structures in each generation during the inverse design optimized by search scheme based on the CNN model and random search [152].

[151]. It is observed that thermal conductance decreases non-monotonically with the increasing nanopillars, due to the competing effects of resonant hybridization and phonon coherence. Previous studies have concluded that the disordered structures lead to lower thermal conductivity than periodic counterparts, due to phonon Anderson localization [157–160]. These studies of machine learning-assisted material design further explore the lower limit of thermal conductivity (conductance) [56,85,152]. However, Wei et al. challenged this accepted knowledge by searching for exceptions [153]. They designed a "two-step" search protocol coupled with a genetic algorithm to search for disordered nanoporous structures with a higher thermal conductivity than the periodic ones [153]. The manual search is expensive and unsuccessful. On the contrary, with much fewer computational costs, the genetic algorithm enabled method successfully searches out disordered structures with higher thermal conductivity. More recently, Chowdhury et al. discovered an unexpected lattice thermal conductivity enhancement in aperiodic superlattices as compared to periodic superlattices by an adaptive search with a CNN-based prediction method [150]. They found that the CNN-based search process can successfully identify the enhancement within two iterations of the search utilizing 200 CPU hours. In contrast, the manual random search fails even after double the simulation hours spent. Although the GA search can identify reasonably high thermal conductivity, but still not exceed the reference superlattice thermal conductivity. These works open a scheme of performing optimizations in the counterintuitive space, and potentially obtain unexpected outcomes and uncover the associated new physics.

Some studies carried out multifunctional optimization of the thermoelectric figure of merit. For example, Masaki et al. carried out structural optimization in porous graphene nanoribbon to achieve high thermoelectric performance combing Green's function method and Bayesian optimization [161]. The efficiency of multifunctional structural optimization is five times higher than that achieved by random search. The optimized thermoelectric figure of merit was found to be up to 11 times of the pristine graphene nanoribbon. Cui et al. also utilized Bayesian optimization with Green's function method to search for the optimal thermoelectric conversion efficiency of defect -graphyne nanoribbons [162]. The efficiency of Bayesian optimization is significantly higher than that of random optimization. Even in the worst round of optimization, the optimal defect -graphyne nanoribbon could be efficiently searched out by only calculating 4.35% of all structure candidates. The thermoelectric figure of merit of the optimal defect -graphyne nanoribbon (length and width are 11.846 nm and 1.453 nm) is as high as 2.315, which is 5 times that of the pristine -graphyne nanoribbon. The structure design for radiative properties has also been investigated.

In addition to the thermal conductivity (conductance), and thermoelectric performance, the optimization of radiative properties has also been performed. For example,Sakurai et al. designed an ultranarrow-band wavelength selective thermal radiator combining the Bayesian optimization and thermal electromagnetic field calculation [163]. The resulting metamaterial is an aperiodic multilayered metamaterial exhibiting a *Q*-factor of 188, which is significantly higher than those of structures empirically designed and fabricated in the past. Guo et al. combined the rigorous coupled wave analysis and Bayesian optimization to design thermal photonic structures for radiative cooling applications [164]. Bayesian optimization method can identify the optimal structure by calculating less than 50 candidates out of 19,683 candidates, while 4500 were needed for a random search. The optimal structure turns out to consist of only Si and SiO_2 (Fig. 3h) and shows spectral selectivity which exactly matches the atmospheric window (Fig. 3i). Hu et al. implemented the Monte Carlo tree search algorithm to maximize the power density and system efficiency of a thermophotovoltaic system by optimizing Tamm emitter structures. Through optimization, the non-trial aperiodic Tamm emitters are obtained and the metal-side one is preferable in terms of the TPV performance [165]. They also optimized the Tamm emitters at the infrared range for achieving a high *Q*-factor and high emissivity simultaneously with Monte Carlo tree search. The optimal structure with a *Q*-factor of 508 and an emissivity peak of 0.92 at 4.225 μm is obtained [166].

3.2.2. *Optimization efficiency*

In optimization problems, the systems for design usually possess a large degree of freedom. The optimization problem is thus usually high-dimensional. In this case, the advanced optimization algorithms are more efficient in searching the optimal structure with desired thermal transport properties than random search or tuning based on physical intuition. On the other hand, the computational or experimental cost is relatively large so that the amount of available data is limited. As such, the number of candidates that can be considered is limited and thus the efficiency of optimization is important [58]. However, current structural optimization with efficient optimization algorithms still suffers from the large computational cost caused by numerical simulations in evaluating the thermal transport properties of candidates and the computations in optimization itself. To improve the efficiency of the whole optimization process, one way is to maximize the efficiency of the optimization algorithm itself. In principle, the complexity of the problem, e.g. the degrees of freedom and the total number of candidates, has a great impact on the performance of a certain algorithm [57]. For example, the Bayesian optimization is extremely effective when the total number of candidates is on the order of several hundred thousand [64,165,167]. When the total number of candidates is much larger or even infinite, the Bayesian optimization becomes less efficient [64,165,167]. Besides, the efficiency also depends on the details of the algorithm, such as the

selection of descriptors, the number of candidates considered in each iteration, and the hyper-parameters in the algorithm [89,90]. For instance, binary flag values, coulomb matrix, and mass matrix descriptors are more efficient than the eigenvalue descriptors to optimize superlattice structures for minimum thermal conductance [64]. Therefore, the efficiency of a specific algorithm is case-sensitive. Even though, it is recommended to select the algorithms based on the complexity of optimization problems and have a careful examination of the selection of those determining factors.

Another different approach is to reduce the cost of evaluating the properties of candidates. One strategy is to optimize hierarchically, i.e., first pre-screening with a crude estimator and then validated by the accurate one. For example, in Ref. [153], the authors proposed a hierarchical "two-step" search protocol to optimize the thermal conductivity of nanoporous graphene Fig. 3.e shows the two-step search process, in which the gray phonon BTE simulation is used to compute the thermal conductivities of candidates as a substitute for computational expensive NEMD simulations. After pre-screening by a genetic algorithm based on BTE, NEMD simulations are applied to validate the optimized configuration Fig. 3.f presents the comparison of optimization output and the overall time cost between the genetic algorithm-enabled search method and the manual search method. The manual search is shown to be expensive (spent 67,200 core hours) and unsuccessful. Whereas in genetic algorithm enabled search, unexpected thermal conductivity enhancement was successfully discovered in certain structures with random pores, at a fraction of the computational cost of the manual search. Another strategy is to build a prediction surrogate model to replace the computational expensive numerical simulations. For example, Wan et al. proposed an inverse design scheme based on a convolutional neural network prediction model to minimize the thermal conductivity of nanoporous graphene [152]. The schematic is shown in Fig. 3j. Firstly, a total of 100 structures are randomly selected from the pool of all candidate structures and their thermal conductivities are computed by MD simulations. Based on these 100 training data, the prediction model is trained as the first generation surrogate prediction model. The first generation of the model is then applied to predict the thermal conductivities of all remaining structures and select new 100 structures with the lowest thermal conductivities. The real thermal conductivities for these 100 new structures are calculated from MD simulations and these data are added into the training set, which is used to train the next generation of the prediction model. The convolutional neural network surrogate prediction model is updated iteratively like this until the optimal structures are searched out. These two strategies can avoid direct computational-expensive calculation and thus accelerate the optimization process Fig. 3.k shows the average thermal conductivity of the top 100 structures in each generation during the inverse design optimized by a search scheme based on the model and random search. It is seen that the average thermal conductivity of the random search approach converges slowly. On the contrary, the search scheme based on the model quickly converges to true optimal value at the 7th generation. This work shows that the search scheme based on the model is able to find the structure of the lowest thermal conductivity more efficiently than the random search [152].

The introduced strategies above are still based on the optimization algorithms that are rule-based approaches including iterative searching case by case, such as genetic algorithms [89], Bayesian optimization [90], Monte Carlo tree search [91]. Numerical simulations are used in each step to compute the thermal property of candidates to help modify the searching process. The efficiency of these stochastic algorithms is limited by their random-search nature and hence is insufficient for complex design. Recently, a quite different design strategy has been used for the design of photonic structures, which is the deep neural network-based design [168–170]. As mentioned in Section 3.1, the neural networks can be used to establish the mapping from structural parameters and the property, which are often referred to as forward-modeling networks. In contrast, another type of neural network takes the property as the input and directly output the structure, which is referred to as inverse-design networks [168–170]. Typical deep learning architectures include multilayer perceptron [171], convolutional network [172], recurrent network [173], generative model [174], etc. These neural networks can discover useful information automatically from a huge amount of data and accomplish the design in a fraction of a second without needing any iterative optimization, in sharp contrast to physics- or rule-based approaches, which may be also suitable for the thermal transport property design.

3.2.3. *Applications of the machine learning search tool*

Owing to the efficient search, the optimization scheme is particularly suitable for material design as introduced in Section 3.2.1. In addition, the optimization scheme can also be helpful for the discovery of new physical mechanisms. For example, in Ref [85]., although lower thermal conductivity in the optimized random structure than periodic structures is expected, the authors studied the underlying mechanism of phonon transport in the optimal aperiodic superlattice. They statistically identified local patterns that are important to reduce thermal conductivity. The local patterns in the optimal structure are weakly correlated, which can separately localize phonons at different frequency ranges [85]. The optimized aperiodic structure is formed by connecting those local patterns with or without overlaps that introduce constructive or destructive interference over broad phonon frequencies. In Ref. [56], the authors discovered two unexpected features of the optimization result. One is the lowest thermal conductivity occurs at a smaller average period in the aperiodic superlattice than in the periodic one. Through analysis, they explained that the location of the average period is the result of a tradeoff between interface density and space for randomizing the layer thicknesses. Another interesting phenomenon is the lowest thermal conductivity occurs at a moderate rather than maximum degree of randomness in the layer thickness. This is because an intermediate degree of randomness creates structures of small and large layer thicknesses interspersed among each other, which is in favor of coherent phonon localization. In Ref. [153], even the outcome of higher thermal conductivity in the optimized random structure than in periodic structures is unexpected and unguaranteed. Further, after knowing the optimal disordered structures, the authors investigated the mechanism that gives rise to the unexpected thermal conductivity enhancement in disordered structures. They inspected the four representative configurations: periodic configuration, two configurations with the highest thermal conductivity obtained from the genetic algorithm enabled two-step search, and also the configuration with lowest thermal conductivity (shown in Fig. 3g). They observed the significant influence of distribution uniformity of pore arrangement and accordingly proposed two structural parameters—shape factor and channel factor, to describe the characteristics of pore arrangement. The unexpected thermal conductivity enhancement in the unusual structures can be attributed to the large shape factor and channel factor that dominate over the phonon localization. In Ref. [150], the mechanism about unexpected enhancement of thermal conductivity of aperiodic superlattice is studied by analyzing the contribution of interfacial resistance. In periodic superlattice, the relatively large layer thicknesses are above the coherence length of most phonons, as a result of which the contribution of coherent phonon transport to the thermal conductivity is quite low. However, in optimized aperiodic structures, some layers are reduced, which increase the coherent phonon contribution, and lower the apparent thermal resistance of the interfaces. Discovering exceptions is a route for advancing sciences but a challenging and risky process. Since it usually involves a well-thought hypothesis, numerous trials, and errors, but most often still ends with no success. The machine learning optimization scheme can discover small probability events to further enlighten the exploration of the intriguing physics behind. Using machine learning to challenge the accepted knowledge may become a routine approach in developing new physical mechanisms in the future.

4. Outlook

Although advancements have been achieved, machine learning for thermal transport is still under development, in which outstanding challenges and great opportunities coexist.

4.1. Small dataset

Machine learning methods are of statistical nature and thus a large amount of data are required, as large as hundreds, thousands, or even millions to obtain accurate models [100]. The inaccessibility of training data not only hinders the learning of relationships but also deteriorates the capability of prediction in the unexplored domain. However, as mentioned in Section 3.1, the available data for thermal transport study are typically much smaller compared to other research fields, such as image processing and industrial manufacturing. Simply expanding the dataset is not only unpractical due to the high cost but also leads to a highly complex model, making it difficult to interpret the underlying physics [93]. One possible solution is to incorporate machine learning models with different fidelities. Compared to the high-fidelity data, the low-fidelity data could be obtained easily/cheaply and can still provide useful information. The multi-fidelity models can achieve the desired accuracy at a reasonable cost [175,176]. Another possible solution is to use the transfer learning technique [177,178]. In transfer learning, when there is limited training data for the target property, an intermediate property, which holds physical connections with the target property, can be first pre-trained using available big data. The final prediction model for the target property can then be obtained by adjusting minor parameters in the pre-trained model using small training data. The above strategies are limited in the data-driven framework, which still requires labeled data. Embedding domain knowledge is another way to reduce data requirements. For example, physics-informed neural networks have been applied to approximate the physical governing equations, such as the Laplace equation [179], Boltzmann transport equation [180], Navier-Stokes equations [181]. In physics-informed neural networks, the physical governing equation is embedded as a constraint of the loss function. Thus, the solutions of the equation are learned in a physics-constrained manner without the need for any labeled training data. This data-free method provides new guidance for the small data in machine learning. The small dataset issue arises from the gap between “big data” that are required for the statistical nature of machine learning and “small data” that can only be collected for large costs. Therefore, the innovation in the way of learning knowledge from data and advancement in physics-based methods is more essential to tackle the issue of “small dataset”.

4.2. Identify effective descriptors

As reviewed before, for both predictive analytics and optimization, the impact of descriptors on the final results is obvious. However, finding the appropriate descriptors is nontrivial. When there is a pool of relevant descriptors, some strategies can be used to reduce the dimension of the feature space and help identify the most representative descriptors, such as the least absolute shrinkage and selection operator [61], principal component analysis [182], and linear discriminant analysis [100]. However, in some cases, the representative capabilities of existing descriptors are not strong enough. There is often a lack of descriptors instead of “too many descriptors”. Therefore, it is significant to develop new descriptors, especially, physics-based descriptors [183]. Physics-based descriptors contain the physical mechanism and are thus powerful for taking advantage of the physical correlations. Furthermore, physics-based descriptors can help investigate the underlying physics and is also important for fundamental theoretical studies. Some works tried to develop proper descriptors for describing the thermal transport problems for specific systems [53,183]. Building physics-based descriptors is not easy, which requires expert knowledge of the studied problems. It is expected that more physics-based descriptors can be identified to help improve the effectiveness of establishing predictive models, searching, and optimization.

4.3. Generate dataset from experiments and validate with experiments

The major approach of combining machine learning methods and physics-based methods is to generate data through physical-based methods, and then use machine learning to learn the relationship between quantities. Nowadays, most machine learning works focus on computational thermal transport due to ease and speed of investigation. Few studies combine machine learning with experimental studies. On the one hand, measuring thermal-physical properties is usually difficult and thus the experimental data is less accessible. On the other hand, experimental results usually lack consistency, caused by the sample differences in synthesis method, experimental conditions, which could increase uncertainties of the data. However, it is still meaningful to explore the potential applications of experiments in machine learning. One solution is to generate more experimental data by high-throughput experiments [184,185]. Another is to integrate experimental data with numerical data through data integration techniques, such as Extract, Transform, Load (ETL) [186,187]. The integrated data may improve data quality, provide a unified view of the studied system, and help gain more meaningful insights. Furthermore, hierarchical machine learning can be adopted, i.e., learning with numerical data and validation with experiments [85,163]. Generally, there are discrepancies between the simulation-based design and actual realization by experiment owing to practical issues in the fabrication process. How to use these predictions and optimizations from machine learning to guide the experiments or produce the actual realization of material with desired thermal transport properties deserves more attention. In the future, we suggest that experimental constraints and uncertainties can be considered at the stage of machine learning prediction and optimization. In this way, there is a better chance the machine learning results may be validated. Nevertheless, experimental validation is critical for machine learning research to make a real impact.

4.4. Discover new physics

Even with many successful applications, few new fundamental mechanisms have been discovered and thus it is difficult to say there are sufficient breakthroughs in employing machine learning to advance our understanding of thermal transport fundamentals. We view it as both a current challenge but also ample future opportunities in this aspect. The real question is what machine learning can help us discover that we cannot without machine learning. With the further integration of these two fields, we hope machine learning can be better combined with a physics-based method to solve real challenges in current technologies. We suggest the opportunity may be in finding new physics. Scientific history has taught us that science has often been advanced by discovering exceptions, such as the discovery of Raman scattering [188], which is an inelastic scattering and a rare event as compared to the known elastic scattering at that time. The outliers or the small-probability events usually contain new physics to be discovered. The outliers stay near the boundaries, while the general case is that we have no or little data near the boundaries. Therefore, it needs induction from those interpolative data points to extrapolative data points to discover new physics. The current machine learning methods seem to be weakly capable of doing extrapolation. With the advancement of algorithms, this problem may be resolved. If machine learning can help us discover new physics in a more systematic manner and without relying on human intuition, it will bring huge potential for thermal transport.

5. Summary

In this perspective, we focus on the recent progress of machine learning-aided thermal transport studies from the aspects of prediction and optimization. We provide our own perspectives about the advantages of machine learning methods in comparison to physics-based methods. The machine learning prediction methods can serve as surrogate models and perform better than the physics-based analytical model for predicting the thermal transport properties. On the other hand, machine learning optimization algorithms can serve as a search tool for optimal structures of desired thermal transport properties, with superior effectiveness compared to random search or tuning based on intuition. Furthermore, we discuss how to improve the accuracy of prediction models by collecting more data, selecting representative descriptors, and suitable learning algorithms accounting for the complexity of the problem and the size of the accessible data. We also summarize possible solutions to improve the design efficiency through a hierarchical optimization process or building a prediction surrogate model as a substitute for computational expensive numerical simulations in the iterative search. Finally, we outline several challenges and opportunities in better utilizing machine learning-based methods to advance the thermal transport research, including developing machine learning methods suitable for the small dataset of the thermal transport field, discovering effective physics-based descriptors, generating dataset from experiments and validating machine learning results with experiments, and making breakthroughs via discovering new physics.

Declaration of Competing Interest

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

Acknowledgment

H.B. and H.W. acknowledge the support by the National Natural Science Foundation of China (No. 52122606).

References

[1] Chen G. Nanoscale energy transport and conversion: a parallel treatment of electrons, molecules, phonons, and photons. Oxford; New York: Oxford University Press; 2005.
[2] Szczech JR, Higgins JM, Jin S. Enhancement of the thermoelectric properties in nanoscale and nanostructured materials. J Mater Chem 2011;21(12):4037–55.
[3] Luo T, Chen G. Nanoscale heat transfer – from computation to experiment. Phys Chem Chem Phys 2013;15(10):3389–412.
[4] Lindsay L, Hua C, Ruan XL, Lee S. Survey of ab initio phonon thermal transport. Mater Today Phys 2018;7:106–20.
[5] Hu M, Yang Z. Perspective on multi-scale simulation of thermal transport in solids and interfaces. Phys Chem Chem Phys 2021;23(3):1785–801.
[6] Shao C, Yu X, Yang N, Yue Y, Bao H. A review of thermal transport in low-dimensional materials under external perturbation: effect of strain, substrate, and clustering. Nanoscale Microscale Thermophys Eng 2017;21(4):201–36.
[7] Zhang Z, Ouyang Y, Cheng Y, Chen J, Li N, Zhang G. Size-dependent phononic thermal transport in low-dimensional nanomaterials. Phys Rep 2020;860:1–26.
[8] Chen XK, Chen KQ. Thermal transport of carbon nanomaterials. J Phys Condens Matter 2020;32(15):153002.
[9] Zhan H, Nie Y, Chen Y, Bell JM, Gu Y. Thermal transport in 3D nanostructures. Adv Funct Mater 2020;30(8):1903841.
[10] Zhao D, Qian X, Gu X, Jajja SA, Yang R. Measurement techniques for thermal conductivity and interfacial thermal conductance of bulk and thin film materials. J Electron Packag 2016;138(4).
[11] Bao H, Chen J, Gu X, Cao B. A review of simulation methods in micro/nanoscale heat conduction. ES Energy Environ 2018;1(39):16–55.
[12] Ma T, Chakraborty P, Guo X, Cao L, Wang Y. First-principles modeling of thermal transport in materials: achievements, opportunities, and challenges. Int J Thermophys 2019;41(1):9.
[13] McGaughey AJH, Kaviany M. Phonon transport in molecular dynamics simulations: formulation and thermal conductivity prediction. Advances in Heat Transfer, Elsevier; 2006. p. 169–255.
[14] Chen J, Zhang G, Li B. How to improve the accuracy of equilibrium molecular dynamics for computation of thermal conductivity? Phys Lett A 2010;374(23):2392–6.
[15] Schelling PK, Phillpot SR, Keblinski P. Comparison of atomic-level simulation methods for computing thermal conductivity. Phys Rev B 2002;65(14):144306.
[16] Zhang W, Fisher TS, Mingo N. The atomistic green's function method: an efficient simulation approach for nanoscale phonon transport. Numer Heat Transf B Fundam 2007;51(4):333–49.
[17] Sadasivam S, Che Y, Huang Z, Chen L, Kumar S, Fisher TS. The atomistic Green's function method for interfacial phonon transport. Annu Rev Heat Transf 2014;17:89–145.
[18] Mingo N, Volz S. Green's function methods for phonon transport through nanocontacts. Thermal nanosystems and nanomaterials. Berlin, Heidelberg: Springer; 2009. p. 63–94.
[19] Wang JS, Agarwalla BK, Li H, Thingna J. Nonequilibrium Green's function method for quantum thermal transport. Front Phys 2014;9(6):673–97.
[20] Narumanchi SVJ, Murthy JY, Amon CH. Submicron heat transport model in silicon accounting for phonon dispersion and polarization. J Heat Transf 2004;126(6):946–55.
[21] Murthy JY, Narumanchi SVJ, Pascual-Gutierrez JA, Wang T, Ni C, Mathur SR. Review of multiscale simulation in submicron heat transfer. Int J Multiscale Comput Eng 2005;3(1):5–32.
[22] Murthy JY, Mathur SR. Computation of sub-micron thermal transport using an unstructured finite volume method. J Heat Transf 2002;124(6):1176–81.
[23] Murthy JY, Mathur SR. Finite volume method for radiative heat transfer using unstructured meshes. J Thermophys Heat Transf 2012;12(3):313–21.
[24] Zahiri S, Zuo J, Shen Y, Bao H. Numerical investigation of ballistic-diffusive heat transfer through a constriction with the Boltzmann transport equation. Appl Therm Eng 2018;141:126–33.
[25] Zhai S, Zhang P, Xian Y, Zeng J, Shi B. Effective thermal conductivity of polymer composites: theoretical models and simulation models. Int J Heat Mass Transf 2018;117:358–74.
[26] Ngo IL, Truong VA. An investigation on effective thermal conductivity of hybrid-filler polymer composites under effects of random particle distribution, particle size and thermal contact resistance. Int J Heat Mass Transf 2019;144:118605.
[27] Cahill DG, Braun PV, Chen G, Clarke DR, Fan S, Goodson KE, Keblinski P, King WP, Mahan GD, Majumdar A, Maris HJ, Phillpot SR, Pop E, Shi L. Nanoscale thermal transport. II. 2003–2012. Appl Phys Rev 2014;1(1):011305.
[28] Liu H, Wang H, Zhang X. A brief review on the recent experimental advances in thermal rectification at the nanoscale. Appl Sci 2019;9(2):344.
[29] Min S, Blumm J, Lindemann A. A new laser flash system for measurement of the thermophysical properties. Thermochim Acta 2007;455(1):46–9.
[30] Ruoho M, Valset K, Finstad T, Tittonen I. Measurement of thin film thermal conductivity using the laser flash method. Nanotechnology 2015;26(19):195706.
[31] Parker WJ, Jenkins RJ, Butler CP, Abbott GL. Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity. J Appl Phys 1961;32(9):1679–84.
[32] Cahill DG, Fischer HE, Klitsner T, Swartz ET, Pohl RO. Thermal conductivity of thin films: measurements and understanding. J Vac Sci Technol A 1989;7(3):1259–66.
[33] Völklein F, Reith H, Meier A. Measuring methods for the investigation of in-plane and cross-plane thermal conductivity of thin films. Phys Status Solidi A 2013;210(1):106–18.
[34] Volklein F, Starz T. Thermal conductivity of thin films-experimental methods and theoretical interpretation. In: Proceedings of the ICT'97, XVI ICT '97 16th international conference on thermoelectrics (Cat. No.97TH8291; 1997. p. 711–8.
[35] Mishra V, Hardin CL, Garay JE, Dames C. A 3 omega method to measure an arbitrary anisotropic thermal conductivity tensor. Rev Sci Instrum 2015;86(5):054902.
[36] Dames C. Measuring the thermal conductivity of thin films: 3 omega and related electrothermal methods. Annu Rev Heat Transf 2013;16:7–49.
[37] Cahill DG. Thermal conductivity measurement from 30 to 750 K: the 3ω method. Rev Sci Instrum 1990;61(2):802–8.
[38] Lee JU, Yoon D, Kim H, Lee SW, Cheong H. Thermal conductivity of suspended pristine graphene measured by Raman spectroscopy. Phys Rev B 2011;83(8):081419.
[39] Balandin AA, Ghosh S, Bao W, Calizo I, Teweldebrhan D, Miao F, Lau CN. Superior thermal conductivity of single-layer graphene. Nano Lett 2008;8(3):902–7.
[40] Chen S, Moore AL, Cai W, Suk JW, An J, Mishra C, Amos C, Magnuson CW, Kang J, Shi L, Ruoff RS. Raman measurements of thermal transport in suspended monolayer graphene of variable sizes in vacuum and gaseous environments. ACS Nano 2011;5(1):321–8.
[41] Capinski WS, Maris HJ, Ruf T, Cardona M, Ploog K, Katzer DS. Thermal-conductivity measurements of GaAs/AlAs superlattices using a picosecond optical pump-and-probe technique. Phys Rev B 1999;59(12):8105–13.
[42] Malen JA, Baheti K, Tong T, Zhao Y, Hudgings JA, Majumdar A. Optical measurement of thermal conductivity using fiber aligned frequency domain thermoreflectance. J Heat Transf 2011;133(8):081601.
[43] Schmidt AJ, Cheaito R, Chiesa M. Characterization of thin metal films via frequency-domain thermoreflectance. J Appl Phys 2010;107(2):024908.
[44] Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. NPJ Comput Mater 2017;3(1):1–13.
[45] Ward L, Wolverton C. Atomistic calculations and materials informatics: a review. Curr Opin Solid State Mater Sci 2017;21(3):167–76.
[46] Hill J, Mulholland G, Persson K, Seshadri R, Wolverton C, Meredig B. Materials science with large-scale data and informatics: unlocking new opportunities. MRS Bull 2016;41(5):399–409.

[47] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature 2018;559(7715):547–55.
[48] Schleder GR, Padilha ACM, Acosta CM, Costa M, Fazzio A. From DFT to machine learning: recent approaches to materials science–a review. J Phys Mater 2019;2 (3):032001.
[49] Ramakrishna S, Zhang TY, Lu WC, Qian Q, Low JSC, Yune JHR, Tan DZL, Bressan S, Sanvito S, Kalidindi SR. Materials informatics. J Intell Manuf 2018;30 (6):2307–26.
[50] Juneja R, Yumnam G, Satsangi S, Singh AK. Coupling the high-throughput property map to machine learning for predicting lattice thermal conductivity. Chem Mater 2019;31(14):5145–51.
[51] Wang X, Zeng S, Wang Z, Ni J. Identification of crystalline materials with ultra-low thermal conductivity based on machine learning study. J Phys Chem C 2020; 124(16):8488–95.
[52] Chen L, Tran H, Batra R, Kim C, Ramprasad R. Machine learning models for the lattice thermal conductivity prediction of inorganic materials. Comput Mater Sci 2019;170:109155.
[53] Chakraborty P, Liu Y, Ma T, Guo X, Cao L, Hu R, Wang Y. Quenching thermal transport in aperiodic superlattices: a molecular dynamics and machine learning study. ACS Appl Mater Interfaces 2020;12(7):8795–804.
[54] Visaria D, Jain A. Machine-learning-assisted space-transformation accelerates discovery of high thermal conductivity alloys. Appl Phys Lett 2020;117(20): 202107.
[55] Ju S, Shiga T, Feng L, Hou Z, Tsuda K, Shiomi J. Designing nanostructures for phonon transport via bayesian optimization. Phys Rev X 2017;7(2):021024.
[56] Roy Chowdhury P, Reynolds C, Garrett A, Feng T, Adiga SP, Ruan X. Machine learning maximized Anderson localization of phonons in aperiodic superlattices. Nano Energy 2020;69:104428.
[57] Ju S, Shiomi J. Materials informatics for heat transfer: recent progresses and perspectives. Nanoscale Microscale Thermophys Eng 2019;23(2):157–72.
[58] Wan X, Feng W, Wang Y, Wang H, Zhang X, Deng C, Yang N. Materials discovery and properties prediction in thermal transport via materials informatics: a mini review. Nano Lett 2019;19(6):3387–95.
[59] Hang Z, Hippalgaonkar K, Buonassisi T, Løvvik OM, Sagvolden E, Ding D. Machine learning for novel thermal-materials discovery: early successes, opportunities, and challenges. ES Energy Environ 2018;2(2):1–8.
[60] Hong Y, Han D, Hou B, Wang X, Zhang J. High-throughput computations of cross-plane thermal conductivity in multilayer stanene. Int J Heat Mass Transf 2021; 171:121073.
[61] Zhan T, Fang L, Xu Y. Prediction of thermal boundary resistance by the machine learning method. Sci Rep 2017;7(1):7109.
[62] Yang H, Zhang Z, Zhang J, Zeng XC. Machine learning and artificial neural network prediction of interfacial thermal resistance between graphene and hexagonal boron nitride. Nanoscale 2018;10(40):19092–9.
[63] Furmanchuk AO, Saal JE, Doak JW, Olson GB, Choudhary A, Agrawal A. Prediction of seebeck coefficient for compounds without restriction to fixed stoichiometry: a machine learning approach. J Comput Chem 2018;39(4): 191–202.
[64] Ju S, Shimizu S, Shiomi J. Designing thermal functional materials by coupling thermal transport calculations and machine learning. J Appl Phys 2020;128(16): 161102.
[65] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson KA. Commentary: the materials project: a materials genome approach to accelerating materials innovation. APL Mater 2013;1(1): 011002.
[66] Curtarolo S, Setyawan W, Wang S, Xue J, Yang K, Taylor RH, Nelson LJ, Hart GLW, Sanvito S, Buongiorno-Nardelli M, Mingo N, Levy O. AFLOWLIB.ORG: a distributed materials properties repository from high-throughput ab initio calculations. Comput Mater Sci 2012;58:227–35.
[67] Hellenbrandt M. The inorganic crystal structure database (ICSD)—present and future. Crystallogr Rev 2004;10(1):17–22.
[68] Saal JE, Kirklin S, Aykol M, Meredig B, Wolverton C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD). JOM 2013;65(11):1501–9.
[69] Otsuka S, Kuwajima I, Hosoya J, Xu Y, Yamazaki M. PoLyInfo: polymer database for polymeric materials design. 2011. In: Proceedings of the international conference on emerging intelligent data and web technologies; 2011. p. 22–9.
[70] Ghiringhelli LM, Vybiral J, Levchenko SV, Draxl C, Scheffler M. Big data of materials science: critical role of the descriptor. Phys Rev Lett 2015;114(10): 105503.
[71] Chauhan NK, Singh K. A review on conventional machine learning vs deep learning. 2018. In: Proceedings of the international conference on computing, power and communication technologies (GUCON); 2018. p. 347–52.
[72] Yosipof A, Shimanovich K, Senderowitz H. Materials informatics: statistical modeling in material science. Mol Inform 2016;35(11–12):568–79.
[73] Rickman JM, Lookman T, Kalinin SV. Materials informatics: from the atomic-level to the continuum. Acta Mater 2019;168:473–510.
[74] Kauwe SK, Graser J, Vazquez A, Sparks TD. Machine learning prediction of heat capacity for solid inorganics. Integr Mater Manuf Innov 2018;7(2):43–51.
[75] Wei H, Zhao S, Rong Q, Bao H. Predicting the effective thermal conductivities of composite materials and porous media by machine learning methods. Int J Heat Mass Transf 2018;127:908–16.
[76] Bang HT, Yoon S, Jeon H. Application of machine learning methods to predict a thermal conductivity model for compacted bentonite. Ann Nucl Energy 2020;142: 107395.
[77] Wang Z, Huang S, Wen G, Liu Q, Tang P. Thermal conductivity prediction and structure-property relationship of CaO-SiO_2-Al_2O_3 ternary system: a combination of molecular dynamics simulations and machine learning. J Mol Liq 2020: 114697.
[78] Loftis C, Yuan K, Zhao Y, Hu M, Hu J. Lattice thermal conductivity prediction using symbolic regression and machine learning. J Phys Chem A 2021;125(1): 435–50.
[79] Zhu MX, Song HG, Yu QC, Chen JM, Zhang HY. Machine-learning-driven discovery of polymers molecular structures with high thermal conductivity. Int J Heat Mass Transf 2020;162:120381.
[80] Dan Y, Zhao Y, Li X, Li S, Hu M, Hu J. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. NPJ Comput Mater 2020;6(1):84.
[81] Yang L, Zhang D, Karniadakis GE. Physics-informed generativeadversarial networks for stochastic differential equations. SIAM Journal on Scientific Computing 2020;42(1):A292–317.
[82] Qian X, Yang R. Machine learning for predicting thermal transport properties of solids. Mater Sci Eng R Rep 2021;146:100642.
[83] Cavazzuti M. Optimization methods. Berlin Heidelberg: Springer; 2013. Berlin, Heidelberg.
[84] Colaço MJ, Orlande HRB, Dulikravich GS. Inverse and optimization problems in heat transfer. J Braz Soc Mech Sci Eng 2006;28(1):1–24.
[85] Hu R, Iwamoto S, Feng L, Ju S, Hu S, Ohnishi M, Nagai N, Hirakawa K, Shiomi J. Machine-learning-optimized aperiodic superlattice minimizes coherent phonon heat conduction. Phys Rev X 2020;10(2):021050.
[86] Chibante R. Simulated annealing: theory with applications. Sciyo. 2010.
[87] Zhang Y, Wang S, Ji G. A comprehensive survey on particle swarm optimization algorithm and its applications. Math Probl Eng 2015;2015:e931256.
[88] Back T. Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. Oxford University Press; 1996.
[89] Jennings PC, Lysgaard S, Hummelshøj JS, Vegge T, Bligaard T. Genetic algorithms for computational materials discovery accelerated by machine learning. NPJ Comput Mater 2019;5(1):46.
[90] Snoek J, Larochelle H, Adams RP. Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems 2012:25.
[91] Browne CB, Powley E, Whitehouse D, Lucas SM, Cowling PI, Rohlfshagen P, Tavener S, Perez D, Samothrakis S, Colton S. A survey of Monte Carlo tree search methods. IEEE Trans Comput Intell AI Games 2012;4(1):1–43.
[92] Rong Q, Wei H, Huang X, Bao H. Predicting the effective thermal conductivity of composites from cross sections images using deep learning methods. Compos Sci Technol 2019;184:107861.
[93] Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. NPJ Comput Mater 2018;4(1):25.
[94] Bao H, Wei H, Zhao S. An improved convolution neural network model to predict the effective thermal conductivity of composite materials. In: Proceedings of the international heat transfer conference16; 2018. p. 6891–901.
[95] Bedolla E, Padierna LC, Castañeda-Priego R. Machine learning for condensed matter physics. J Phys Condens Matter 2020;33(5):053001.
[96] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L, Zdeborová L. Machine learning and the physical sciences. Rev Mod Phys 2019;91 (4):045002.
[97] Oliynyk AO, Buriak JM. Virtual issue on machine-learning discoveries in materials science. Chem Mater 2019;31(20):8243–7.
[98] Ray S. A quick review of machine learning algorithms. International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). IEEE; 2019. p. 35–9.
[99] Kalidindi SR, Medford AJ, McDowell DL. Vision for data and informatics in the future materials innovation ecosystem. JOM 2016;68(8):2126–37.
[100] Song Z, Chen X, Meng F, Cheng G, Wang C, Sun Z, Yin WJ. Machine learning in materials design: algorithm and application. Chin Phys B 2020;29(11):116103.
[101] Seber GAF, Lee AJ. Linear regression analysis. John Wiley & Sons; 2012.
[102] Menard S. Applied logistic regression analysis. Sage; 2002.
[103] Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004;14(3):199–222.
[104] Bishop CM. Neural networks for pattern recognition. Oxford University Press; 1995.
[105] Wu S, Kondo Y, Kakimoto MA, Yang B, Yamada H, Kuwajima I, Lambard G, Hongo K, Xu Y, Shiomi J, Schick C, Morikawa J, Yoshida R. Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. NPJ Comput Mater 2019;5(1):1–11.
[106] Carrete J, Li W, Mingo N, Wang S, Curtarolo S. Finding unprecedentedly low-thermal-conductivity half-heusler semiconductors via high-throughput materials modeling. Phys Rev X 2014;4(1):011019.
[107] Toher C, Plata JJ, Levy O, de Jong M, Asta M, Nardelli MB, Curtarolo S. High-throughput computational screening of thermal conductivity, Debye temperature, and Gr\"uneisen parameter using a quasiharmonic Debye model. Phys Rev B 2014;90(17):174107.
[108] van Roekeghem A, Carrete J, Oses C, Curtarolo S, Mingo N. High-throughput computation of thermal conductivity of high-temperature solid phases: the case of oxide and fluoride perovskites. Phys Rev X 2016;6(4):041061.
[109] Xia Y, Hegde VI, Pal K, Hua X, Gaines D, Patel S, He J, Aykol M, Wolverton C. High-throughput study of lattice thermal conductivity in binary rocksalt and zinc blende compounds including higher-order anharmonicity. Phys Rev X 2020;10 (4):041029.

[110] Tewari A, Dixit S, Sahni N, Bordas SPA. Machine learning approaches to identify and design low thermal conductivity oxides for thermoelectric applications. Data Centric Eng 2020;1.
[111] Curtarolo S, Hart GLW, Nardelli MB, Mingo N, Sanvito S, Levy O. The high-throughput highway to computational materials design. Nat Mater 2013;12(3):191–201.
[112] Qian X, Peng S, Li X, Wei Y, Yang R. Thermal conductivity modeling using machine learning potentials: application to crystalline and amorphous silicon. Mater Today Phys 2019;10:100140.
[113] Chan H, Sasikumar K, Srinivasan S, Cherukara M, Narayanan B, Sankaranarayanan SKRS. Machine learning a bond order potential model to study thermal transport in WSe_2 nanostructures. Nanoscale 2019;11(21):10381–92.
[114] Zhang C, Sun Q. Gaussian approximation potential for studying the thermal conductivity of silicene. J Appl Phys 2019;126(10):105103.
[115] Zhang Y, Shen C, Long T, Zhang H. Thermal conductivity of h-BN monolayers using machine learning interatomic potential. J Phys Condens Matter 2020;33(10):105903.
[116] Sosso GC, Deringer VL, Elliott SR, Csányi G. Understanding the thermal properties of amorphous solids using machine-learning-based interatomic potentials. Mol Simul 2018;44(11):866–80.
[117] Mortazavi B, Podryabinkin EV, Novikov IS, Roche S, Rabczuk T, Zhuang X, Shapeev AV. Efficient machine-learning based interatomic potentials for exploring thermal conductivity in two-dimensional materials. J Phys Mater 2020;3(2):02LT02.
[118] Korotaev P, Novoselov I, Yanilkin A, Shapeev A. Accessing thermal conductivity of complex compounds by machine learning interatomic potentials. Phys Rev B 2019;100(14):144308.
[119] Xie H, Yan J, Gu X, Bao H. A scattering rate model for accelerated evaluation of lattice thermal conductivity bypassing anharmonic force constants. J Appl Phys 2019;125(20):205104.
[120] Cheng T, Jaramillo-Botero A, An Q, Ilyin DV, Naserifar S, Goddard WA. First principles-based multiscale atomistic methods for input into first principles nonequilibrium transport across interfaces. Proc Natl Acad Sci 2019;116(37):18193–201.
[121] Jia W, Wang H, Chen M, Lu D, Lin L, Car R, Weinan E, Zhang L. Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning. In: Proceedings of the SC20: international conference for high performance computing, networking, storage and analysis. IEEE; 2020. p. 1–14.
[122] Behler J. Perspective: machine learning potentials for atomistic simulations. J Chem Phys 2016;145(17):170901.
[123] Handley CM, Behler J. Next generation interatomic potentials for condensed systems. Eur Phys J B 2014;87(7):152.
[124] Liu YB, Yang JY, Xin GM, Liu LH, Csányi G, Cao BY. Machine learning interatomic potential developed for molecular simulations on thermal properties of β-Ga_2O_3. J Chem Phys 2020;153(14):144501.
[125] Dai FZ, Wen B, Sun Y, Xiang H, Zhou Y. Theoretical prediction on thermal and mechanical properties of high entropy $(Zr_{0.2}Hf_{0.2}Ti_{0.2}Nb_{0.2}Ta_{0.2})C$ by deep learning potential. J Mater Sci Technol 2020;43:168–74.
[126] Gu X, Zhao CY. Thermal conductivity of single-layer $MoS_{2(1-x)}Se_{2x}$ alloys from molecular dynamics simulations with a machine-learning-based interatomic potential. Comput Mater Sci 2019;165:74–81.
[127] Mortazavi B, Podryabinkin EV, Roche S, Rabczuk T, Zhuang X, Shapeev AV. Machine-learning interatomic potentials enable first-principles multiscale modeling of lattice thermal conductivity in graphene/borophene heterostructures. Mater Horiz 2020;7(9):2359–67.
[128] Behler J. Constructing high-dimensional neural network potentials: a tutorial review. Int J Quantum Chem 2015;115(16):1032–50.
[129] Bartók AP, Payne MC, Kondor R, Csányi G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. Phys Rev Lett 2010;104(13):136403.
[130] Thompson AP, Swiler LP, Trott CR, Foiles SM, Tucker GJ. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. J Comput Phys 2015;285:316–30.
[131] Li R, Lee E, Luo T. A unified deep neural network potential capable of predicting thermal conductivity of silicon in different phases. Mater Today Phys 2020;12:100181.
[132] Rodriguez A, Liu Y, Hu M. Spatial density neural network force fields with first-principles level accuracy and application to thermal transport. Phys Rev B 2020;102(3):035203.
[133] Plimpton S. Fast parallel algorithms for short-range molecular dynamics. J Comput Phys 1995;117(1):1–19.
[134] Fan Z, Chen W, Vierimaa V, Harju A. Efficient molecular dynamics simulations with many-body potentials on graphics processing units. Comput Phys Commun 2017;218:10–6.
[135] T. Lookman, F.J. Alexander Lookman T, Balachandran PV, Xue D, Pilania G, Shearman T, Theiler J, Gubernatis JE, Hogden J, Barros K, BenNaim E, Alexander FJ, Rajan K. A perspective on materials informatics: state-of-the-art and challenges. Information science for materials discovery and design. Cham: Springer International Publishing; 2016. p. 3–12.
[136] Lookman T, Alexander FJ, Rajan K. Information science for materials discovery and design. Springer; 2016.
[137] Liu Y, Guo B, Zou X, Li Y, Shi S. Machine learning assisted materials design and discovery for rechargeable batteries. Energy Storage Mater 2020;31:434–50.
[138] Liu Y, Zhao T, Ju W, Shi S. Materials discovery and design using machine learning. J Mater 2017;3(3):159–77.
[139] Gubernatis JE, Lookman T. Machine learning in materials design and discovery: examples from the present and suggestions for the future. Phys Rev Mater 2018;2(12):120301.
[140] Zhou T, Song Z, Sundmacher K. Big data creates new opportunities for materials research: a review on methods and applications of machine learning for materials design. Engineering 2019;5(6):1017–26.
[141] Hansson J, Nilsson TMJ, Ye L, Liu J. Novel nanostructured thermal interface materials: a review. Int Mater Rev 2018;63(1):22–45.
[142] Pichanusakorn P, Bandaru P. Nanostructured thermoelectrics. Mater Sci Eng R Rep 2010;67(2):19–63.
[143] Yang J, Xi L, Qiu W, Wu L, Shi X, Chen L, Yang J, Zhang W, Uher C, Singh DJ. On the tuning of electrical and thermal transport in thermoelectrics: an integrated theory–experiment perspective. NPJ Comput Mater 2016;2(1):1–17.
[144] Cahill DG, Ford WK, Goodson KE, Mahan GD, Majumdar A, Maris HJ, Merlin R, Phillpot SR. Nanoscale thermal transport. J Appl Phys 2003;93(2):793–818.
[145] Yan J, Wei H, Xie H, Gu X, Bao H. Seeking for low thermal conductivity atomic configurations in SiGe alloys with bayesian optimization. ES Energy Environ 2020;8(4):56–64.
[146] Kitai K, Guo J, Ju S, Tanaka S, Tsuda K, Shiomi J, Tamura R. Designing metamaterials with quantum annealing and factorization machines. Phys Rev Res 2020;2(1):013319.
[147] Sha W, Xiao M, Zhang J, Ren X, Zhu Z, Zhang Y, Xu G, Li H, Liu X, Chen X. Robustly printable freeform thermal metamaterials. Nat Commun 2021;12(1):1–8.
[148] Hu R, Huang S, Wang M, Luo X, Shiomi J, Qiu CW. Encrypted thermal printing with regionalization transformation. Adv Mater 2019;31(25):1807849.
[149] Hu R, Zhou S, Li Y, Lei DY, Luo X, Qiu CW. Illusion thermotics. Adv Mater 2018;30(22):1707237.
[150] Roy Chowdhury P, Ruan X. Unexpected thermal conductivity enhancement in aperiodic superlattices discovered using active machine learning. NPJ Comput Mater 2022;8(1):12.
[151] Wan X, Ma D, Pan D, Yang L, Yang N. Optimizing thermal transport in graphene nanoribbon based on phonon resonance hybridization. Mater Today Phys 2021;20:100445.
[152] Wan J, Jiang JW, Park HS. Machine learning-based design of porous graphene with low thermal conductivity. Carbon N Y 2020;157:262–9.
[153] Wei H, Bao H, Ruan X. Genetic algorithm-driven discovery of unexpected thermal conductivity enhancement by disorder. Nano Energy 2020;71:104619.
[154] Wang Y, Huang H, Ruan X. Decomposition of coherent and incoherent phonon conduction in superlattices and random multilayers. Phys Rev B 2014;90(16):165406.
[155] Dieb TM, Ju S, Yoshizoe K, Hou Z, Shiomi J, Tsuda K. MDTS: automatic complex materials design using Monte Carlo tree search. Sci Technol Adv Mater 2017;18:498–503.
[156] Liu Y, Hu R, Wang Y, Ma J, Yang Z, Luo X. Big-data-accelerated aperiodic Si/Ge superlattice prediction for quenching thermal conduction via pattern analysis. Energy AI 2021;3:100046.
[157] Hu S, Zhang Z, Jiang P, Chen J, Volz S, Nomura M, Li B. Randomness-induced phonon localization in Graphene heat conduction. J Phys Chem Lett 2018;9(14):3959–68.
[158] Luckyanova MN, Mendoza J, Lu H, Song B, Huang S, Zhou J, Li M, Dong Y, Zhou H, Garlow J, Wu L, Kirby BJ, Grutter AJ, Puretzky AA, Zhu Y, Dresselhaus MS, Gossard A, Chen G. Phonon localization in heat conduction. Sci Adv 2018;4(12):eaat9460.
[159] Cloutier SG, Guico RS, Xu JM. Phonon localization in periodic uniaxially nanostructured silicon. Appl Phys Lett 2005;87(22):222104.
[160] Williams ML, Maris HJ. Numerical study of phonon localization in disordered systems. Phys Rev B 1985;31(7):4508–15.
[161] Yamawaki M, Ohnishi M, Ju S, Shiomi J. Multifunctional structural design of graphene thermoelectrics by Bayesian optimization. Sci Adv 2018;4(6):eaar4192.
[162] Cui C, Ouyang T, Tang C, He C, Li J, Zhang C, Zhong J. Bayesian optimization-based design of defect gamma-graphyne nanoribbons with high thermoelectric conversion efficiency. Carbon N Y 2021;176:52–60.
[163] Sakurai A, Yada K, Simomura T, Ju S, Kashiwagi M, Okada H, Nagao T, Tsuda K, Shiomi J. Ultranarrow-band wavelength-selective thermal emission with aperiodic multilayered metamaterials designed by bayesian optimization. ACS Cent Sci 2019;5(2):319–26.
[164] Guo J, Ju S, Shiomi J. Design of a highly selective radiative cooling structure accelerated by materials informatics. Opt Lett OL 2020;45(2):343–6.
[165] Hu R, Song J, Liu Y, Xi W, Zhao Y, Yu X, Cheng Q, Tao G, Luo X. Machine learning-optimized Tamm emitter for high-performance thermophotovoltaic system with detailed balance analysis. Nano Energy 2020;72:104687.
[166] Xi W, Liu Y, Song J, Hu R, Luo X. High-throughput screening of a high-Q mid-infrared Tamm emitter by material informatics. Opt Lett OL 2021;46(4):888–91.
[167] Dieb TM, Ju S, Shiomi J, Tsuda K. Monte Carlo tree search for materials design and discovery. MRC 2019;9(2):532–6.
[168] Ma W, Liu Z, Kudyshev ZA, Boltasseva A, Cai W, Liu Y. Deep learning for the design of photonic structures. Nat Photon 2021;15(2):77–90.
[169] Ma W, Cheng F, Liu Y. Deep-learning-enabled on-demand design of chiral metamaterials. ACS Nano 2018;12(6):6326–34.
[170] Liu D, Tan Y, Khoram E, Yu Z. Training deep neural networks for the inverse design of nanophotonic structures. ACS Photonics 2018;5(4):1365–9.
[171] Ramchoun H, Ghanou Y, Ettaouil M, Idrissi MAJ. Multilayer perceptron: architecture optimization and training. Int J Interact Multimed Artif Intell 2016;4:26–30 (Special Issue on Artificial Intelligence Underpinning).

[172] Xie T, Grossman JC. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. Phys Rev Lett 2018;120(14):145301.
[173] Zhou Q, Yang C, Liang A, Zheng X, Chen Z. Low computationally complex recurrent neural network for high speed optical fiber transmission. Opt Commun 2019;441:121–6.
[174] Kingma DP, Welling M. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 2013.
[175] Forrester AIJ, Sóbester A, Keane AJ. Multi-fidelity optimization via surrogate modelling. Proc R Soc A 2007;463(2088):3251–69.
[176] Fernández-Godino MG, Park C, Kim NH, Haftka RT. Review of multi-fidelity models. arXiv preprint arXiv:1609.07196 2016.
[177] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2010;22(10):1345–59.
[178] Kailkhura B, Gallagher B, Kim S, Hiszpanski A, Han TYJ. Reliable and explainable machine-learning methods for accelerated material discovery. NPJ Comput Mater 2019;5(1):1–9.
[179] Farimani AB, Gomes J, Pande VS. Deep learning the physics of transport phenomena. arXiv preprint arXiv:1709.02432 2017.
[180] Li R, Lee E, Luo T. Physics-informed neural networks for solving multiscale mode-resolved phonon Boltzmann transport equation. Mater Today Phys 2021;19:100429.
[181] Jin X, Cai S, Li H, Karniadakis GE. NSFnets (Navier-stokes flow nets): physics-informed neural networks for the incompressible Navier-stokes equations. J Comput Phys 2021;426:109951.
[182] Shahdoosti HR, Ghassemian H. Combining the spectral PCA and spatial PCA fusion methods by an optimal filter. Inf Fusion 2016;27:150–60.
[183] Wei H, Bao H, Ruan X. Machine learning prediction of thermal transport in porous media with physics-based descriptors. Int J Heat Mass Transf 2020;160:120176.
[184] Liu Y, Hu Z, Suo Z, Hu L, Feng L, Gong X, Liu Y, Zhang J. High-throughput experiments facilitate materials innovation: a review. Sci China Technol Sci 2019;62(4):521–45.
[185] Hattrick-Simpers JR, Gregoire JM, Kusne AG. Perspective: composition–structure–property mapping in high-throughput experiments: turning data into knowledge. APL Mater 2016;4(5):053211.
[186] Lenzerini M. Data integration: a theoretical perspective. Association for Computing Machinery; 2002. p. 233–46.
[187] Pokhriyal N, Jacques DC. Combining disparate data sources for improved poverty prediction and mapping. Proc Natl Acad Sci 2017;114(46):E9783–92.
[188] Raman CV. A new radiation. Indian J Phys 1928;2:387–98.