International Journal of Heat and Mass Transfer 160 (2020) 120176

Contents lists available at ScienceDirect

International Journal of Heat and Mass Transfer

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

Machine learning prediction of thermal transport in porous media with physics-based descriptors

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

ARTICLE INFO

Article history:
Received 13 April 2020
Revised 19 June 2020
Accepted 6 July 2020

Keywords:
Porous media
Effective thermal conductivity
Physics-based descriptors
Machine learning
Support vector regression
Gaussian process regression

ABSTRACT

Understanding the thermal transport mechanism in porous media is important for various engineering and industrial applications. The effective thermal conductivity of porous media is known to be related to the morphology of porous structures. However, existing effective medium approaches usually miss the morphology effects, and numerical simulations are expensive and not physically intuitive. Machine learning methods have recently been successful in predicting effective thermal conductivity, but the lack of descriptors limits physical insights. In this work, we investigate structural features that have significant effects on thermal transport in porous media and identify five physics-based descriptors to characterize the structural features: shape factor, bottleneck, channel factor, perpendicular nonuniformity, and dominant paths. These descriptors can effectively quantify the anisotropy of pore morphology and strongly correlate with effective thermal conductivities. The proposed descriptors are incorporated into machine learning models to predict the effective thermal conductivity of porous media, and the results are shown to be fairly accurate. They provide new insights into the thermal transport mechanisms in complex heterogeneous media.

1. Introduction

Porous media, such as fibrous composites, granular materials, packed beds, and foam, have a variety of applications in material, aerospace, biological and medical engineering [1,2]. Effective thermal conductivity is one of the key thermophysical properties of porous media, especially for thermal management applications [3]. There have been numerous previous investigations on heat conduction in porous media through experiments [4], analytical methods [5], or numerical simulations [6,7]. The effective thermal conductivity of porous media is known to depend on the thermal conductivity of the composition phase, porosity, and porous morphology [8]. However, the effect of the latter is complicated. Effective medium approaches often ignore it, while numerical methods capture it but do not offer physical insights. Identifying and quantifying morphology features that are correlated to thermal conductivity is vital for understanding the heat conduction mechanism in porous media.

With the recent advances of data-driven techniques, investigating the structure-property relationship through machine learning methods has attracted significant research interest [9–17]. Machine learning methods have been successfully applied to obtain surrogate models for predicting the effective thermal conductivity of composite materials and porous media [18,19]. Machine learning methods can generally be divided into two categories. One is the deep learning methods based on convolutional neural network (CNN) [20], which can automatically extract structural feature maps of porous media and accurately predict the effective thermal conductivity. However, because feature maps [20] are not physically interpretable, CNN can only serve as a prediction model without physical understanding [18]. The other category is conventional machine learning methods, such as support vector regression (SVR) [21] and Gaussian process regression (GPR) [22]. The accuracy of these methods heavily relies on manually selected structural descriptors [23]. Some existing descriptors can effectively quantify structural features, but the physics is not intuitive, such as the two-point correlation function [24,25]. Other descriptors can have clear physical meanings, such as pore size, aspect ratio, and orientation distribution [26,27]. These physics-based descriptors are more favorable because they can guide understanding of the thermal transport mechanism in porous media. These descriptors are also commonly used in analytical models to include specific morphological features [28]. In order to achieve a better prediction accuracy for complex porous media and at the same time guide the understanding of heat transfer mechanism, developing physics-

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

https://doi.org/10.1016/j.ijheatmasstransfer.2020.120176
0017-9310/

Nomenclature

ε	Porosity
c_d	Core distribution probability
g_i	Directional growth probability
A_x	matrix cross-section along x direction
A_y	matrix cross-section along y direction
SF	Shape factor
BL	Bottleneck
PN	Perpendicular nonuniformity
CF	Channel factor
DP	Dominant paths

based descriptors to quantify the complex morphologies in porous media is highly desirable.

In this work, the structural features of high and low thermal conductivity complex porous media are first examined. We propose five physics-based descriptors that are intuitive and able to capture the structural features related to thermal transport. The strong correlations between these descriptors and the effective thermal conductivity are demonstrated. The effective thermal conductivities of complex porous media are then accurately predicted using conventional machine learning methods. The rest of the manuscript is organized as follows. In Section 2, the structural features that have an important effect on the effective thermal conductivity of porous media are investigated with a large dataset composed of porous structures and corresponding thermal conductivities. In Section 3, five physics-based descriptors are identified to represent the significant structural features. In Section 4, the correlations between descriptors and thermal conductivities are presented. The physical significance of the descriptors in describing the structural features is discussed. In Section 5, machine learning analysis with the proposed descriptors is demonstrated.

2. Significant structural features

We focus on thermal transport in two-dimensional (2D) porous media with granular pores. Quartet structure generation set (QSGS) is a random internal morphology and structure generation-growth method, which can reflect the stochastic distribution characteristics of most porous media [29]. Therefore, QSGS is adopted here to generate granular porous configurations that closely resemble realistic granular porous media. [29]. In the 2D QSGS generation process, there are three major input parameters, which are porosity ε, core distribution probability c_d, and growth probability in eight directions g_i, $i = 1...8$ [30]. c_d is used to control the averaged pore size. Directional growth probabilities are used to control the anisotropy of the pore shape. Fig. 1(a) shows the eight growth directions of pores, where g_{1-4} are the main directional growth probabilities, and g_{5-8} are the diagonal directional growth probabilities. Fig. 1(b) shows an example of the generated 2D porous media, where the pores (white phase) are discontinuously distributed in the continuous matrix (blue phase).

Studies have shown that effective thermal conductivity can be significantly affected by different morphologies of pores [31,32]. As such, we intentionally generate structures with different morphological characteristics, including porosity, pore size, shape, and distribution. Specifically, the porosities ε include 10%, 20%, 30%, 40%, and 50%. The pore size is controlled by taking different core distribution probability c_d, including 0.0125 ε, 0.1 ε, and 0.025 ε. The pore shape is controlled by different directional growth probabilities in eight directions, which include 41 different cases. Because the generation process is random [30], different distributions of pores can be naturally included. We generate five structures for each combination of these input parameters to span the different distributions. In total, 3075 structures are generated for the whole dataset.

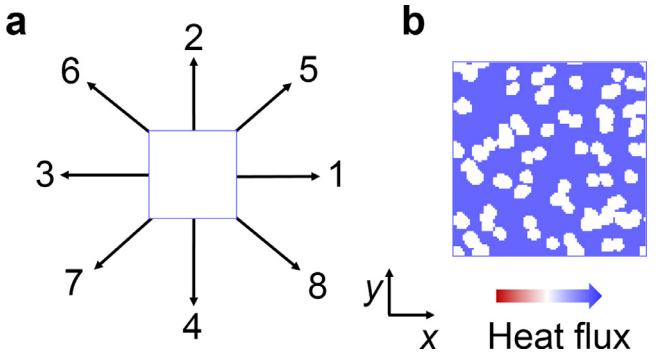

Fig. 1. Generation of porous structure with the QSGS method. (a) Eight growth directions of each point for 2D systems. The direction 1, 2, 3 and 4 denote the main growth directions and direction 5, 6, 7 and 8 denote the diagonal growth directions. (b) An example of QSGS generated 2D porous structure. The blue and white region denotes the matrix (solid) and pore phase. In FEM simulation, the heat flux is set along the x direction.

The lengths and widths of all configurations are set to 100 pixels so that the total size is 100 × 100, as shown in Fig. 1(b). It is assumed that heat transfer is along the x direction. To obtain the effective thermal conductivity, we consider macroscopic heat transfer by solving the heat diffusion equation and applying Fourier's law [33]. In the macroscopic regime, size does not affect the values of effective thermal conductivity, so its unit is arbitrary [34]. Similarly, temperature difference does not affect the effective thermal conductivity [34], so fixed temperature boundary conditions with a dimensionless temperature of 1 and 0 are assigned to the left and right boundary, respectively. Periodic boundary conditions are used for the top and bottom boundaries. Without loss of generality, the dimensionless thermal conductivity values of matrix and pores are fixed as 100 and 1, respectively. The 2D heat diffusion equation is solved with the finite element method (FEM), which is implemented with a python package developed by the FEniCS Project [35]. In the FEM simulation, each pixel is further discretized into 4 elements to ensure numerical convergence.

Fig. 2(a) shows the effective thermal conductivities of the total 3075 configurations (red hollow circles) with different porosities. The results from the parallel (magenta line) and series model (black curve) are also plotted, serving as the upper and lower bounds of effective thermal conductivity [36]. Thermal conductivity values of all configurations cover almost the entire range between the upper and lower bounds. Fig. 2(a) also shows the effective thermal conductivities predicted by the Maxwell model [37] (blue curve) and the Bruggeman model [38] (green curve). It can be clearly seen that their accuracies are limited because they only include porosity as the structural feature and can be only applicable to isotropic media [28].

In order to consider the structural effect on the effective thermal conductivity, we further examine the detailed morphology of different porous structures. Fig. 2(b–j) shows typical porous structures and their corresponding thermal conductivities. These values are also indicated in Fig. 2(a). Structures (b), (c) and (d) all have isotropic pores and their thermal conductivity values decrease with the increase of porosity. The values are close to the Bruggeman model, which confirms the capacity of the Bruggeman model in estimating the thermal conductivity of isotropic structures. Structures (e), (f), and (g) have the same porosity, similar pore size, and shape, but differ in the distribution of pores, causing a significant difference in the thermal conductivity. Structures (h), (i), and (j) have the same porosity and similar pore size but differ in the

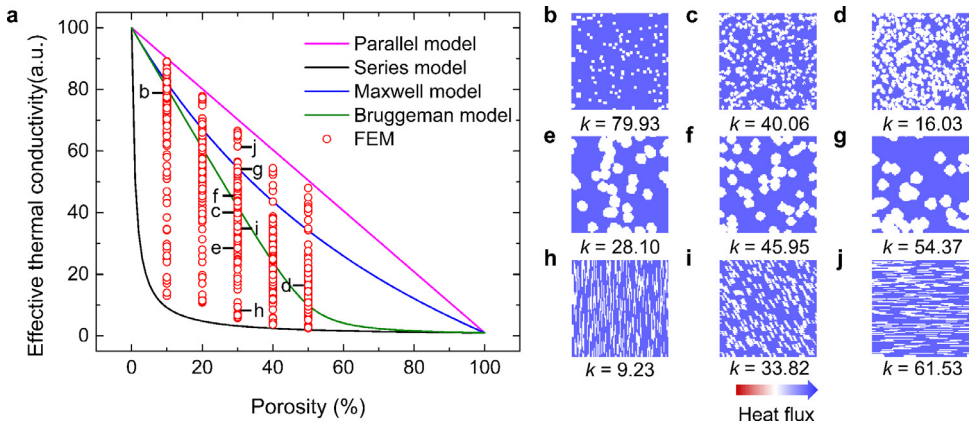

Fig. 2. Thermal conductivities and typical porous structures. (a) Thermal conductivities of the 3075 porous configurations as a function of the porosity, computed by the FEM method and four analytical models: parallel model (magenta line), series model (black curve), Maxwell model (blue curve) and Bruggeman model (green curve). (b–j) Typical porous structures with different porosity, pore size, shape, and arrangements. The corresponding thermal conductivity is shown below the figure and marked in (a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

anisotropies of pore shape, also inducing large difference in thermal conductivity. These results indicate that besides porosity, the pore shape and distribution can significantly affect the heat transfer in porous media.

Therefore, for complex porous media, the detailed morphology or pore arrangement should be considered. Existing structural descriptors, such as aspect ratio, the number of clusters, and the orientation angle of a cluster's principal axis [27], are difficult to quantify in our system. Even they can be quantified; the relationship between the descriptors and thermal transport is unclear. In our previous investigation of nanoporous graphene, two structural descriptors, shape factor and channel factor, are found to be strongly correlated to the thermal transport property [39], besides considerations of phonon ballistic and wave effects. Inspired by that work but neglecting ballistic and wave effects of heat conduction since we consider diffusive transport here, we propose five structural descriptors attempting to capture the structural differences of the typical structures shown in Fig. 2.

3. Physics-based descriptors

Five descriptors are proposed to describe the pore shape and distribution that have significant effects on the heat conduction. In this section, we first introduce the basis of descriptors and then discuss the definitions and physical meanings of these descriptors.

3.1. Basis of descriptors

For a typical structure shown in Fig. 3(a), we can accumulate the matrix and pore phases along the x and y direction, respectively. The rearranged structures are shown in Fig. 3(b) and (c), where the matrix cross-section along the x, y direction, $A_x(y)$ and $A_y(x)$, are also shown. The subscript denotes the aligning direction of the cross-section and the variable is shown in the bracket. They serve as the basis to build our five descriptors.

3.2. Descriptors

The five descriptors we propose are the shape factor (*SF*), bottleneck (*BN*), channel factor (*CF*), perpendicular nonuniformity (*PN*), and dominant paths (*DP*). Their definitions are given below.

3.2.1. Shape factor

The shape factor is a useful concept to describe heat conduction in two-dimensional or three-dimensional systems by the graphical method [33]. If 2D or 3D heat conduction is treated with the quasi-1D approximation, the shape factor can be calculated as [33]

$$SF = 1/\int \frac{1}{A_y}\mathrm{d}x. \tag{1}$$

This definition should not be confused with the shape factor in some literature [32], where it is used to describe the sphericity or circularity of pore shape in three-dimensional or two-dimensional porous media.

3.2.2. Bottleneck

The bottleneck is defined as

$$BN = \min A_y, \tag{2}$$

which is the shortest matrix cross-section along y direction. The graphical illustration of bottleneck is shown in Fig. 4(a). It is inferred that the porous configuration with a large piece of connected pore phase along the y direction can have a small bottleneck. The bottleneck effect has been illustrated in several research works about nanoporous structures [40,41]. For structures with circular pores, either periodic or disordered, the bottleneck can be identified from the shortest distance between pores along

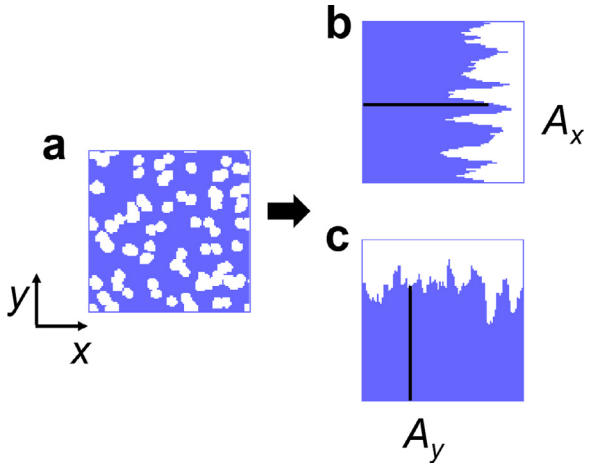

Fig. 3. The schematic of matrix distribution. (a) An example of the morphology of porous structure. (b), (c) The rearranged representations of configuration (a), where the matrix and pore phases are accumulated along the x and y direction, respectively. A_x and A_y are the matrix cross-sections along x and y direction (black lines), respectively.

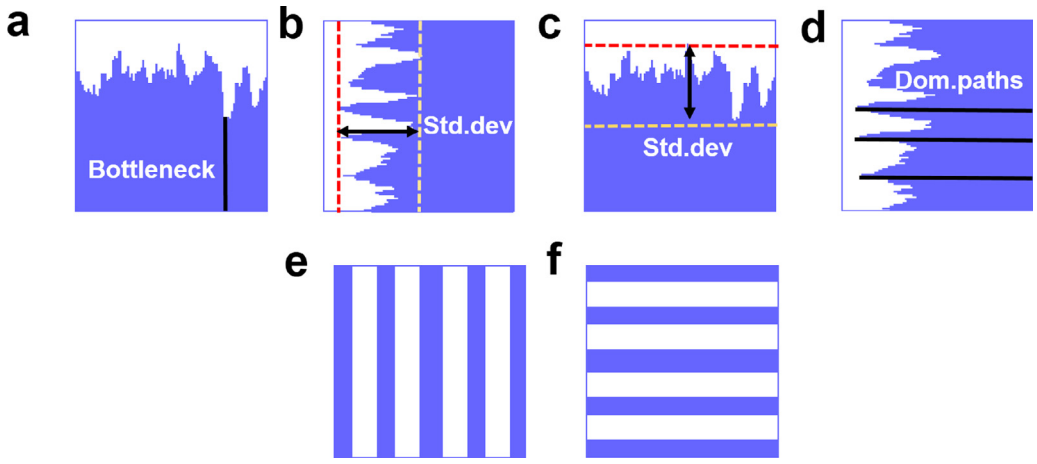

Fig. 4. Graphical illustration of proposed descriptors: (a) Bottleneck. (b) Channel factor. "Std.dev" denotes standard deviation. (c) Perpendicular nonuniformity. (d) Dominant paths. "Dom.paths" denotes dominant paths. (e) Series structure. (f) Parallel structure.

the path that is perpendicular to heat flow [40,41]. For porous materials with random heterogeneous microstructure, it is difficult to identify the bottleneck precisely with a similar approach. We believe the definition as shown in Eq. (2) can sufficiently reflect the blocking effect caused by the narrow channel on the heat flow.

3.2.3. Channel factor

The channel factor is defined as the standard deviation of A_x with probability density function $p(A_x)$, which is given by

$$CF = \sqrt{\int (A_x - \mu)^2 p(A_x) \mathrm{d}y}, \text{ where } \mu = \int A_x p(A_x) \mathrm{d}y. \tag{3}$$

The graphical illustration of the channel factor is shown in Fig. 4(b). The channel factor quantifies the distribution uniformity of matrix cross-sections along the direction parallel to the heat flow. According to the definition, a more uniform distribution of A_x (e.g. the value of A_x is close to constant) results in a smaller value of channel factor. In our previous work about nanoporous material with circular pores [39], the uniformity of pore distribution along the direction parallel to the heat flow is defined as the "channel factor". Here, the terminology is kept consistent.

3.2.4. Perpendicular nonuniformity

The perpendicular nonuniformity is defined as the standard deviation of A_y with probability density function $p(A_y)$, which is given by

$$PN = \sqrt{\int (A_y - \nu)^2 p(A_y) \mathrm{d}x}, \text{ where } \nu = \int A_y p(A_y) \mathrm{d}x. \tag{4}$$

The graphical illustration of the perpendicular nonuniformity is shown in Fig. 4(c). We note that in a previous work analyzing heat transfer in polycrystals, the grain boundaries directly blocking the heat flow have a significant effect on the thermal conductivity [42]. These grain boundaries are similar to the cross section area A_y defined in porous media. To include the complex and stochastic morphologies, the perpendicular nonuniformity is proposed to explicitly quantify the distribution nonuniformity of matrix cross-section area A_y. According to the definition, a more uniform distribution of matrix A_y (e.g. the value of A_y is close to constant) leads to a smaller value of perpendicular nonuniformity.

3.2.5. Dominant paths

The dominant paths are defined as

$$DP = \int_{\lambda} A_x \mathrm{d}y, \tag{5}$$

which is the sum of a certain percent (λ) of the longest matrix cross-sections along x direction. In this work, λ is set as 30%. The graphical illustration of dominant paths is shown in Fig. 4(d). Previous studies have shown continuous heat transfer paths formed in a highly conductive matrix phase can significantly enhance the thermal transport in porous materials [36,43]. To include this effect, we propose the dominant paths to indicate the amount of efficient heat flow paths in the porous structures.

Table 1
The values of descriptors of the series and parallel structure. "*SF*", "*BN*", "*CF*", "*PN*", and "*DP*" denotes shape factor, bottleneck, channel factor, perpendicular nonuniformity, and dominant paths, respectively. ε denotes the porosity, ranging from 0 to 1. The values of the bottleneck, channel factor, perpendicular nonuniformity, and dominant path are normalized by the domain size.

Structure	*SF*	*BN*	*CF*	*PN*	*DP*
Series	0	0	0	$\varepsilon(1-\varepsilon)$	$1-\varepsilon$
Parallel	$1-\varepsilon$	$1-\varepsilon$	$\varepsilon(1-\varepsilon)$	0	1

Figs. 4(e) and (f) show the series and parallel structures and the corresponding values of descriptors are presented in Table 1. Obviously, the values of descriptors for the series and parallel structure are quite different, which indicates that the descriptors can effectively distinguish the structural difference between the two benchmark structures. Because the series (parallel) structure has the lowest (highest) thermal conductivity along the x direction and the highest (lowest) thermal conductivity along the y direction. They can be regarded as two extremes of anisotropic structures. Similar to the effective thermal conductivity, the series structure and parallel structure can serve as the bounds for the anisotropy, which are quantified by the descriptors. As shown in Table 1, the series (parallel) structure serves as the lower (higher) bound for shape factor, bottleneck, channel factor, and dominant paths and higher (lower) bound for perpendicular nonuniformity for a certain porosity.

4. Correlation between the descriptor and thermal conductivity

Next we investigate the correlations between each descriptor and the thermal conductivities, and the results are shown in Fig. 5. It can be seen that the shape factor (Fig. 5a), bottleneck (Fig. 5b) and dominant paths (Fig. 5e) have good positive correlations with thermal conductivities, while the channel factor (Fig. 5c) and the perpendicular nonuniformity (Fig. 5d) have moderate negative correlations with thermal conductivities.

In order to quantify the correlations, Pearson's correlation coefficient r [44] and Spearman's rank correlation coefficients ρ [45] are calculated for each descriptor. The Pearson's correlation coefficient is used for measuring the linear correlation between two variables, which ranges from -1 to 1. A value of 1, 0 and -1 implies a perfect positive linear, no linear and perfect negative linear correlation between the variables, respectively [44]. The Spearman's rank correlation coefficient is used for measuring the monotonic correlation between two variables (whether linear or not), which is equal to the Pearson's correlation between the rank values of the two variables [45].

Table 2 presents the two correlation coefficients for each descriptor. For all the descriptors, the absolute values of Spearman's correlation coefficient ρ are higher than that of Pearson's correlation coefficient r. This is because the Spearman's correlation is

Table 2
Pearson's correlation coefficient r and Spearman's correlation coefficient ρ between descriptors and effective thermal conductivities. "*SF*", "*BN*", "*CF*", "*PN*", and "*DP*" denotes shape factor, bottleneck, channel factor, perpendicular nonuniformity, and dominant paths, respectively.

Descriptors	*SF*	*BN*	*CF*	*PN*	*DP*
r	0.934	0.939	−0.242	−0.663	0.951
ρ	0.952	0.940	−0.338	−0.688	0.960

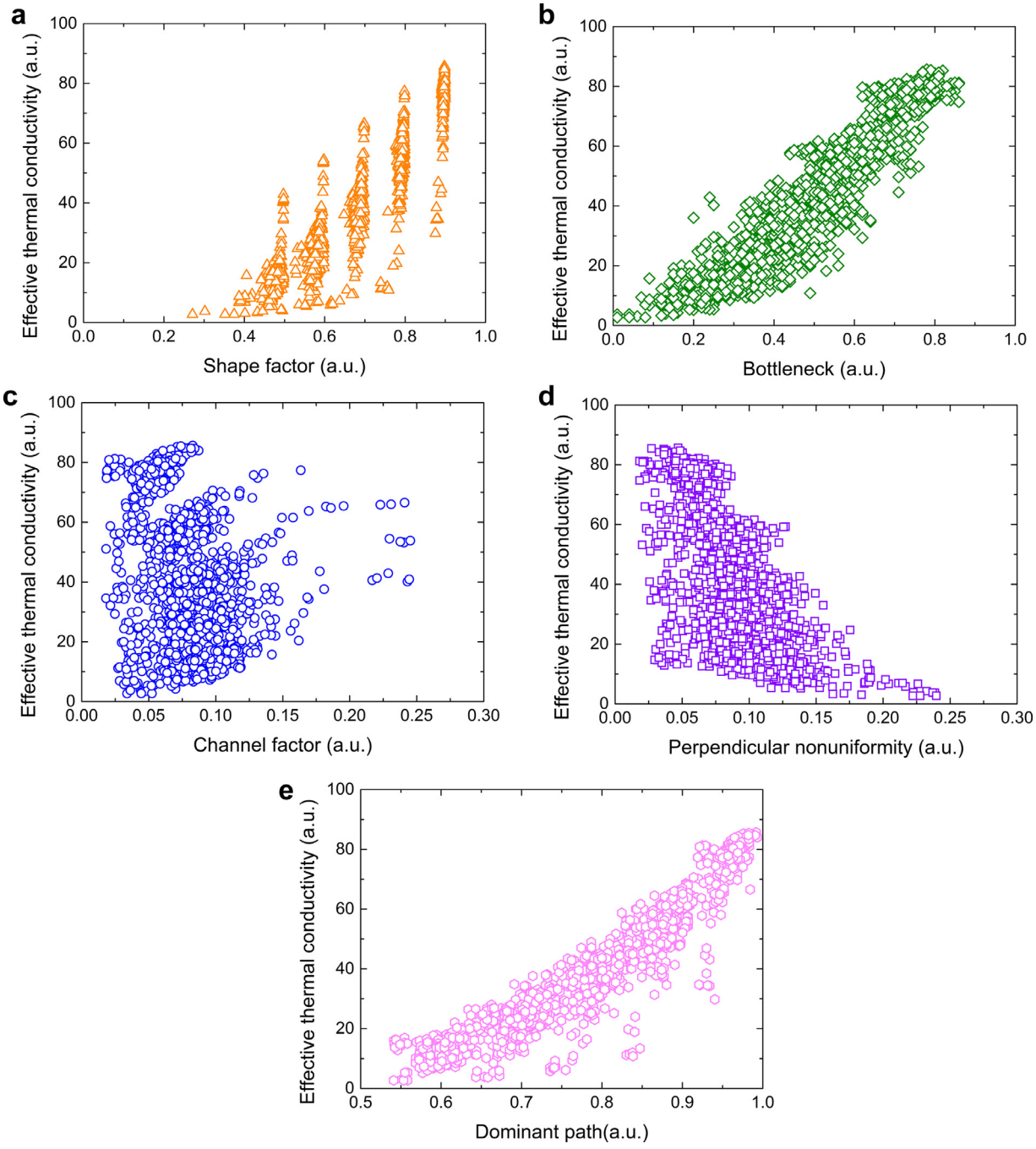

Fig. 5. Correlations between descriptors and effective thermal conductivities: (a) shape factor, (b) bottleneck, (c) channel factor, (d) perpendicular nonuniformity, (e) dominant paths.

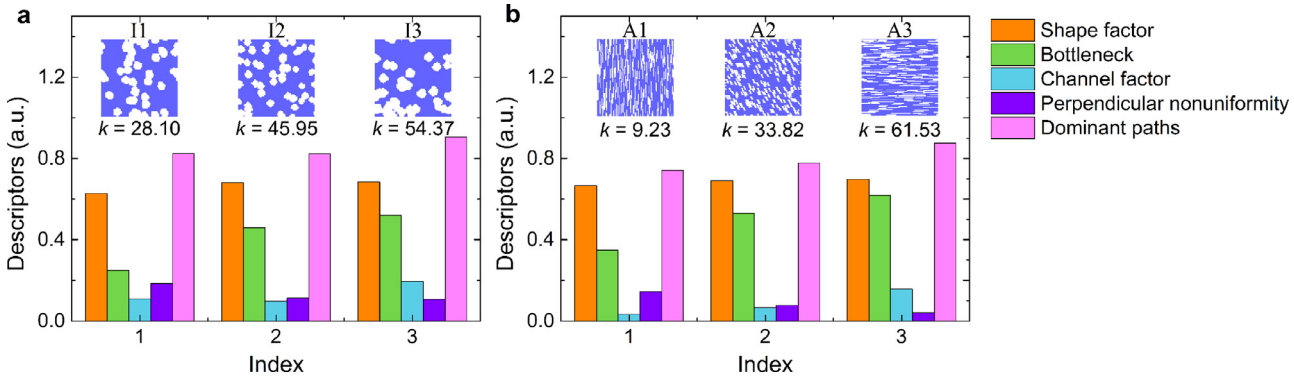

Fig. 6. The descriptors of porous structures with different pore distribution (a) and shapes (b). The corresponding configurations and thermal conductivities are shown at the top of the figures.

related to the value of ranks but not the values of variables and is less sensitive to the outliers than the Pearson's correlation. Both correlation coefficients indicate that the shape factor, bottleneck, and dominant paths have strong positive correlations with thermal conductivities. The perpendicular nonuniformity and channel factor have a moderate and weak negative correlation with thermal conductivities, respectively.

In our previous study about nanoporous graphene [39], the correlation coefficient of channel factor is 0.73, which is higher than the corresponding value in this study. There are two reasons accounting for such a difference. The first reason is that the two studies consider different transport regimes. Here we consider diffusive regime, while in nanoporous graphene phonon transport is mainly in the ballistic regime [39]. The channel effect is more important in that regime due to the classical size effect [46]. On the other hand, the value of the correlation coefficient is also related to the dataset. The correlation coefficient shown in Table 2 is based on the dataset with different pore sizes and anisotropies. In our previous work [39], the pore size is the same and the configurations only differ in pore arrangement. The two factors above collectively result in a lower correlation coefficient of channel factor in comparison with the previous study.

In order to investigate how the morphological features could be quantified by descriptors, we further take an examination of the detailed morphologies of porous structures with different pore distributions and shapes. Fig. 6(a) presents the values of descriptors corresponding to three porous structures with different distributions of isotropic pores. The configurations are shown above the figure, named as I1, I2, and I3, respectively. For configuration I1, some pores are packed densely normal to the heat flow direction, making the structure anisotropic and close to the series structure. For configuration I3, the long channels parallel to the heat flow direction induce the anisotropy, which is close to the parallel structure. In terms of descriptors, the parallel-like structure (I3) has larger values of shape factor, bottleneck and dominant paths and a smaller value of perpendicular nonuniformity than the series-like structure (I1). The significant differences of descriptors between the configurations indicate that the pore distribution can be effectively quantified by the descriptors. Fig. 6(b) shows the values of descriptors corresponding to three porous structures with different anisotropic pores, which are shown above, named as A1, A2, and A3, respectively. Similarly, for the parallel-like structure (A3), the values of shape factor, bottleneck, and dominant paths are larger and the perpendicular nonuniformity is smaller than that of series-like structure (A1). Thus, these descriptors can also effectively distinguish the different anisotropies of pore shape.

As indicated in Section 3, the series and parallel structure are two extremes of anisotropic structures, of which the values of descriptors can serve as the bounds for all the isotropic and anisotropic structures. Therefore, the effect of anisotropy on the thermal transport can be explained by the degree of similarity to the series or parallel structure, which is described by the descriptors. The structure has larger (smaller) values of shape factor, bottleneck and dominant paths or a smaller (larger) value of perpendicular nonuniformity may be more similar to the parallel (series) structure and thus have a high (low) thermal conductivity. From series-like structure to parallel-like structure, the changes of the distribution of matrix will induce the increase of shape factor, bottleneck, dominant path, and decrease of perpendicular nonuniformity according to their definitions, and in turn increase the thermal conductivity, as shown in Fig. 5 and Table 2.

5. Machine learning prediction

We further evaluate the predictive capability of these descriptors for porous structures with machine learning methods. Two conventional machine learning methods are implemented, which are Support Vector Regression (SVR) [47] and Gaussian Process Regression (GPR) [48]. SVR and GPR are two typical nonlinear regression methods that can establish relationships between descriptors and targeted quantity based on statistical learning of the given dataset. The principles of these two methods can be referred to our previous work [18]. The python package scikit-learn [49] is employed to carry out the regression analysis. The five descriptors proposed above are considered in machine learning models, without any additional descriptors.

The dataset described in Section 2 is randomly divided into the training and test sets, which have 2460 and 615 pieces of data, respectively. The training set is used to obtain the relationship between descriptors and thermal conductivities, in which the five-fold cross-validation is applied to avoid over-fitting [50]. The test set is used to estimate the accuracy of the trained model. The relative prediction error is given by

$$\text{Relative prediction error} = \left(\hat{y}_k - y_k\right)/\hat{y}_k \times 100\%, k = 1...K. \quad (6)$$

Besides, three metrics are used for accuracy evaluation, which is the root mean square error (RMSE), mean absolute error (MAE) and determination coefficient (R^2) [19]. The definitions are given by

$$\text{RMSE} = \sqrt{\frac{1}{K}\sum_{k}^{K}\left(\frac{\hat{y}_k - y_k}{\bar{y}_k}\right)^2} \times 100\%, \quad (7)$$

$$\text{MAE} = \frac{1}{K}\sum_{k}^{K}\left|\frac{\hat{y}_k - y_k}{\bar{y}_k}\right| \times 100\%, \quad (8)$$

Table 3
The RMSE, MAE, and R^2 of the training set and test set predicted with five proposed descriptors and only porosity.

		RMSE (%)		MAE (%)		R^2	
		SVR	GPR	SVR	GPR	SVR	GPR
Proposed descriptors	Training	14.95	15.91	9.67	10.52	0.98	0.98
	Test	16.02	17.54	12.28	12.80	0.98	0.97
Only with porosity	Training	57.11	59.71	22.46	22.13	0.92	0.90
	Test	64.10	63.31	24.15	26.61	0.87	0.87

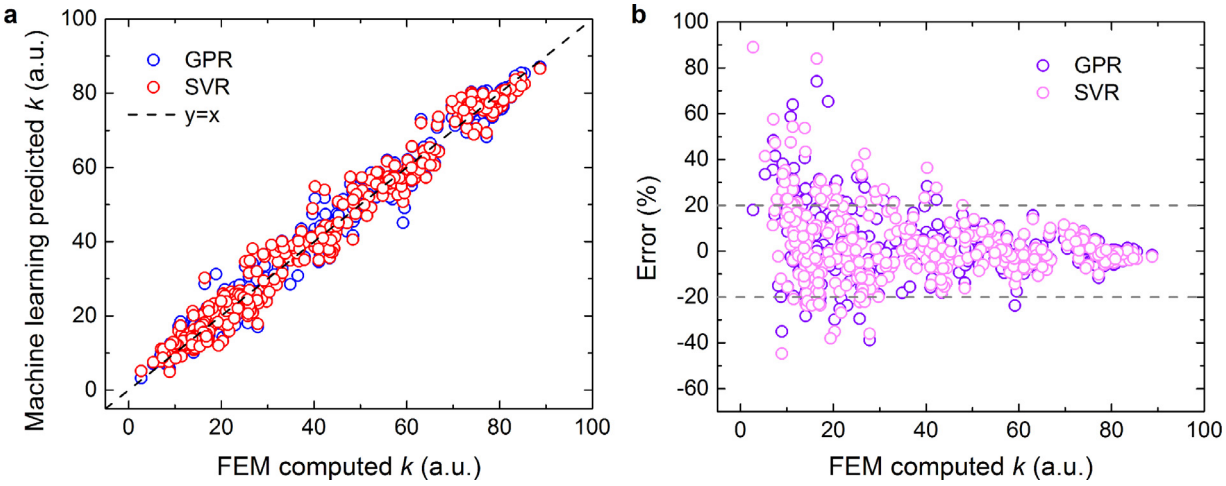

Fig. 7. Machine learning predicted results of the test set. (a) Comparison between the FEM computed effective thermal conductivities and SVR, GPR predicted thermal conductivities. (b) Relative prediction errors with respect to the FEM computed thermal conductivities.

$$R^2 = 1 - \frac{\sum_k^K (\hat{y}_k - y_k)^2}{\sum_k^K (\bar{y}_k - y_k)^2}. \quad (9)$$

In Eq. (6)~(9), $\hat{y}_k$ denotes the accurate value of the effective thermal conductivity, computed by the FEM simulation.$y_k$denotes the predicted value of effective thermal conductivity. $\bar{y}_k$ denotes the average value of the effective thermal conductivity in the set. k denotes the data index and K denotes the number of data in the set.

After training with machine learning methods, the surrogate models for predicting the effective thermal conductivity of porous structures based on five descriptors can be obtained. Accordingly, we can obtain the errors of the predictive model for the training set and test sets. Fig. 7(a) shows the comparison between the FEM computed thermal conductivities and machine learning methods predicted thermal conductivities of the test sets. The machine learning predicted thermal conductivities are close to the FEM computed ones for almost the whole range from 2 to 90, indicating that the predictive model is accurate. Fig. 7(b) shows the relative prediction errors (Eq. (6)) with respect to the computed thermal conductivities. Most relative prediction errors are within the range of −20–20%, which are acceptable in engineering. In addition, both figures show that the prediction results of GPR and SVR are close. We note the relative prediction errors increase with decreasing thermal conductivity and are large when thermal conductivity values are lower than 20. The reason is that we minimize the absolute error instead of relative error in the training process of SVR and GPR. It is possible to reduce them if we choose the relative error to be the loss function.

Table 3 presents the RMSE, MAE, and R^2 of the training set and test set predicted by SVR and GPR method with the five proposed descriptors. It can be seen that the RMSE, MAE, and R^2 of training and test set are close, indicating that both predictive models are well fitted. The value of R^2 is close to 1, which is consistent with the results shown in Fig. 7(a). The errors of SVR and GPR are similar, which is consistent with the results shown in Fig. 7(b). It is clear that the predictive models obtained based on the proposed descriptors are accurate for predicting the effective thermal conductivities of porous media.

We also evaluate the predicting performance with only the porosity as a descriptor, which is also shown in Table 3. Clearly, the errors with only porosity as a descriptor are much larger than that with the five proposed descriptors. From the above results, it can be seen that the five descriptors can fairly accurately predict the effective thermal conductivity of porous media with complex microstructures. Furthermore, the predictive power does not rely on specific machine learning methods.

6. Conclusions

In conclusion, we recognize the significant effect of the pore distribution and shape on effective thermal conductivities of porous media, and propose five descriptors with explicit physical meanings: shape factor, bottleneck, channel factor, perpendicular nonuniformity, and dominant paths. They are found to strongly correlate to effective thermal conductivities of porous media and significantly improve the prediction accuracy compared to effective thermal conductivities based on porosity alone. These physics-based descriptors can be easily extracted from the porous structure. They can serve as good indicators for structural anisotropy, thereby providing intuitive picture about the high (or low) thermal conductivity of porous media.

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.

CRediT authorship contribution statement

Han Wei: Software, Investigation, Writing - original draft, Visualization. **Hua Bao:** Methodology, Data curation, Writing - review & editing, Supervision. **Xiulin Ruan:** Conceptualization, Writing - review & editing, Validation, Supervision.

Acknowledgment

H.B. acknowledges the support from the National Natural Science Foundation of China (No. 51676121) and Guangdong Province Key Area R&D Program (2019B010940001), and H.W.acknowledges the SJTU Zhiyuan Honor Ph.D. fellowship program. Simulations were performed with computing resources granted by HPC (π) from Shanghai Jiao Tong University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijheatmasstransfer.2020.120176.

References

[1] R.S. Jagjiwanram, Effective thermal conductivity of highly porous two-phase systems, Appl. Therm. Eng. 24 (17) (2004) 2727–2735.
[2] M. Wang, N. Pan, J. Wang, S. Chen, Mesoscopic simulations of phase distribution effects on the effective thermal conductivity of microgranular porous media, J. Colloid Interface Sci. 311 (2) (2007) 562–570.
[3] M.J. Sailor, Porous Silicon in Practice: Preparation, Characterization and Applications, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2011.
[4] J.K. Carson, S.J. Lovatt, D.J. Tanner, A.C. Cleland, Experimental measurements of the effective thermal conductivity of a pseudo-porous food analogue over a range of porosities and mean pore sizes, J. Food Eng. 63 (1) (2004) 87–95.
[5] N. Burger, A. Laachachi, M. Ferriol, M. Lutz, V. Toniazzo, D. Ruch, Review of thermal conductivity in composites: mechanisms, parameters and theory, Prog. Polym. Sci. 61 (2016) 1–28.
[6] W.-Z. Fang, H. Zhang, L. Chen, W.-Q. Tao, Numerical predictions of thermal conductivities for the silica aerogel and its composites, Appl. Therm. Eng. 115 (2017) 1277–1286.
[7] J.-W. Luo, L. Chen, T. Min, F. Shan, Q. Kang, W. Tao, Macroscopic transport properties of Gyroid structures based on pore-scale studies: permeability, diffusivity and thermal conductivity, Int. J. Heat Mass Transf. 146 (2020) 118837.
[8] M. Kaviany, Principles of Heat Transfer in Porous Media, Springer Science & Business Media, 2012.
[9] A. Gupta, A. Cecen, S. Goyal, A.K. Singh, S.R. Kalidindi, Structure–property linkages using a data science approach: application to a non-metallic inclusion/steel composite system, Acta Mater. 91 (2015) 239–254.
[10] A. Seko, H. Hayashi, K. Nakayama, A. Takahashi, I. Tanaka, Representation of compounds for machine-learning prediction of physical properties, Phys. Rev. B 95 (14) (2017) 144110.
[11] Z. Yang, Y.C. Yabansu, R. Al-Bahrani, W.-k. Liao, A.N. Choudhary, S.R. Kalidindi, A. Agrawal, Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets, Comput. Mater. Sci. 151 (2018) 278–287.
[12] Z. Yang, Y.C. Yabansu, D. Jha, W.-k. Liao, A.N. Choudhary, S.R. Kalidindi, A. Agrawal, Establishing structure-property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches, Acta Mater. 166 (2019) 335–345.
[13] Y.-J. Wu, L. Fang, Y. Xu, Predicting interfacial thermal resistance by machine learning, NPJ Comput. Mater. 5 (1) (2019) 56.
[14] X. Wan, W. Feng, Y. Wang, H. Wang, X. Zhang, C. Deng, N. Yang, Materials discovery and properties prediction in thermal transport via materials informatics: a mini review, Nano Lett. 19 (6) (2019) 3387–3395.
[15] S. Ju, J. Shiomi, Materials informatics for heat transfer: recent progresses and perspectives, Nanosc. Microsc. Thermophys. Eng. 23 (2) (2019) 157–172.
[16] Y. Wang, S. Liu, J. Cheng, X. Wan, W. Feng, N. Yang, C. Zou, A new machine learning algorithm to optimize a reduced mechanism of 2-butanone and the comparison with other algorithms , ES Mater. Manuf. 6 (0) (2019) 28–37.
[17] P. Roy Chowdhury, C. Reynolds, A. Garrett, T. Feng, S.P. Adiga, X. Ruan, Machine learning maximized Anderson localization of phonons in aperiodic superlattices, Nano Energy 69 (2020) 104428.
[18] H. Wei, S. Zhao, Q. Rong, H. Bao, Predicting the effective thermal conductivities of composite materials and porous media by machine learning methods, Int. J. Heat Mass Transf. 127 (2018) 908–916.
[19] Q. Rong, H. Wei, X. Huang, H. Bao, Predicting the effective thermal conductivity of composites from cross sections images using deep learning methods, Compos. Sci. Technol. 184 (2019) 107861.
[20] N. Lubbers, T. Lookman, K. Barros, Inferring low-dimensional microstructure representations using convolutional neural networks, Phys. Rev. E 96 (5) (2017).
[21] D. Basak, S. Pal, D.C. Patranabis, Support Vector Regression, Neural Inf. Process. 11 (10) (2007) 22.
[22] C.E. Rasmussen, H. Nickisch, Gaussian Processes for Machine Learning (GPML) Toolbox, 5.
[23] H. Xu, D.A. Dikin, C. Burkhart, W. Chen, Descriptor-based methodology for statistical characterization and 3D reconstruction of microstructural materials, Comput. Mater. Sci. 85 (2014) 206–216.
[24] G. Saheli, H. Garmestani, B.L. Adams, Microstructure design of a two phase composite using two-point correlation functions, J. Comput.-Aided Mater. Des. 11 (2–3) (2004) 103–115.
[25] Calculation of Two-Point Correlation Functions, in: Applied RVE Reconstruction and Homogenization of Heterogeneous Materials, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016, pp. 15–42.
[26] H. Xu, Y. Li, C. Brinson, W. Chen, A descriptor-based design methodology for developing heterogeneous microstructural materials system, J. Mech. Des. 136 (5) (2014) 051007-051007-051012.
[27] R. Bostanabad, Y. Zhang, X. Li, T. Kearney, L.C. Brinson, D.W. Apley, W.K. Liu, W. Chen, Computational microstructure characterization and reconstruction: review of the state-of-the-art techniques, Prog. Mater. Sci. 95 (2018) 1–41.
[28] S. Zhai, P. Zhang, Y. Xian, J. Zeng, B. Shi, Effective thermal conductivity of polymer composites: theoretical models and simulation models, Int. J. Heat Mass Transf. 117 (2018) 358–374.
[29] M. Wang, J. Wang, N. Pan, S. Chen, Mesoscopic predictions of the effective thermal conductivity for microscale random porous media, Phys. Rev. E 75 (3) (2007) 036702.
[30] T. Xie, Y.-L. He, Heat transfer characteristics of silica aerogel composite materials: structure reconstruction and numerical modeling, Int. J. Heat Mass Transf. 95 (2016) 621–635.
[31] J. Ordonez-Miranda, J.J. Alvarado-Gil, Effect of the pore shape on the thermal conductivity of porous media, J. Mater. Sci. 47 (18) (2012) 6733–6740.
[32] X. Yang, T. Lu, T. Kim, Effective thermal conductivity modelling for closed–cell porous media with analytical shape factors, Transp. Porous Media 100 (2) (2013) 211–224.
[33] F.P. Incropera, A.S. Lavine, T.L. Bergman, D.P. DeWitt, Fundamentals of Heat and Mass Transfer, Wiley, 2007.
[34] Z. Tong, M. Liu, H. Bao, A numerical investigation on the heat conduction in high filler loading particulate composites, Int. J. Heat Mass Transf. 100 (2016) 355–361.
[35] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E. Rognes, G.N. Wells, The FEniCS project version 1.5, Arch. Numer. Softw. 3 (100) (2015).
[36] J.K. Carson, S.J. Lovatt, D.J. Tanner, A.C. Cleland, Thermal conductivity bounds for isotropic, porous materials, Int. J. Heat Mass Transf. 48 (11) (2005) 2150–2158.
[37] J.C. Maxwell, J.J. Thompson, A treatise on electricity and magnetism, Clarendon (1904).
[38] D. Bruggeman, Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen. III. Die elastischen Konstanten der quasiisotropen Mischkörper aus isotropen Substanzen, Ann. Phys. 421 (2) (1937) 160–178.
[39] H. Wei, H. Bao, X. Ruan, Genetic algorithm-driven discovery of unexpected thermal conductivity enhancement by disorder, Nano Energy (2020) 104619.
[40] T. Feng, X. Ruan, Ultra-low thermal conductivity in graphene nanomesh, Carbon N Y 101 (2016) 107–113.
[41] G. Romano, J.C. Grossman, Phonon bottleneck identification in disordered nanoporous materials, Phys. Rev. B 96 (11) (2017).
[42] Q. Hao, Influence of structure disorder on the lattice thermal conductivity of polycrystals: a frequency-dependent phonon-transport study, J. Appl. Phys. 111 (1) (2012) 014309.
[43] On identifying optimal heat conduction topologies from heat transfer paths analysis, International Communications in Heat and Mass Transfer, 90 (2018) 93–102.
[44] A.P. King, R.J. Eckersley, Chapter 2 - Descriptive Statistics II: bivariate and Multivariate Statistics, in: A.P. King, R.J. Eckersley (Eds.) Statistics for Biomedical Engineers and Scientists.
[45] J.L. Myers, A. Well, N.J. Mahwah (Ed.), Lawrence Erlbaum Associates, 2003.
[46] S. Zahiri, J. Zuo, Y. Shen, H. Bao, Numerical investigation of ballistic-diffusive heat transfer through a constriction with the Boltzmann transport equation, Appl. Therm. Eng. 141 (2018) 126–133.
[47] H. Drucker, C.J.C. Burges, L. Kaufman, A. Smola, V. Vapnik, Support Vector Regression Machines, 7.
[48] K.P. Murphy, Machine Learning: A Probabilistic Perspective, MIT press, 2012.
[49] F. Pedregosa, G.l. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, C.d. Duchesnay, Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (null) (2011) 2825–2830.
[50] C.M. Bishop, Pattern Recognition and Machine Learning, springer, 2006.