Energy and AI 3 (2021) 100046

Contents lists available at ScienceDirect

Energy and AI

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

Big-data-accelerated aperiodic Si/Ge superlattice prediction for quenching thermal conduction via pattern analysis

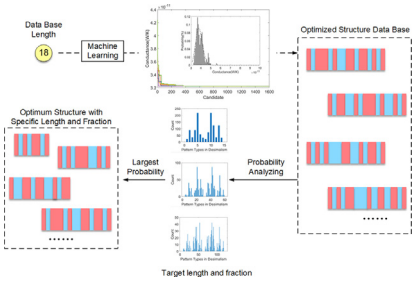
Yida Liu [a], Run Hu [a,*], Yan Wang [b,*], Jinglong Ma [a], Zhangcan Yang [a], Xiaobing Luo [a]

[a] School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[b] Department of Mechanical Engineering, University of Nevada, Reno, NV 89557, USA

HIGHLIGHTS

- We propose a method to accelerate the optimization of quenching the thermal conductance of aperiodic superlattice structure via pattern analysis.
- The thermal conductivity of the optimized structure in pattern analysis is almost the same as that in machine learning.
- The proposed pattern analysis can effectively reduce the computational time consumed in the multi-variable problem of machine learning.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 November 2020
Received in revised form 20 December 2020
Accepted 21 December 2020
Available online 30 December 2020

Keywords:
Aperiodic superlattice
Pattern analysis
Thermal conductivity
Machine learning
Atomic Green's function

ABSTRACT

Thermal conductivity of material is one of the basic physical properties and plays an important role in manipulating thermal energy. In order to accelerate the prediction of material structure with desired thermal property, machine learning algorithm has been widely adopted. However, in the optimization of multivariable material structure such as different lengths or proportions, the machine learning algorithm may be required to be reconducted again and again for each variable, which will consume a lot of computing resources. Recently, it has been found that the thermal conductivity of aperiodic superlattices is closely related to the degree of the structural randomness, which can also be reflected in their local pattern structures. Inspired by these, we propose a new pattern analysis method, in which machine learning only needs to be carried out for one time, and through which the optimal structure of different variables with low thermal conductivity can be obtained. To verify the method, we compare the thermal conductivities of the structure obtained by pattern analysis, conventional machine learning, and previous literature, respectively. The pattern analysis method is validated to greatly reduce the prediction time of multivariable structure with high enough accuracy and may promote further development of material informatics.

1. Introduction

Thermal conductivity (TC) is one of the basic physical properties of materials and plays an important role in manipulating thermal energy, as highlighted by high-TC materials for thermal management of high-power electronics and low-TC materials for thermal barriers and thermoelectric applications [1,2]. Stimulated by the profound and wide applications of thermal materials, the design, fabrication, and optimization materials with desired TC has attracted tremendous attentions and has made significant progress in the last two decades, especially on the micro- and nanoscale heat transport. For non-metal solids, the quanta of heat, i.e. phonons, like electrons and photons, also possess particle-wave duality, and thereby a plenty of theoretical tools has been successively proposed to characterize and quantify the particle-like (incoherent) and wave-like (coherent) phonon transport properties, like Boltzmann transport equation (BTE), molecular dynamics, Monte Carlo simulations, atomistic Green's function (AGF), etc. To manipulate the phonon transport properties, there are mainly two categories of strategies to tune the thermal conductivity of materials. The first is to manipulate phonon scatterings by engineering nanoparticles [3], impurities [4], roughness [5], and defects [6] into materials. However, the cou-

* Corresponding authors.
E-mail addresses: hurun@hust.edu.cn (R. Hu), yanwang@unr.edu (Y. Wang).

https://doi.org/10.1016/j.egyai.2020.100046

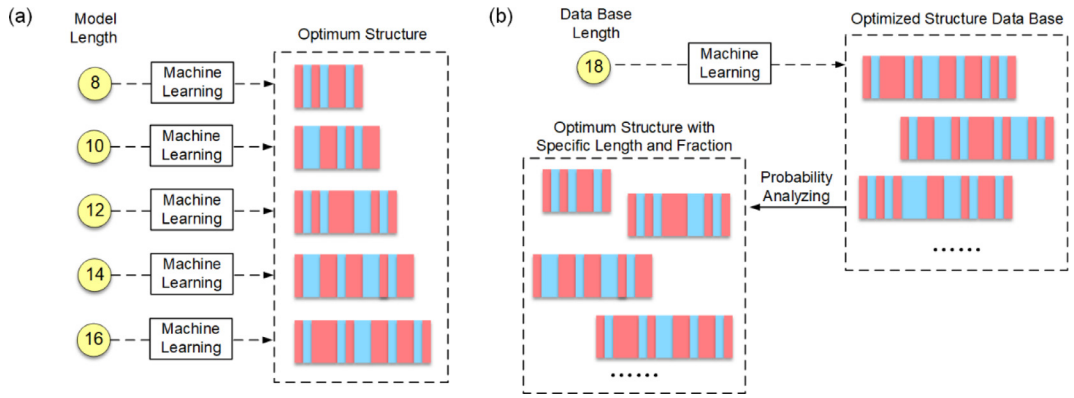


Fig. 1. (a) Illustration of the repeated machine learning to achieve optimum superlattice structures with different lengths. (b) Illustration of the proposed pattern analysis to achieve optimum superlattice structures with different lengths by once machine learning process.

pling between these factors often makes it difficult to identify the effect of each individual factor on the overall heat transfer process. The other strategy focuses on manipulating phonon spectrum and scattering by creating superlattices [7,8] and nanoscale phononic crystals [9]. These design methods introduce multiple interfaces into the material, which makes it very challenging to find the optimal structure. In 2014, Wang et al. observed greatly reduced thermal conductivity of conceptual Lennard-Jones aperiodic superlattices compared to their periodic counterparts, which was attributed to the localization of coherent phonons [10]. In 2015, Qiu et al. compared the influence of aperiodicity on thermal conductivities of Si/Ge superlattices and found that the thermal conductivity was greatly reduced in aperiodic Si/Ge superlattices [11]. These studies have revealed the great influence of aperiodicity on thermal conductivity of superlattices. Nonetheless, there still lacks a rigorous analysis on the optimization of aperiodic superlattice structure, because there are a formidably large number of possible aperiodic superlattice structures, necessitating extensive computational time and resources to calculate the thermal conductivity of each candidate. In this regard, some researchers tactfully combined the machine learning algorithm [12] with thermal conductivity calculation under the framework of material informatics [13] to accelerate the superlattice structure optimization [14]. Notably, Ju et al. in 2017 optimized the structure of Si/Ge superlattices to achieve minimum thermal conductance at different lengths via AGF and Bayesian optimization [15]. Although machine learning can greatly shorten the time to find the optimal structure, it needs to be conducted repeatedly for each variable for a multivariable problem; thus, the accumulated computational cost is still huge. For instance, as shown in Fig. 1a, for a superlattice with length of 8, we can do machine learning to find its optimum structure. When changing the superlattice length to 10, 12, 14, or 16, we have to restart a new machine learning process to find the corresponding optimum structures. It is implied that machine learning optimization must be conducted for many times separately to obtain the optimum structure of each length, which is still time-consuming and computationally expensive. How to improve the computational efficiency of such multivariable optimization has become an urgent problem to be solved. In 2020, Chakraborty et al. revealed the relationship between the randomness degree of the aperiodic superlattice and the thermal conductivity based on the swapping method, and defined the randomness parameters which have been proved to be the effective structural parameters to predict thermal conductivity in the process of machine learning [16]. Later, Hu et al. minimized the coherent phonon heat conduction in GaAs/AlAs aperiodic superlattice by machine-learning optimization and experimental validation and revealed the relationship between the optimal structure and its local pattern structure [17].

The above works well proves that the feature of low thermal conductivity of the aperiodic superlattice is closely related to the degree of its structural randomness, which also can be reflected in its local pattern structure. In other words, the process of optimizing the superlattice with low thermal conductivity as the goal is the process of enhancing its structural randomness. Therefore, the short disordered local pattern structures, which is extracted from the long disordered structures in some way, should also be superior for searching for ultra-low thermal conductivity. Based on these, we propose a pattern analysis optimization method which is of great potential in overcoming the multivariable issue in machine learning. The general process of the optimization method with superlattice is illustrated in Fig. 1b. Firstly, a long superlattice, say length of 18, is optimized by machine learning. The superlattice sequences with ascending low thermal conductivity of the optimization results are recorded to form a big database. Then, to find a relatively short-length optimized superlattice, we statistically count the appearance number of target length or target fraction sequences in the database and predict the optimal superlattice structures with the largest probability. To simply validate this method, we use AGF method to calculate the thermal conductance of a one-dimensional (1D) diatomic chain model to generate the long-length superlattice database and predict the optimal short-length superlattices, which is then validated by comparing with the optimized structure in machine learning and previous literature [15]. Finally, the same comparison is also carried out in the three-dimensional (3D) system based on molecular dynamics (MD) simulations, which demonstrates the feasibility of pattern analysis in the actual system. The proposed big-data-accelerated pattern analysis is validated to effectively reduce the computational time consumed in the multivariable problem of machine learning, which enables the prediction of more complex and better-performing structures of a specific material, and opens a new door for the further development of material informatics.

2. Methodology

2.1. Atomistic Green's function

In AGF method [18], the phonon transmission function is defined as:

$$\Xi(\omega) = \mathrm{Tr}\left[\Gamma_1 G_I \Gamma_2 G_I^H\right] \tag{1}$$

where ω represents the phonon frequency, G_I is the interfacial region subset of the overall Green's function and G_I^H is the conjugate transpose of G_I. The $\Gamma_1 = i(\sum_1 - \sum_1^H)$ and $\Gamma_2 = i(\sum_2 - \sum_2^H)$ are the "escape rate" for phonons leaking into the interfacial region from the left lead and right lead respectively [19], and $\sum_1$ and $\sum_2$ are the self-energy matrices associated with left lead and right lead respectively, which can be calculated from surface Green's function. Thus, the transmission of the full spectrum can be calculated.

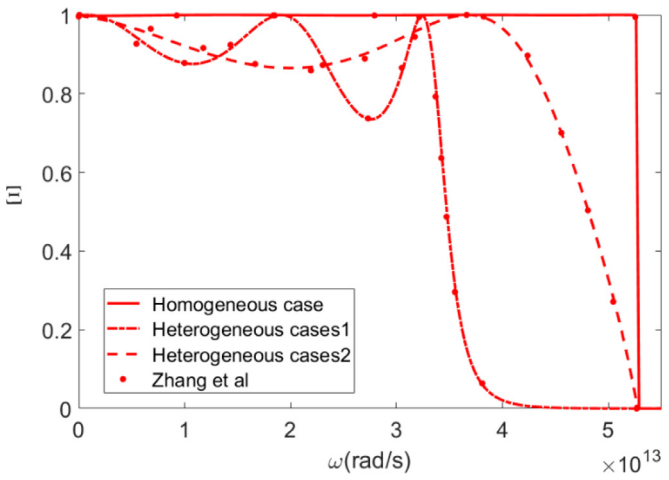


Fig. 2. Validation of the AGF calculations with the literature [21].

As the left and right lead set at different temperatures T_L and T_R, the net heat flow through the interfacial region can be given by the Landauer formula [20]:

$$J=\int_0^\infty \frac{\hbar\omega}{2\pi}\Xi(\omega)\left[f_{\mathrm{BE}}(\omega,T_L)-f_{\mathrm{BE}}(\omega,T_R)\right]d\omega \tag{2}$$

in which, $f_{\mathrm{BE}}(\omega,T_{L,R})=\{\exp[\hbar\omega/(k_BT_{L,R})]-1\}^{-1}$ is the Bose-Einstein distributions of phonons and the k_B represents the Boltzmann constant. For small temperature differences, the thermal conductance of the interface region at average temperature T can be obtained by the following formula:

$$\sigma=\int_0^\infty \frac{\hbar\omega}{2\pi}\Xi(\omega)\frac{\partial f_{\mathrm{BE}}}{\partial T}d\omega. \tag{3}$$

In order to verify the correctness of our code, we reproduce the model in the literature [21] and the validation results are shown in Fig. 2 where the dot represents the result in the literature, and the line represents the result of our calculation. It is obviously that our calculation is in good agreement with the results given by Zhang et al. [21].

2.2. Machine learning

During the machine learning process, we use the open-source python library MDTS [22] based on the Monte Carlo Tree Search (MCTS) method. The MCTS optimization process is performed by building a search tree in which the tree nodes can be built one by one based on evaluation of each simulated case. There are two vital information in each node: the estimated value of the simulation results and the number of times each node has been accessed. The MCTS process can be divided into the following steps: selection, expansion, simulation and backpropagation. In the first selection step, from the root node, the optimized child nodes are selected recursively based on the small upper confidence bound (UCB) score until the leaf node is reached. The UCB score is defined as:

$$u_i=\frac{V_i}{n_i}-b\sqrt{\frac{2\ln N_{parent}}{n_i}} \tag{4}$$

in which, V_i is the sum of the simulated values of all structures calculated from the node during all accesses, n_i is the number of times the node has been accessed, b is tunable bias parameter used to adjust the balance between tree exploration and exploitation, and N_{parent} represents the number of times its parents nodes have been accessed. In the second expansion step, a number of children nodes are added to the leaf node. In the third simulation step, a playout from the expanded nodes is selected randomly and then its thermal conductance will be calculated. In the last backpropagation step, the parameters V_i, n_i, N_{parent} of the nodes are updated along the path back to root.

2.3. Molecular dynamics

In the process of 3D model calculation, the MD simulations are performed using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [23]. The interatomic interactions of atoms in the superlattice are described by the Tersoff potential [24]. The time step is set as 1 fs. First, the whole system is equilibrated in the NPT ensemble at zero pressure with the temperature increasing from 100 K to 300 K in 50 ps and then maintained at 300 K for another 100 ps using Nose-Hoover thermostats [25]. Afterwards, based on the optimized structure, the system is relaxed in the NVT ensemble at 300 K for 1 ns using Langevin thermostats [26]. With all the optimizations done, the NEMD method is conducted in the system to calculate the thermal conductivity. In the NEMD process, the total system is simulated in the NVE ensemble except the fixed ends for 5 ns and the hot bath and the cold bath are maintained at 320 K and 280 K respectively. When the system reaches equilibrium, the thermal conductivity can be calculated as:

$$\kappa=\frac{dQ/dt}{S(\Delta T/L)} \tag{5}$$

where dQ/dt denotes the average thermal power from hot bath to cold bath, S is the cross-sectional area, ΔT is the temperature difference between the hot bath and cold bath, which is 40 K in this paper and L is the length of the interfacial region.

3. Results and discussions

To explain the proposed pattern analysis method, we take the optimization of 14-atom-long chain from the 18-atom-long chain database as an example, of which the corresponding idiographic flowchart is shown in Fig. 3. First, the diatomic chain structures are encoded in binary digits, where digit 0 represents Si atom and digit 1 represents Ge atom, which are later used as descriptors in the machine learning optimization. As for the evaluator, the thermal conductance of the diatom chain calculated by AGF is used to quantitatively evaluate the performance of each structure for the machine learning. The theoretical framework of AGF method has been discussed in methodology section. Here, both the left and right leads in the AGF model are fixed as Si atoms, which can be adjusted for further calculations. For a diatom chain with a length of N atoms, there will be 2^N possible candidates. After the optimization of the 18-atom-long chain structures via the Monte Carlo tree search (MCTS) method [21], the top 100 structures with the ascending thermal conductance are selected to form the database. Then, we set the target length of the short pattern as 14 and the probability analysis of 14-atom-long chain structure in the database is carried out to sum up the occurrence times in the top 100 structures of the 18-atom-long chain. By statistical probability analysis, the pattern with the highest probability is taken as the final optimized structure. Since the 1D AGF model is bilaterally symmetric (i.e., the thermal conductivities of symmetric structures are the same), short patterns with bilateral symmetry are thus regarded as the same during pattern analysis. Consequently, the optimal structures always appear in symmetric pairs in the final frequency analysis. As shown in Fig. 3, the optimized 14-atom-long pattern is 11011010101001 and 10010101011011 by the pattern analysis.

To examine whether the obtained 14-atom-long structure is the optimized one, the machine learning for the 14-atom-long chain by MCTS algorithm is also carried out. The distribution histogram of thermal conductance of all candidates, as shown in the inset of Fig. 4(a), resembles a logarithmic Gaussian distribution. The probability decreases as the thermal conductance reaches the minimum, reaffirming the effectiveness of MCTS algorithm in searching for the structure with minimum thermal conductance. In the process of MCTS, 5 dependent rounds of machine learning optimizations for 14-atom-long diatomic chains are conducted and the evolution process of the minimum thermal conductance are recorded. As shown in Fig. 4(a), the optimized thermal conductance converges quickly, within 400 structures, which is about 3%

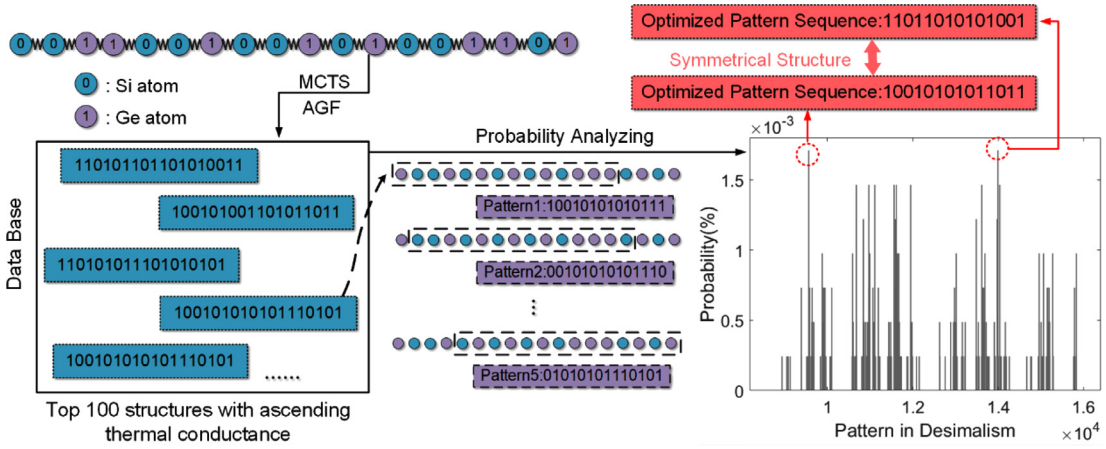


Fig. 3. Pattern analysis flow chart. For the convenience of statistics, the abscissa in the frequency graph is the corresponding number after the transformation of pattern binary sequence structure into decimal system.

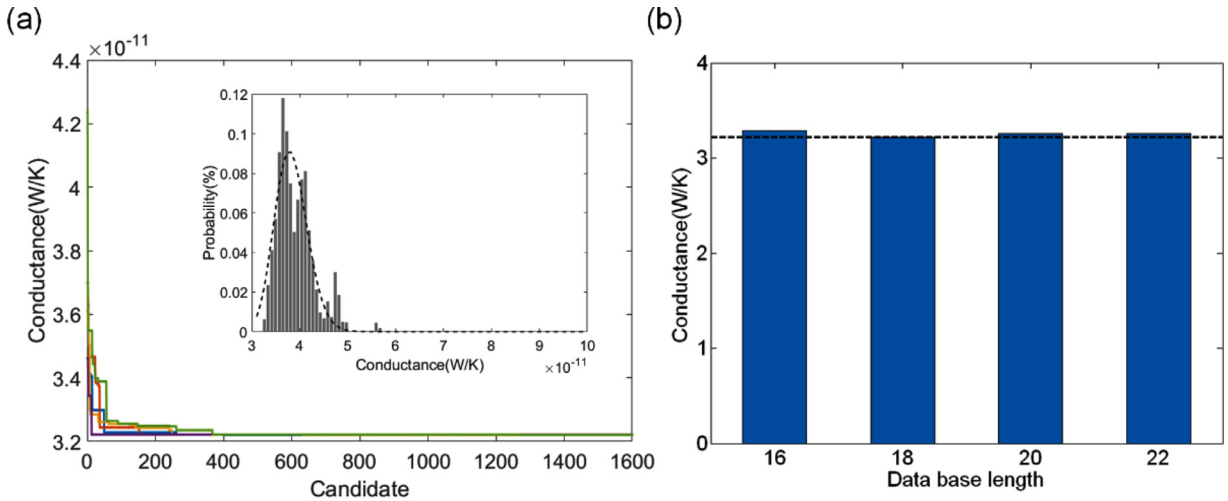


Fig. 4. (a) The 5 rounds of machine learning optimization. The lines in different colors represent the optimization process of different rounds. The inset is the distribution histogram of thermal conductance of all candidates where the dotted black line is the fitting curve of the thermal conductance distribution. (b) Comparison of thermal conductance of 14-atom-long chain optimized by different length databases, the dashed line represents the minimal conductance optimized by MCTS.

of the total number of the candidates (16,384). Surprisingly, it is found that the structure with the minimum thermal conductance by MCTS is 11011010101001, which is completely consistent with the optimized structure by pattern analysis method. The comparison of thermal conductance of 14-atom-long chain optimized by different length databases is made to explore the database length dependence in pattern analysis method. As shown in Fig. 4(b), The 14-atom-long chain optimized by different database differs little, proving that the pattern analysis method is not affected by the database length (as long as the database length is longer than the length of the structure to be optimized). Here we choose the 18-atom-long database just as an example.

To further explore the physical mechanisms, the atomic chain models of pure Si structure, periodic Si/Ge superlattice structure, and pattern optimized aperiodic Si/Ge superlattice structure are constructed and their corresponding phonon transmission are analyzed in details, which is shown in Fig. 5. In pure Si, since there is no interface scattering, phonon transmission in the entire frequency range below the cutoff frequency is always unity, resulting in the maximum thermal conductance of pure Si. In both the periodic structure and the pattern structure, phonon transmission decreases to a certain extent, and the Fabry-Pérot oscillations appear in the low frequency region [27]. For such 1D design structures, there are two main factors that affect the thermal conductance: the number of interfaces in the model and the length of each layer L in the model (i.e., the distance between two interfaces). It is known that phonon scattering increases as the number of interfaces increases, thereby reducing the model thermal conductance. Furthermore, Fabry-Pérot oscillations will increase as L increases. When L is small, the resonance will only appear at a few specific frequencies; as L increases, the frequency range displaying resonance will gradually extend. Finally, when L is large enough, resonances will occur over the entire frequency range. Due to the appearance of Fabry-Pérot oscillations, the thermal conductivity of the model will gradually decrease with the increase of L and eventually converges [15]. For an atom chain of a given length, the number of interfaces will decrease as L increases, and vice versa, thus these two parameters are competitive. In the process of pattern analysis, the models in the optimized long chain database via MCTS are all aperiodic, and the thermal conductance of the structures is significantly lower than that of the periodic ones. The reason is that the aperiodic structure has more degrees of freedom and can be allocated with different numbers of interfaces and lengths of layers. Similarly, the optimum short chain structure is taken from the pattern that appears most frequently in the optimized long chain database obtained from machine learning, resulting in a low thermal conductance of the optimized short chain structure.

To validate optimization performance of the pattern analysis method, pattern optimized structure is compared with the machine

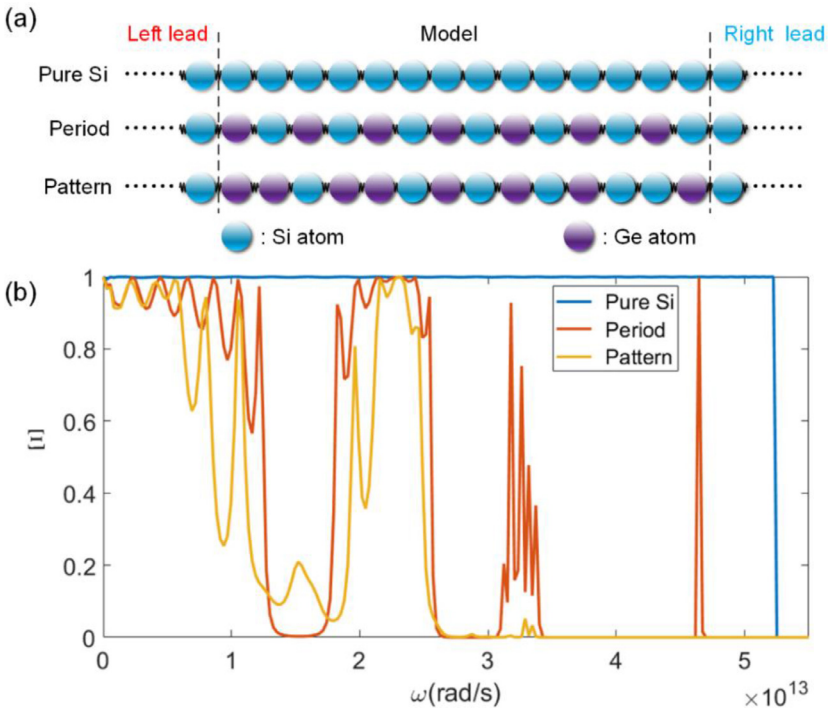


Fig. 5. (a) Schematic diagram of pure Si atomic chain, periodic atomic chain and pattern optimized atomic chain at Si-Si interface, in which the blue atoms represent Si and the purple atoms represents Ge. (b) Phonon transmission of pure Si, periodic and the pattern optimized atomic chain.

learning optimized structure and the optimized structure in previous literatures [15], in which, the machine learning process has already been described in the second paragraph of this section. Here, the normalized thermal conductivity—which is defined as the thermal conductivity of a structure divided by the thermal conductivity of a pure Si structure of equal length—of the three different structures (i.e., the pattern optimized structure, the machine learning optimized structure and the optimized structure in previous literature [15]) of the same length of 14 has been calculated as the comparative indices. Thermal conductivity is the product of thermal conductance and system length; therefore, the normalized thermal conductivity of the structure is the same as its normalized thermal conductance. It should be noted that the AGF calculation in this paper is based on 1D atomic chains, while the models in the previous literature [15] are 3D structures, thus the obtained optimized structures may be different with slightly different normalized thermal conductance. Though the superlattice has changed in dimensions, its degree of structural randomness has been be preserved, which should lead to ultra-low thermal conductance. And the main purpose of the pattern method is to find the optimal degree of randomness, which has been discussed in the first section. As shown in Fig. 8a, for the 14-atom-long chain, the normalized thermal conductivities of the three structures are almost the same, which further proves the validity of the pattern analysis method.

The above analysis results are based on the 1D atomic chain model which may be inapplicable to practical applications. Hence, we transform the aforementioned three 14-atom-long chain into experimentally feasible 3D superlattice to further verify whether the optimized model via pattern analysis method can be directly applied to interfacial materials in practice. We choose non-equilibrium molecular dynamics (NEMD) method to calculate the thermal conductivity of the 3D superlattice. The superlattice structure is also constructed according to the optimized structure binary code, where 1 represents a Ge layer and 0 represents an Si layer. The Si layer and the Ge layer have the same size, and each layer consist of 1×6×6 unit cells (UCs) in the x, y, z axis directions, respectively. The lattice constant of Si-Ge is assumed to be exactly the same, i.e., 5.43 Å, since the effect of lattice mismatch on thermal transport in Si-Ge binary structures is negligible [28]. The NEMD model is shown in Fig. 6a, where the blue atoms represent Si and the purple ones represent Ge, and periodic boundary conditions are used in all three directions. The structure binary code of the 3D aperiodic superlattice model is the same as the 1D pattern optimized model (i.e. 11011010101001), adjacent to which are hot bath and cold bath respectively. Both ends of the model are fixed regions with a thickness of 2UC, where the atoms are fixed to prevent the bulk motion of the whole system. Specific simulation steps have been collectively described in the methodology section. Fig. 6b records the temperature field in the x direction of the model after simulation reaches equilibrium, and the dotted line indicates the boundary position between the thermostat and the interfacial region. Due to the thermal resistance generated by phonon scattering at Si-Ge interface, temperature is not linearly distributed in the interfacial region and temperature jump occurs at some positions. Fig. 7a-d respectively shows the sketch of pure Si structure, machine learning optimized structure, pattern optimized structure and optimized structure in previous literature [15], and the detailed numerical normalized thermal conductivity calculated by 1D AGF and 3D NEMD of each structure is indicated on the left side of the figure. Similar to the previous results, although the structures obtained by different optimization methods are not the same, there is little difference in normalized thermal conductivity. It has been found in our previous work [16] that the thermal conductivity of the aperiodic superlattice is inversely correlated with the randomness of its structure, and in the structures with the same length, there exist many, if not a tremendous large number of, sufficiently random aperiodic superlattices, of which the thermal conductivity is very close to the minimal thermal conductivity. Thus, it is not surprising to find that different optimization approaches lead to different structures, but all of them possess similar thermal conductivity. Owing to the relationship between thermal conductivity and structural randomness in aperiodic superlattice, the process of aperiodic structure with low thermal conductivity is actually the process of enhancing its structural randomness. From this point of view, the pattern analysis based on the 1D AGF is able to effectively search for a sufficiently random structure which also exhibits low thermal conductivity in 3D superlattice. It can be noted from the 1D and 3D verifications that the superlattice randomness is the key for quenching thermal conductivity of superlattice [16]. Moreover, our previous work [17] also indicates that the optimized aperiodic superlattice is composed by several short-chain local structures with or

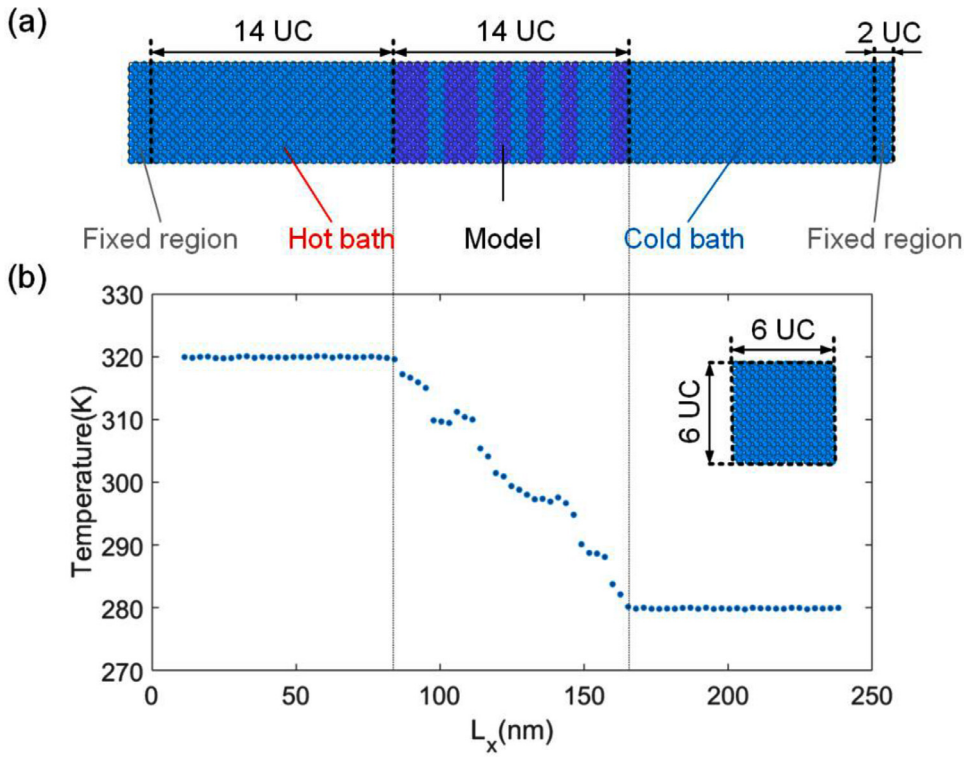


Fig. 6. (a) Front view of the 14 UC superlattice computational model and its regional distribution. (b) Temperature distribution of the model along the *x* direction, where the inset is a side view of the model.

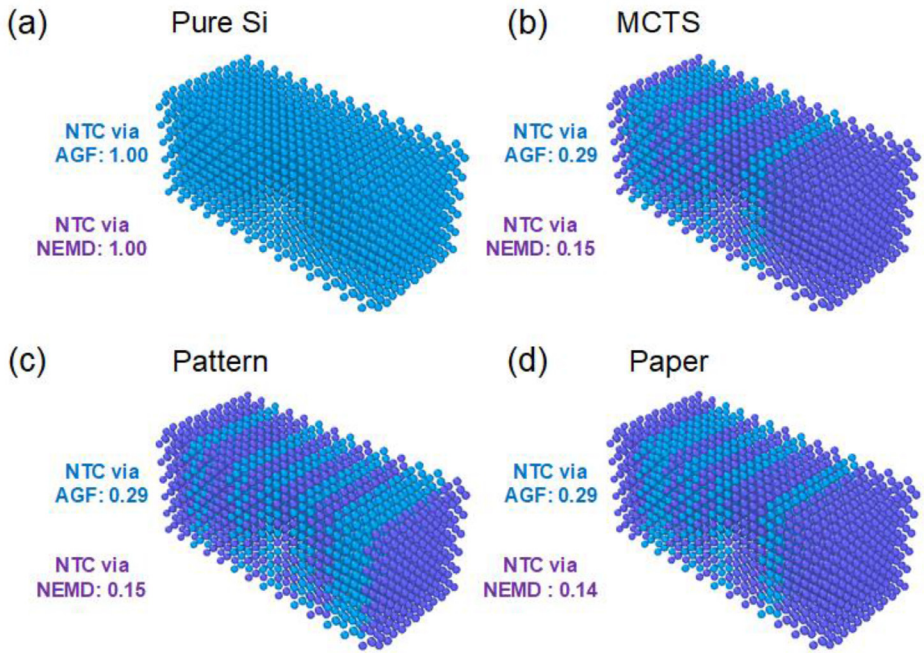


Fig. 7. 3D sketch of (a) pure Si, (b) optimized structure obtained by MCTS, (c) optimized structure obtained by pattern analysis and (d) optimized structure in previous literature [14]. The left side is their respective normalized thermal conductivity (NTC) via AGF or NEMD.

without overlapping, which lay the foundation of underlying physics for the pattern analysis.

The above verification is only based on the optimized superlattice with length of 14 unit layers. We extend the type of calculation model to verify the universality of the pattern analysis method based on the same 18-atom-long chain database. The pattern analysis is performed for various lengths from 8 to 16; and the Si/Ge atomic proportion is set to be equal and variable respectively for each length. Fig. 8a and Fig. 8c shows the normalized thermal conductivity of three models with a variable Si/Ge atomic proportion via 1D AGF and 3D NEMD for the same structures predicted by MCTS, pattern analysis, and previous results, respectively. Fig. 8b and Fig. 8d shows the normalized thermal conductivity of three models with a fixed Si/Ge fraction (1:1) via 1D AGF and 3D NEMD, respectively. The trends in the four figures are roughly the same, as the model length increases, the normalized thermal conductivity of the optimized model decreases gradually. It is evident that at the same model lengths and same atomic proportions situation, the three kinds of optimized models are not identical, which leads to different normalized thermal conductivity. However, the difference of the three models is negligible for 1D atomic chain or 3D structures. What's more, the computational cost of the pattern analysis is worth mentioning. Since the pattern analysis only involves simple counting analysis, it can save

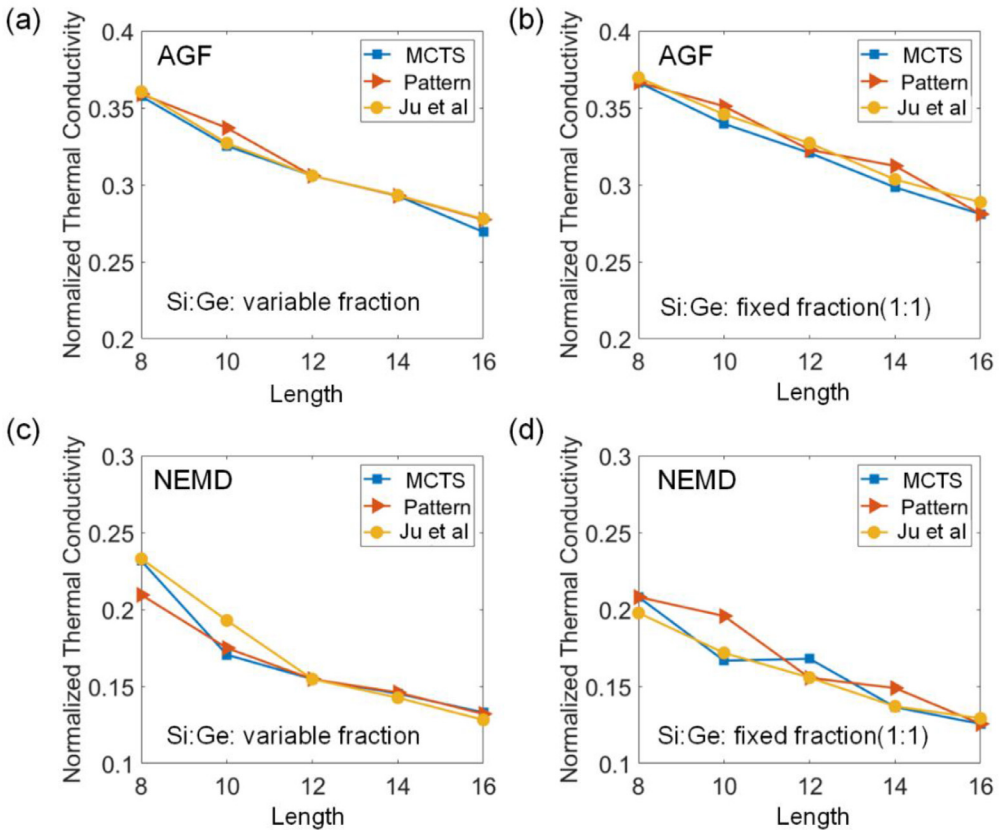


Fig. 8. (a-b) Comparison of normalized thermal conductivity of the different optimized models with variable lengths via 1D AGF calculation. (c-d) Normalized thermal conductance comparison of various optimized structures verified by 3D NEMD calculation. The Si/Ge atoms fraction are fixed at 1:1 in (b) and (d), while the fraction is set to be variable in (a) and (c).

a lot of computing time compared to machine learning optimization. Take the 14-atom-long chain optimization as an example, it costs about 88 core minutes by MCTS optimizing the structure, while only 28 core minutes are used by pattern analysis, which was only one-third of the former. Thus, we conclude that the proposed pattern analysis method is effective to find the optimized structure with low thermal conductivity, and it is easy to implement, flexible to extend, low computation cost, and high efficiency with machine learning optimization.

It should be noted that though, in this work, the pattern method is used to predict the binary-atom superlattice structure, more complicated cases may also be achieved actually. We also can add different types of atoms, layer thickness, atom ratios by adding other digits like “1″, “2″, “3″, and so on. For example, we can use the volume fraction of silicon atom as variable in Si/Ge superlattice, and define digit “1″ for 10% silicon content in each layer, “2″ for 20% silicon content in each layer and so on. Then, one possible structure may be marked as a digit sequence like “0–2–1–…–3″. Once the structure descriptor is defined, the following steps are the same for optimizing the superlattice. Since this type of structure is much more complex, the number of structure possibilities will be enormous, which increase the computation expense to obtain the long-chain machine learning database.

4. Conclusion

In summary, we propose a pattern analysis method to accelerate the optimization of quenching thermal conductivity of Si/Ge superlattice. First, the optimized structure database with long structure is constructed through machine learning, then the target length and proportion of the desired optimized structure are set. The structure with the greatest probability in the optimized database is taken as the final optimized structure. In order to verify the effectiveness of this method, we construct 1D atomic chain based on AGF method, which is then validated by comparing with the corresponding 3D structure via NEMD simulations. Finally, by comparing the pattern optimized structure with the optimized structure in machine learning and other literatures, we find that the thermal conductivity difference between the three structures is tiny, which proves the effectiveness of the pattern analysis method and structural randomness is the key parameter for quenching thermal conductivity of superlattice. The proposed pattern analysis is validated to effectively reduce the computational time consumed in the multivariable problem of machine learning, which enables the prediction of more complex and better-performing structures of a specific material, and opens a new avenue for the further development of material informatics.

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.

Acknowledgements

This work was supported by National Natural Science Foundation of China (52076087), and the Ministry of Science and Technology of the People’s Republic of China (2017YFE0100600), and Wuhan City Science and Technology Program (2020010601012197).

References

[1] Luckyanova MN, Garg J, Esfarjani K, Jandl A, Bulsara MT, Schmidt AJ, et al. Coherent phonon heat conduction in superlattices. Science 2012;338:936.
[2] 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.
[3] Lin Z, Yin A, Mao J, Xia Y, Kempf N, He Q, et al. Scalable solution-phase epitaxial growth of symmetry-mismatched heterostructures on two-dimensional crystal soft template. Sci Adv 2016;2:e1600993.
[4] Chen S, Wu Q, Mishra C, Kang J, Zhang H, Cho K, et al. Thermal conductivity of isotopically modified graphene. Nat Mater 2012;11:203.

[5] Merabia S, Termentzidis K. Thermal boundary conductance across rough interfaces probed by molecular dynamics. Phys Rev B 2014;89:054309.
[6] Liu Y, Hu C, Huang J, Sumpter BG, Qiao R. Tuning interfacial thermal conductance of graphene embedded in soft materials by vacancy defects. J Chem Phys 2015;142:244703.
[7] Chowdhury PR, 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.
[8] Wang Y, Gu C, Ruan X. Optimization of the random multilayer structure to break the random-alloy limit of thermal conductivity. Appl Phys Lett 2015;106:073104.
[9] Yang L, Yang N, Li B. Reduction of thermal conductivity by nanoscale 3D phononic crystal. Sci Rep 2013;3:1143.
[10] Wang Y, Huang H, Ruan X. Decomposition of coherent and incoherent phonon conduction in superlattices and random multilayers. Phys Rev B 2014;90:165406.
[11] Qiu B, Chen G, Tian Z. Effects of aperiodicity and roughness on coherent heat conduction in superlattices. Nanoscale Microscale Therm 2015;19:272.
[12] Hu R, Song J, Liu Y, Xi W, Zhao Y, Yu X, et al. Machine learning-optimized Tamm emitter for high-performance thermophotovoltaic system with detailed balance analysis. Nano Energy 2020;72:104687.
[13] Rajan K. Materials informatics. Mater Today 2012;15:470.
[14] Seko A, Togo A, Hayashi H, Tsuda K, Chaput L, Tanaka I. Prediction of low-thermal–conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization. Phys Rev Lett 2015;115:205901.
[15] Ju S, Shiga T, Feng L, Hou Z, Tsuda K, Shiomi J. Designing nanostructures for phonon transport via Bayesian optimization. Phys Rev X 2017;7:021024.
[16] Chakraborty P, Liu Y, Ma T, Guo X, Cao L, Hu R, et al. Quenching thermal transport in aperiodic superlattices: a molecular dynamics and machine learning study. ACS Appl Mater Interfaces 2020;12:8795.
[17] Hu R, Iwamoto S, Feng L, Ju S, Hu S, Ohnishi M, et al. Machine-learning-optimized aperiodic superlattice minimizes coherent phonon heat conduction. Phys Rev X 2020;10:021050.
[18] 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.
[19] Datta S. Quantum transport: atom to transistor. New York: Cambridge University Press; 2005.
[20] Landauer R. Electrical resistance of disordered one-dimensional lattices. Philos Mag 1970;21:863.
[21] 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;51:333.
[22] 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.
[23] Plimpton S. Fast parallel algorithms for short-range molecular dynamics. J Comput Phys 1995;117:1.
[24] Tersoff J. Modeling solid-state chemistry: interatomic potentials for multicomponent systems. Phys Rev B 1990;39:5566.
[25] Berendsen HJC, Postma JPM, Gunsteren WFV, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684.
[26] Schneider T, Stoll E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. Phys Rev B 1978;17:1302.
[27] Hyldgaard P. Resonant thermal transport in semiconductor barrier structures. Phys Rev B 2004;69:193305.
[28] Gudiksen MS, Lauhon LJ, Wang J, Smith DC, Lieber CM. Growth of nanowire superlattice structures for nanoscale photonics and electronics. Nature 2002;415:617.