Physically informed deep homogenization neural network for unidirectional multiphase/multi-inclusion thermoconductive composites

Jindong Jiang[a,1], Jiajun Wu[b,1], Qiang Chen[c,*], George Chatzigeorgiou[c], Fodil Meraghni[c]

[a] *Arts et Métiers Institute of Technology, LCFC, HESAM University, F-57070 Metz, France*
[b] *Arts et Métiers Institute of Technology, PIMM, HESAM University, F-75013 Paris, France*
[c] *Arts et Métiers Institute of Technology, CNRS, Université de Lorraine, LEM3-UMR7239, F-57000 Metz, France*

Received 11 January 2023; received in revised form 24 January 2023; accepted 25 February 2023
Available online xxxx

Abstract

Elements of the periodic homogenization framework and deep neural network were seamlessly connected for the first time to construct a new micromechanics theory for thermoconductive composites called physically informed Deep Homogenization Network (DHN). This method utilizes a two-scale expansion of the temperature field of spatially uniform composites in terms of macroscopic and fluctuating contributions. The latter is estimated using deep neural network layers. The DHN is trained on a set of collocation points to obtain the fluctuating temperature field over the unit cell domain by minimizing a cost function given in terms of residuals of strong form steady-state heat conduction governing differential equations. Novel use of a periodic layer with several independent periodic functions with adjustable training parameters ensures that periodic boundary conditions of temperature and temperature gradients at the unit cell edges are exactly satisfied. Automatic differentiation is utilized to correctly compute the fluctuating temperature gradients. Homogenized properties and local temperature and gradient distributions of unit cells reinforced by unidirectional fiber or weakened by a hole are compared with finite-element reference results, demonstrating remarkable correlation but without discontinuities associated with temperature gradient distributions in the finite-element simulations. We also illustrate that the DHN enhanced with transfer learning provides a substantially more efficient and accurate simulation of multiple random fiber distributions relative to training the network from scratch.

1. Introduction

An objective of micromechanics is to predict the local and homogenized response of composite and heterogeneous materials from the knowledge of mechanical and physical properties of the individual phases, in conjunction with the microstructure characteristics such as the arrangement, shape and orientation of the inclusions [1]. The advantages of micromechanics are several-fold. The main advantage is that, once the constituent phase properties

* Corresponding author.
E-mail addresses: qiang.chen@ensam.eu, kirkcq@gmail.com (Q. Chen).
[1] These authors contributed equally to this work.

are known, micromechanics allows the prediction of the multi-axial and non-proportional anisotropic response (both elastic and inelastic) of composites that are often difficult to measure experimentally, hence the cost of laborious experimental campaigns of composites is significantly reduced [2]. Furthermore, micromechanics simulations accelerate the development and discovery of engineered materials with target thermo-mechanical and physical properties for a specific application by selecting candidate materials through a reverse engineering technique. Another important but thus-far underrated role micromechanics may play is the analysis and optimization of composite materials or structural components in a multi-scale setting when incorporated into structural analysis tools [3–6].

The central idea of micromechanics theories for heterogeneous media is the determination of the localization relations, which relate the local stresses or strains per phase to the macroscopic (or applied) stress or strain states. In general, micromechanics techniques may be divided into three broad categories, cf. Pindera et al. [7]. The first category includes microstructural-detail-free micromechanics models, which are based on various simplifications of the material's microstructure [8,9]. Solutions for displacement and stress fields of an inclusion embedded in an effective medium, or in a matrix that is further embedded as a whole in an effective medium, are derived to obtain the closed-form expressions of effective moduli of the material-at-large. These approaches are commonly called the self-consistent scheme [10], the three-phase model [11], and the Mori–Tanaka method [12,13].

Nonperiodic and periodic micromechanics methodologies accounting for explicit geometric representations of material microstructures have also been developed to characterize statistically homogeneous and periodic composites, respectively. The nonperiodic homogenization schemes are based on the representative volume element (RVE) concept. Effective moduli of the composites are determined by applying homogeneous displacement or traction boundary conditions, along with the satisfaction of the equivalence of these boundary conditions. The composite cylinder/sphere assemblage (CCA/CSA) proposed by Hashin and Rosen [14] is perhaps the best-known in this category. Another well-known RVE-based model is the equivalent inhomogeneity technique proposed by Mogilevskaya et al. [15].

Alternatively, the periodic homogenization schemes are based on the repeating unit cell (RUC) concept. The RUC is the basic building block that can be replicated in two or three dimensions to form the entire materials. The effective or homogenized moduli of the entire array under macroscopic loading are identical to those of an arbitrary RUC subjected to periodic boundary conditions. The finite-element method is the prevailing technique because of its ability to model complex microstructures with multiple and arbitrarily shaped inclusions [16–18]. The finite-volume direct averaging micromechanics scheme and its various forms are yet another technique for the analysis of heterogeneous materials [19–21]. The key feature of the finite-volume method is the satisfaction of the governing differential equations in an averaged sense in the discretized RUC domain. However, both finite-element and finite-volume techniques require the discretization of fiber and matrix phases of the unit cell in the determination of displacement and stress fields, an issue that leads to significant computational costs in the case of unit cells with multiple random distributions of fibers that require large discretization. The locally exact homogenization technique (LEHT), developed by Pindera and his coworkers [22,23], employs exact solutions of governing differential equations via Fourier series representation of displacement fields and the balanced variational principle to enforce periodic boundary conditions, rather than unit cell discretization used in the finite-element and finite-volume technique. An important limitation of the LEHT is that only single-inclusion unit cells in square and hexagonal arrays can be considered. In response to this issue, a hybrid homogenization technique was proposed to model multiple random fiber distributions [24]. Eigenstrain-based periodic homogenization techniques for multiphased and multi-inclusion periodic composites have also been developed by Lages and Marques [25,26]. Nevertheless, simulating the random fiber distributions has not been reported yet either. The readers are referred to the review articles provided by Pindera et al. [7], Saeb et al. [27], and Chen et al. [28] for additional references in this area.

In recent years, deep neural networks (DNN) have found vast applications in the modeling of elastic and inelastic behaviors of materials due to the potential benefits offered by them. In general, deep neural networks aim to establish a hypothesis function using a series of network layers. These layers are connected via a sequence of linear matrix multiplication and nonlinear mapping with activation functions in order to obtain the underlying relationship between the input–output data pairs. The DNN has been demonstrated to be a successful approach in the micromechanical analysis of lattice structures [29], unidirectional composites [30] and short fiber reinforced composites [31,32], nanocomposites [33], etc.

The physically informed deep neural network (PINN) model, proposed by Raissi et al. [34], has emerged in recent years as a viable alternative to the classical numerical solution of the partial differential equations with prescribed

boundary conditions (such as the finite-element, finite difference or finite-volume techniques). The contributions of Rabczuk and his coworkers [35–38] have spurred extensive applications of PINN theory in the solution of boundary value problems in mechanics. Different from the conventional techniques for solving Navier's equations analytically or numerically, the machine learning approaches use multiple neural network layers to look for functions that satisfy the governing equations and the specified boundary conditions by minimizing the cost functions [39–43].

Herein, for the first time, elements of the periodic homogenization framework and deep neural network were employed to construct a new micromechanics theory for thermoconductive periodic composites called physically informed deep homogenization neural network. In the case of the unit cell problem, the DHN developed herein aims to look for the fluctuating temperature field that satisfies the strong form heat condition equation and the periodicity boundary conditions, which differs from any of the aforementioned periodic micromechanics schemes. Specifically, the DHN deals directly with the minimization of cost functions expressed in terms of the averaged PDE residuals, which are sampled randomly on a set of collocation points. Therefore, this method does not rely on conventional computationally extensive mesh discretizations. It should be emphasized that, in the micromechanics theory, periodicity boundary conditions play a vital role in yielding accurate unit cell solutions. The often-used penalty method in the PINN approach may easily cause significant errors in the unit cell solution because the boundary conditions can be only enforced approximately [30]. Therefore, a method for implementing exactly the periodic boundary, developed by Dong and Ni [44], is adapted to the DHN framework. The main contributions of the present manuscript include:

- construction of a novel physically informed deep neural network model for the micromechanical analysis of composites in the periodic homogenization framework
- implementation of exact periodic boundary conditions to infinite order at the unit cell boundaries upon the use of a periodic layer with a set of periodic functions
- demonstration that the proposed network model, enhanced with transfer learning, is capable of predicting accurately the local and homogenized response of composites with multiple random inclusions

We note that an alternative approach for solving the unit cell problem with periodicity boundary conditions is the proper generalized decomposition (PGD) method [45]. The PDG method, however, employs the rectangular discretization of the unit cell hence it may not be suited for dealing with inclusion with curved surfaces. It should be also mentioned that the physics-informed deep neural network model developed in this work attempts to minimize a loss function that is tailored to the underlying differential operator, which is significantly different from purely data-driven approaches where the machine learning techniques are merely used as a black box [46]. Two related advantages are the generalizability of the trained network model for the unseen training data and the rapid convergence of the neural network with a small training dataset.

The rest of the manuscript is organized as follows: Section 2 outlines the theoretical framework for the zeroth-order periodic homogenization for thermoconductive composites and its finite-element implementation. Section 3 describes the physically informed deep homogenization network developed in this investigation that takes into account explicitly the periodic boundary conditions required by periodic homogenization. Section 4 presents the comparison of the predictive capabilities of the developed homogenization scheme and the finite-element method vis-à-vis the local temperature and temperature gradient distributions and effective thermal conductive coefficient for unidirectional fiber or porous composites. The effect of network parameters and training dataset size was studied extensively and the best neural network model was identified. Section 5 illustrates the capabilities of the developed neural network model, enhanced with transfer learning, for simulating the composites with multiple random inclusions. A discussion of the proposed technique is presented in Section 6. Section 7 draws the pertinent conclusions.

2. Periodic homogenization theory for thermoconductive composites

The scope of this section is to define the problem under consideration. Hereafter, some preliminary notes concerning the homogenization theory of thermo-conductive composites and its finite-element formulation are described.

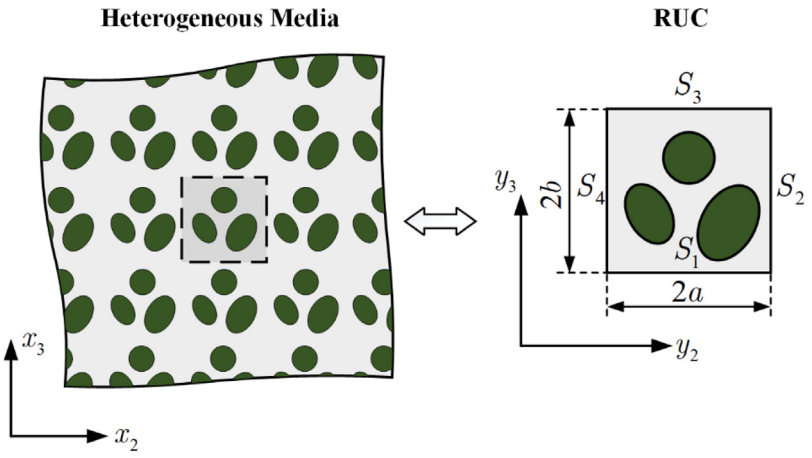


Fig. 1. A multiphase periodic array characterized by the smallest building block or the unit cell.

2.1. Preliminaries

Let us consider a heterogeneous solid reinforced by unidirectional fibers which are embedded in a matrix phase. If the fibers are periodically dispersed in the matrix, the overall response of the heterogeneous media is identical to that of a repeating unit cell (RUC) subjected to periodic boundary conditions, as shown in Fig. 1. The latter can be replicated in two dimensions to form the entire heterogeneous materials.

Following the zeroth-order homogenization theory [6,26,47–49], the local variations of temperature can be described using a two-scale expansion involving the global and microscopic coordinates, $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$, representing the averaged $\overline{T}(\mathbf{x})$ and microstructure-induced fluctuating $\widetilde{T}(\mathbf{y})$ contributions, respectively:

$$T(\mathbf{x}, \mathbf{y}) = \overline{T}(\mathbf{x}) + \widetilde{T}(\mathbf{y}) \tag{1}$$

In the above equation, $\overline{T}(\mathbf{x}) = \overline{H}_i x_i \ (i = 1, 2, 3)$, $\overline{H}_i$ denotes the macroscopic temperature gradient. $\widetilde{T}(\mathbf{y})$ is a periodic function in $y_2 - y_3$ plane. Accordingly, the local temperature gradients are expressed in terms of average and fluctuating contributions as follows:

$$H_1 = \overline{H}_1, H_2 = \overline{H}_2 + \widetilde{H}_2 = \overline{H}_2 + \frac{\partial \widetilde{T}}{\partial y_2}, H_3 = \overline{H}_3 + \widetilde{H}_3 = \overline{H}_3 + \frac{\partial \widetilde{T}}{\partial y_3} \tag{2}$$

It should be noted that $\partial \widetilde{T}/\partial y_1 = 0$ since there is no fast variation in the fiber longitudinal direction.

Under steady-state conditions, the heat flux $\mathbf{q}$ is related to the temperature gradient using Fourier's law as:

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = - \begin{bmatrix} k_{11} & 0 & 0 \\ 0 & k_{22} & 0 \\ 0 & 0 & k_{33} \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} \tag{3}$$

where k_{11}, k_{22} and k_{33} denote the thermal conductivity coefficients for orthotropic materials.

The steady-state heat conduction governing differential equation expressed in terms of fluctuating temperature reads:

$$\frac{\partial}{\partial y_2} k_{22} \left(\overline{H}_2 + \frac{\partial \widetilde{T}}{\partial y_2} \right) + \frac{\partial}{\partial y_3} k_{33} \left(\overline{H}_3 + \frac{\partial \widetilde{T}}{\partial y_3} \right) = 0 \tag{4}$$

For a unit cell depicted in Fig. 1, Eq. (4) is solved subject to the periodicity boundary condition of temperature and its gradients:

$$\widetilde{T}(-a, y_3) = \widetilde{T}(a, y_3), \qquad \widetilde{T}(y_2, -b) = \widetilde{T}(y_2, b) \tag{5}$$

$$\frac{\partial}{\partial y_2}\widetilde{T}\left(-a, y_3\right) = \frac{\partial}{\partial y_2}\widetilde{T}\left(a, y_3\right), \qquad \frac{\partial}{\partial y_3}\widetilde{T}\left(y_2, -b\right) = \frac{\partial}{\partial y_3}\widetilde{T}\left(y_2, b\right) \tag{6}$$

where $\forall y_2 \in [-a, a]$ and $\forall y_3 \in [-b, b]$. It should be noted that Eq. (6) ensures that the periodicity of the normal heat flux density along the unit cell faces in the pertinent direction is satisfied:

$$Q\left(S_1\right) + Q\left(S_3\right) = 0, \quad Q\left(S_2\right) + Q\left(S_4\right) = 0 \tag{7}$$

where $Q\left(S_i\right) = \left[q_2 \cdot n_2 + q_3 \cdot n_3\right]^{(i)}$. $\left(n_2, n_3\right)^{(i)}$ denotes the unit normal vector of the ith unit cell face.

Finally, following the classical homogenization theory, the homogenized constitutive equations of the unit cell with a volume V, under steady-state conditions, can be expressed in terms of effective constants, designated by an asterisk, Eq. (8). This equation relates the macroscopic temperature gradient $\overline{\mathbf{H}}$ to the macroscopic heat flux density $\overline{\mathbf{q}}$ evaluated as volume-averaging of their corresponding local quantities as follows:

$$\overline{\mathbf{q}} = -\mathbf{k}^* \cdot \overline{\mathbf{H}} \tag{8}$$

where

$$\overline{\mathbf{q}} = \frac{1}{V}\int_V \mathbf{q}\mathrm{d}V, \quad \overline{\mathbf{H}} = \frac{1}{V}\int_V \mathbf{H}\mathrm{d}V \tag{9}$$

2.2. Finite-element homogenization

In the finite-element framework [19,50,51], the fluctuating temperature field $\widetilde{T}^{(q)}$ can be expressed in terms of nodal fluctuating temperature and shape function as follows:

$$\widetilde{T}^{(q)} = \sum_{p=1}^{n_p} N_p\left(\eta, \xi\right)\widetilde{T}^{o(p,q)} \tag{10}$$

where $\widetilde{T}^{o(p,q)}$ denotes the nodal fluctuating temperature at the pth node of the qth element. $N_p\left(\eta, \xi\right)$ represents the shape function and n_p is the total number of modes in the qth subvolume. Applying the differential operator $\partial = (\partial/\partial \mathbf{y})$ to the fluctuating temperature $\widetilde{T}^{(q)}$ expressed in terms of nodal temperatures using the matrix notation $\widetilde{T}^{(q)} = \mathbf{N}^{(q)}\widetilde{\mathbf{T}}^{o(q)}$, the fluctuating temperature gradient $\widetilde{\mathbf{H}}^{(q)}$ is obtained in the form:

$$\widetilde{\mathbf{H}}^{(q)} = \partial\mathbf{N}^{(q)}\widetilde{\mathbf{T}}^{o(q)} = \mathbf{B}^{(q)}\widetilde{\mathbf{T}}^{o(q)} \tag{11}$$

where $\widetilde{\mathbf{T}}^{o(q)} = \left[\widetilde{T}^{o(1,q)}, \ldots, \widetilde{T}^{o\left(n_p,q\right)}\right]^{\mathrm{T}}$. $\mathbf{B}^{(q)}$ indicates the temperature–temperature gradient matrix,

$$\mathbf{B}^{(q)} = \left[\partial\widetilde{N}_1, \ldots, \partial\widetilde{N}_p, \ldots \partial\widetilde{N}_{n_p}\right]^{(q)} \tag{12}$$

with

$$\partial\widetilde{N}_p = \left[0, \frac{\partial N_p}{\partial y_2}, \frac{\partial N_p}{\partial y_3}\right]^{\mathrm{T}}, \tag{13}$$

which is used to construct the potential energy integral at the element level.

For a thermally conductive unit cell with a volume V, the total potential energy can be expressed as:

$$\pi = \frac{1}{2}\int_V \overline{\mathbf{H}}^{\mathrm{T}} : \mathbf{k} : \overline{\mathbf{H}}\mathrm{d}V + \frac{1}{2}\int_V \widetilde{\mathbf{H}}^{\mathrm{T}} : \mathbf{k} : \widetilde{\mathbf{H}}\mathrm{d}V + \int_V \widetilde{\mathbf{H}}^{\mathrm{T}} : \mathbf{k} : \overline{\mathbf{H}}\mathrm{d}V - \overline{\mathbf{q}}^{\mathbf{T}} : \overline{\mathbf{H}} \cdot V \tag{14}$$

Substituting Eq. (11) into Eq. (14) and minimization of the total potential energy with respect to the nodal temperatures, $\partial\pi/\partial\widetilde{\mathbf{T}}^o = 0$, yield the global system of equations for the common unknown nodal temperatures:

$$\mathbb{K}\widetilde{\mathbf{T}}^o = \overline{\mathbb{F}} \tag{15}$$

where

$$\mathbb{K} = \int_V \mathbf{B}^{\mathbf{T}} : \mathbf{k} : \mathbf{B}\mathrm{d}V, \overline{\mathbb{F}} = -\left(\int_V \mathbf{B}^{\mathbf{T}} : \mathbf{k}\mathrm{d}V\right)\overline{\mathbf{H}} \tag{16}$$

The local conductivity matrices are assembled to form a global system of equations by enforcing the continuities of fluctuating temperatures at common nodes of adjacent elements, as well as the periodicity conditions at the mirrored faces of a repeating unit cell.

3. Physically informed deep homogenization network

The central idea of the physically informed deep homogenization network is to approximate a continuous function that maps the space coordinates (y_2, y_3) to the fluctuating temperature $\widetilde{T}$:

$$\widetilde{T} = \mathrm{DHN}\left(y_2, y_3\right), \forall \left(y_2, y_3\right) \in V \tag{17}$$

for periodic arrays whose microstructural information is incorporated by the spatial dependency of the local conductivities $k_{22}\left(y_2, y_3\right)$ and $k_{33}\left(y_2, y_3\right)$. The unit cell for the periodic arrays is subjected to given macroscopic temperature gradients $\overline{H}_2$ and $\overline{H}_3$.

A feedforward network is constructed by connecting neurons between layers. The parameters of the neural network $\boldsymbol{\theta}^*$ are determined by minimizing a loss function $\mathcal{L}\left(\widetilde{T}\right)$ with contributions from the PDE residual $\mathcal{L}_{PDE}$ and the penalties representing the periodicity boundary conditions, $\mathcal{L}_{bc1}$ and $\mathcal{L}_{bc2}$,

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\mathrm{ArgMin}}\ \mathcal{L}\left(\widetilde{T}\right) = \mathcal{L}_{PDE} + \gamma_1 \mathcal{L}_{bc1} + \gamma_2 \mathcal{L}_{bc2} \tag{18}$$

where

$$\mathcal{L}_{PDE} = \frac{1}{V}\int_V \left[\frac{\partial}{\partial y_2} k_{22}\left(\overline{H}_2 + \frac{\partial \widetilde{T}}{\partial y_2}\right) + \frac{\partial}{\partial y_3} k_{33}\left(\overline{H}_3 + \frac{\partial \widetilde{T}}{\partial y_3}\right)\right]^2 \mathrm{d}V \tag{19}$$

$$\mathcal{L}_{bc1} = \frac{1}{2a}\int_{-a}^{a}\left[\widetilde{T}\left(y_2, -b\right) - \widetilde{T}\left(y_2, b\right)\right]^2 \mathrm{d}y_2 + \frac{1}{2b}\int_{-b}^{b}\left[\widetilde{T}\left(-a, y_3\right) - \widetilde{T}\left(a, y_3\right)\right]^2 \mathrm{d}y_3 \tag{20}$$

$$\mathcal{L}_{bc2} = \frac{1}{2a}\int_{-a}^{a}\left[\frac{\partial}{\partial y_3}\widetilde{T}\left(y_2, -b\right) - \frac{\partial}{\partial y_3}\widetilde{T}\left(y_2, b\right)\right]^2 \mathrm{d}y_2 + \frac{1}{2b}\int_{-b}^{b}\left[\frac{\partial}{\partial y_2}\widetilde{T}\left(-a, y_3\right) - \frac{\partial}{\partial y_2}\widetilde{T}\left(a, y_3\right)\right]^2 \mathrm{d}y_3 \tag{21}$$

In the above equations, γ_1 and γ_2 denote the penalty coefficients, which play an indispensable role in yielding a good neural network solution. Nonetheless, the enforcement of the periodicity conditions in the penalty method is still only approximate.

Herein, in order to exactly enforce the periodicity boundary conditions given by Eqs. (5) and (6), we employ the technique developed by Dong and Ni [44] for representing periodic functions. This method is based on the representation of arbitrary functions with a set of independent sinusoidal functions with adjustable (training) parameters using a neural network layer. For a periodic function in $y_2 - y_3$ plane with periods of $2a$ and $2b$ in y_2 and y_3 directions, respectively, the periodic layer is defined as:

$$\begin{aligned}
v_{2i}\left(y_2\right) &= \sigma\left[A_{2i}\cos\left(\omega_2 y_2 + \phi_{2i}\right) + c_{2i}\right], \quad 1 \le i \le m \\
v_{3i}\left(y_3\right) &= \sigma\left[A_{3i}\cos\left(\omega_3 y_3 + \phi_{3i}\right) + c_{3i}\right], \quad 1 \le i \le m \\
q_j\left(y_2, y_3\right) &= \sigma\left[\sum_{i=1}^{m} v_{2i}\left(y_2\right) W_{ij}^{(2)} + \sum_{i=1}^{m} v_{3i}\left(y_3\right) W_{ij}^{(3)} + B_j\right], \quad 1 \le j \le n
\end{aligned} \tag{22}$$

where

$$\omega_2 = \frac{\pi}{a}, \quad \omega_3 = \frac{\pi}{b} \tag{23}$$

are constants with prescribed periods $2a$ and $2b$, respectively. $\sigma\left[\cdot\right]$ represent the nonlinear activation function. The nonlinear activation function ensures that $v_{2i}\left(y_2\right)$ and $v_{3i}\left(y_3\right)$ contain not only the frequencies ω_2 and ω_3, but also components with higher frequencies with common periods in the pertinent direction. m and n are hyper-parameters of the periodic layer and q_j denotes the output of this layer. The training parameters include:

$$A_{2i}, A_{3i}, \phi_{2i}, \phi_{3i}, c_{2i}, c_{3i}, W_{ij}^{(2)}, W_{ij}^{(3)}, B_j \tag{24}$$

where $1 \le i \le m$, $1 \le j \le n$.

The output of the periodic layer then passes through multiple sequential residual learning blocks and is finally summed up in the output layer to calculate the fluctuating temperature $\widetilde{T}$, as shown in Fig. 2. The operations in the residual block are defined by:

$$q_j^l = \sigma\left(\sum_{i=1}^{n} W_{ij}^l q_j^{l-1} + B_j^l + q_j^{l-1}\right) \tag{25}$$

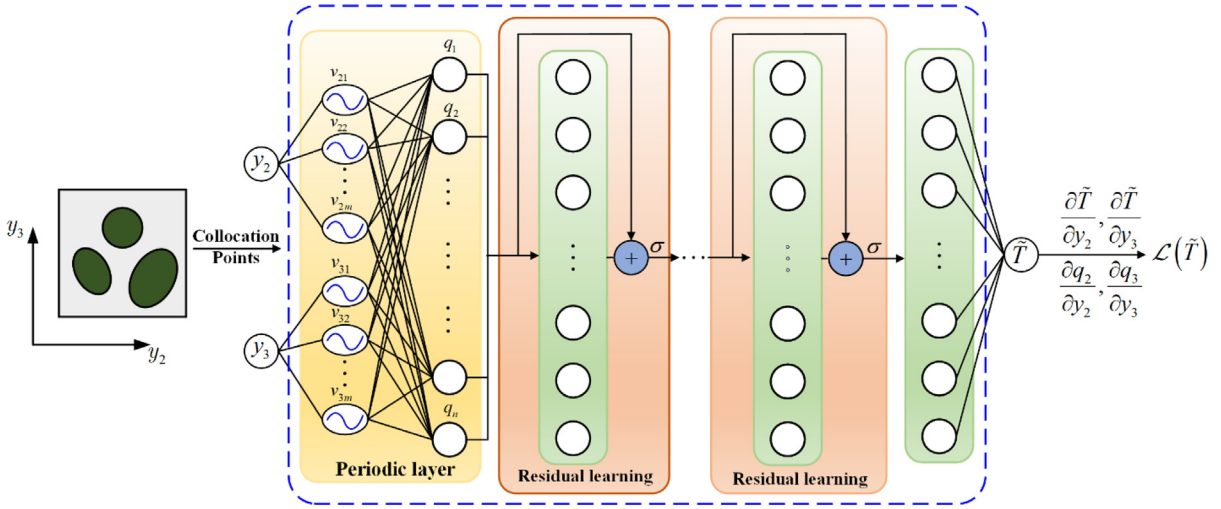


Fig. 2. Physically informed deep neural network acting as surrogate homogenization model for periodic microstructured materials.

Within each residual block, the input from the previous layer/block q_j^{l-1} is added up to the output of a weighted fully connected layer before the application of the activation function to form the residual block output q_j^l. This direct mapping of block input to output, also known as skip connection or shortcut, can help to propagate input information through a complex network, avoid gradient vanishing problems and present better convergence performance compared to classical network structures with concatenated fully connected layers. Note that for the simplicity of the network structure definition, the dimension of the hidden layer in the residual block has been kept the same as in the periodic layer (n).

The periodic layer defined in Eq. (22) permits the exact satisfaction of periodicity boundary conditions of fluctuating temperature to infinite order. As such, the penalty terms Eqs. (20) and (21) can be omitted, further facilitating training and convergence of the neural network model. In summary, the DHN-based neural network solution is accomplished in several steps:

(1) Selecting a microstructure and macroscopic temperature gradient;

(2) Decomposing the temperature field into averaged and fluctuating contributions using Eq. (1);

(3) Generating collocation points and building the neural network layers;

(4) Evaluating the loss function given by Eq. (19). The periodicity boundary conditions are automatically satisfied through Eqs. (22) and (23);

(5) Computing the temperature gradients Eq. (2) and heat flux distributions Eq. (3) after the loss function is minimized;

(6) Computing the homogenized properties Eq. (8).

4. Numerical results

In this section, to demonstrate the predictive capabilities of the neural network-based micromechanics model developed in this work, the deep homogenization network approach is employed to predict the temperature and temperature gradient distributions of unit cells containing cylindrical fibers or porosity subjected to periodic boundary conditions. They are extensively verified against unit cell solutions generated by the in-house finite-element homogenization technique. For the sake of simplicity, both fiber and matrix are considered isotropic and their thermal conductive coefficients are $k^f = 40$ $\mathrm{Wm^{-1}K^{-1}}$ and $k^m = 10$ $\mathrm{Wm^{-1}K^{-1}}$, respectively.

It should be mentioned that an important feature of functions to be estimated through the periodic neural network approach is that while the fluctuating temperature is continuous within the entire unit cell domain, the temperature gradients are discontinuous, which marks the boundary of the fibers due to the jump of fiber/matrix thermal conductivity coefficients. Since the neural network solution is infinitely differentiable, it cannot provide an exact representation of the functions with discontinuities [30,52]. Therefore, we seek a reasonable approximation

Table 1
Network parameters for choosing the best model.

Network	Neurons in periodic layer	Hidden layers (excluding the first periodic layer)	Hidden neurons	Trainable parameters
1	10	3	20	1761
2	30	3	20	2681
3	20	2	20	1801
4	20	3	10	871
5	20	3	20	2221
6	20	3	30	4171
7	20	5	10	1091
8	20	5	20	3061
9	20	5	30	6031

of fiber/matrix thermal conductivity transition. For an N-phase composite, the coordinates of the center of the ith fiber are given by $(o_2, o_3)^i$. The thermal conductivities are artificially smoothened as:

$$k_{22} = k_{33} = \sum_{i=1}^{N} c_1 \left[c_2 + \tanh \frac{R - \sqrt{\left(y_2 - o_2^i\right)^2 + \left(y_3 - o_3^i\right)^2}}{\delta} \right] + c_3 \tag{26}$$

In the above equation, if c_2 is set to 1 directly, the matrix thermal conductivities are equal to $k_{22}^m = k_{33}^m = k^m = c_3$ whereas for the fiber $k_{22}^f = k_{33}^f = k^f = 2c_1 + c_3$. R denotes the fiber radius. δ denotes the smoothness of the material transition. In a rough but successful approximation, $\delta = 0.05R$ will be used in the following simulation unless otherwise stated.

4.1. Effect of network hyperparameters

The input dataset size and neural network hyperparameters, namely, the number of hidden layers, the number of neurons per layer and the activation function, play a critical role in yielding an accurate neural network solution. To approximate $\widetilde{T}(y_2, y_3)$, we employ a feed-forward deep neural network as shown in Fig. 2. The latter takes a set of spatial coordinates (y_2, y_3) and outputs the fluctuating temperature $\widetilde{T}(y_2, y_3)$ distributions that satisfy exactly the periodicity boundary conditions of the zeroth-order homogenization theory. As described in Section 3, the second network layer is the periodic layer. In order to highlight the effect of network hyper-parameters on network performance, we consider 9 neural network architectures which are built with different network depths and hidden neurons to cover a wide range of parameter space as summarized in Table 1. The hyperbolic tangent function $\sigma(x) = \tanh(x)$ is always used as an activation function except for the output layer where no activation function is applied.

As illustrated in Fig. 3, the first investigated problem is a unidirectionally-oriented composite containing 20% fiber volume fraction, with fiber arranged in a square array, which is subjected to a macroscopic temperature gradient by $\overline{H}_2 = 1$ K/m. The remarkable property mismatch between the constituent phases yields large variations of temperature and temperature gradient in the vicinity of the fibers (interphase), hence providing a very demanding test of the correctness and robustness of the neural network approach. A dataset with 10k randomly-distributed collocation points was generated using the Monte-Calo simulation, with the ranges of each coordinate in $y_2, y_3 \in [-0.5, 0.5]$. A test dataset with 280×280 unseen grid data was utilized to test the network model. The neural network was trained on the Google Colab with GPU acceleration, which provides Intel Xeon CPU @ 2.00 GHz with 1 core, 2 threads, 13 GB RAM and one Tesla T4 GPU of 16 GB RAM, 2560 CUDA Cores @1250 MHz. The ADAM (adaptive momentum) optimizer is utilized to minimize the loss function expressed in Eq. (19). The learning rate is prescribed as 0.01 and decays by a factor of 0.5 for each 1000 epoch. For each network architecture, three training runs were performed with different random seed values for neural network initialization, to better evaluate the overall performance of each neural network architecture.

The training progress for each network model at multiple restarts is illustrated in Fig. 4. The pointwise loss distributions for each network model with the lowest overall loss values are compared in Fig. 5. It is evident that

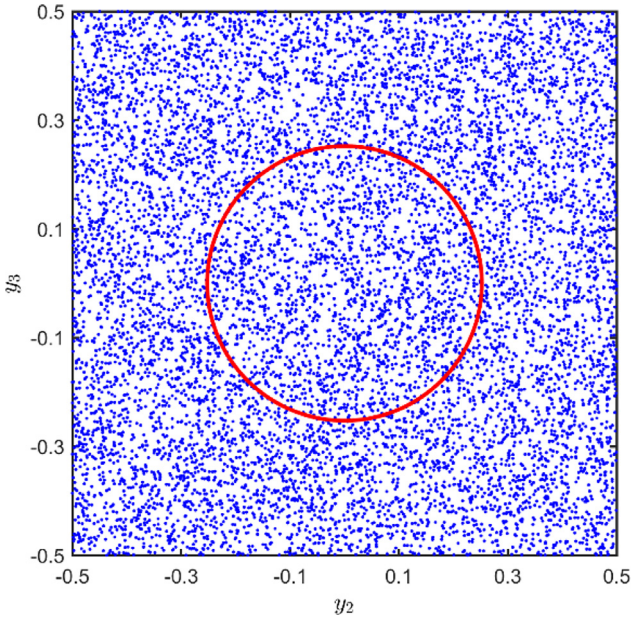


Fig. 3. Collocation point distributions generated using Monte-Carlo simulation for evaluating PDE residuals.

the loss functions for networks 4, 5, 6, 8, and 9 at multiple restarts are less likely to be affected by variations in the seed values for the neural network weights and biases, while the repeatability of the loss functions for networks 1, 2, 3 and 7 is inferior to the former ones. A close examination of Figs. 4 and 5 reveals that networks 5, 6, 8, and 9 yield relatively lower overall loss value and loss distributions. In contrast, networks 1 3, 4, and 7 predict significant loss values in the vicinity of the fiber/matrix interface where the temperature gradients are the highest due to the fiber/matrix property mismatch. Therefore, networks 5, 6, 8, and 9 retain the most robust candidates for the considered unit cell solution. Network 8 converges very fast relative to networks 5, 6, and 9 with relatively smaller training parameters. In what follows, network 8 is chosen for conducting all the simulations in the sequel.

Accurate characterization of temperature field distributions and concomitant identification of maximum temperature that occurs along the fiber/matrix interface in composites play a key role in developing failure criteria for this type of material. Fig. 6 presents comparison of the full-field fluctuating temperature $\widetilde{T}$ and temperature gradients $\widetilde{H}_2$ and $\widetilde{H}_3$ distributions generated by the physically informed DHN and 8-noded FEM methods under transverse thermal loading by $\overline{H}_2 = 1$ K/m. To quantify errors between the deep neural network and finite-element predictions, the differences between the two approaches are also enclosed in Fig. 6. It should be emphasized that DHN predictions were evaluated on a separate and hold-out dataset with 280×280 grid points that were not used in the training process. As observed, the fluctuating temperature distributions predicted by the deep neural network and finite-element techniques show a high level of accordance. Both characters and magnitudes are accurately captured by the network-based approach with negligible differences in the entire unit cell domain. While the temperature gradients predicted by the DHN approach are smoothly varying in the vicinity of the fiber/matrix interface region, the finite-element results exhibit important discontinuities as observed in Fig. 6(b) and (c) despite that the corresponding temperature fields are continuous. This is because the conventional finite-element method can only satisfy the C^0 continuity of the field variables, while the calculated temperature gradients are not necessarily continuous from one element to another. In contrast, the neural network-based results are infinitely differentiable, satisfying exactly the continuities of temperature and its gradients. It is worth noting that the solution principles between the finite-element and deep neural network approaches are fundamentally different. While the minimization of the total potential energies within the finite-element framework leads to the ultimate satisfaction of the unit cell's global conservation with sufficient mesh refinement (weak-form solution), the DHN is based on the direct minimization of the PDE loss function (strong-form solution), further lending credence to the deep neural network method's rigorous validation and the ensuing conclusions.

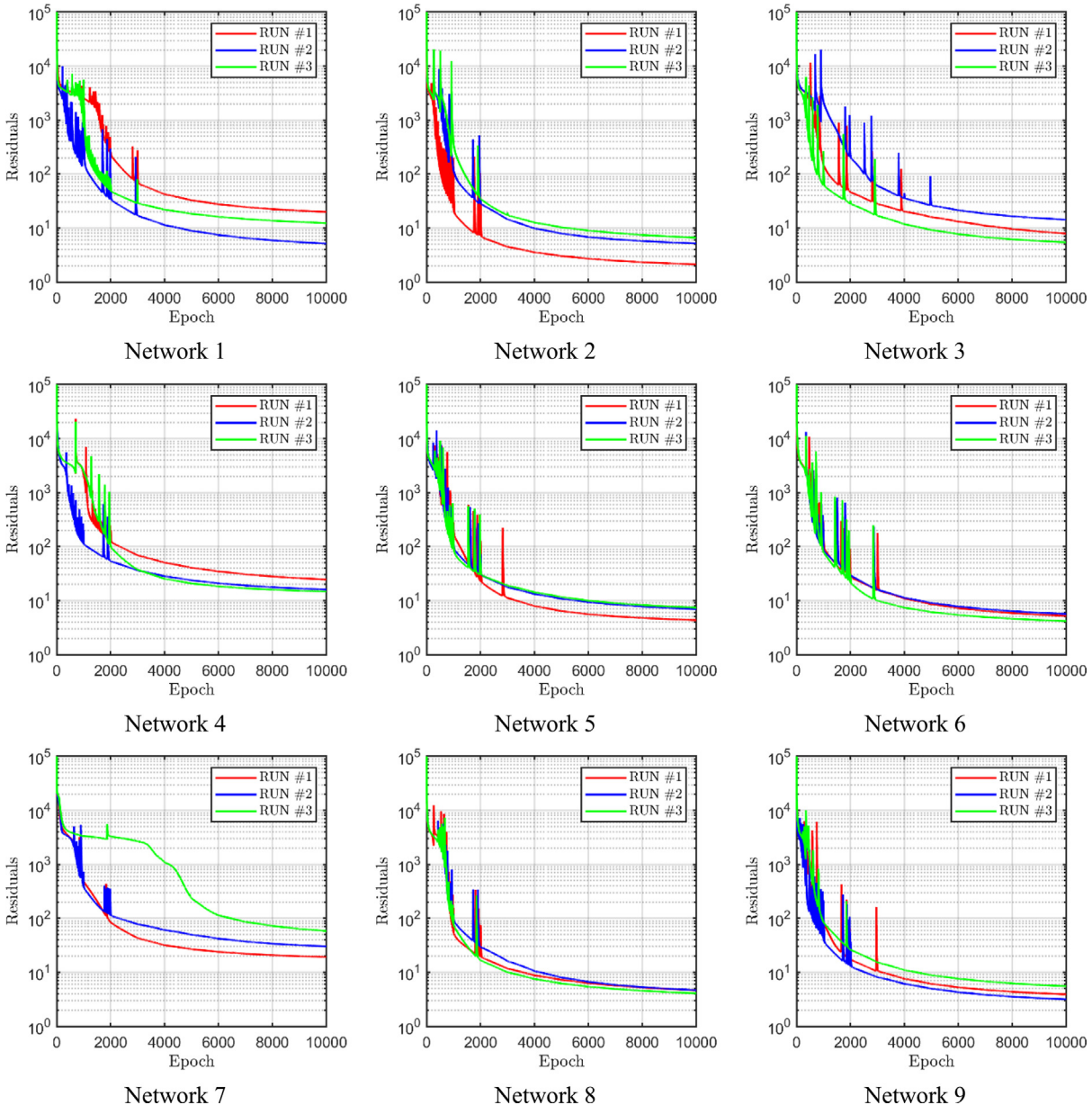


Fig. 4. Effect of network depth and hidden neurons on the network performance.

4.2. Effect of dataset size

To demonstrate the effect of training dataset size on the neural network performance, the best network model (Network 8) described in Section 4.1 is employed to predict the full-field temperature and its gradient distributions with three different dataset sizes, namely 2k, 5k, and 15k which were generated with Monte-Carlo simulations. Fig. 7 illustrates the corresponding performance plots for the selected dataset size for three training attempts. The training loss function for the dataset with 10k points is also enclosed in Fig. 7 for comparison. Generally speaking, while the computational cost increases with the size of the dataset, the loss value is lower when more data points have been used in the training process. It is also clear that the networks trained with 2k and 5k data points are more likely to be affected by the seed values for the weights and biases relative to the networks trained with 10k and

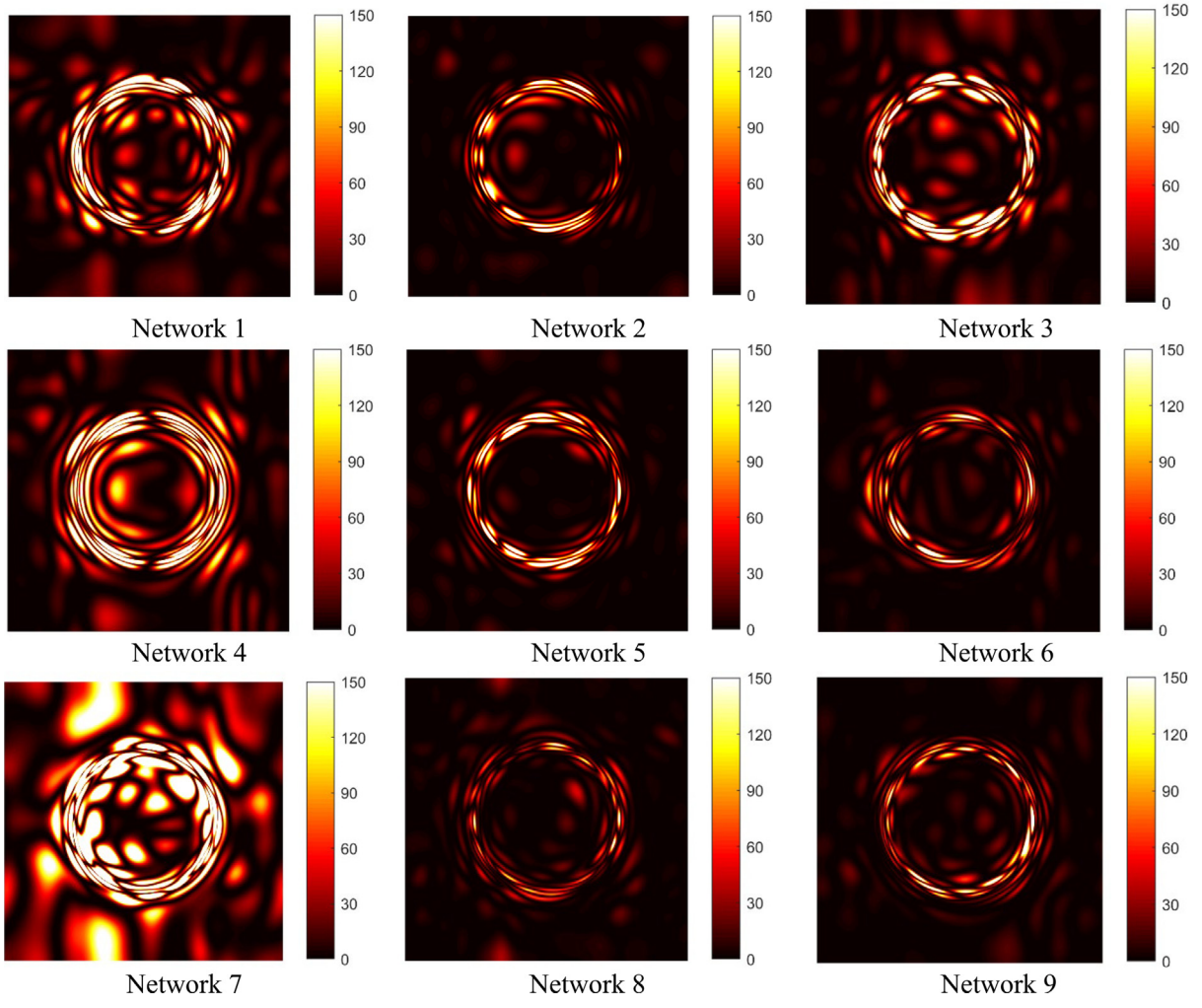


Fig. 5. Comparison of the PDE residual distributions for different network models.

15k data points at different training attempts. The loss functions obtained from 10k and 15k training data points are roughly the same with the ultimate loss value of 10k data points even lower than that of 15k data points.

Fig. 8 compares the transverse fluctuating temperature gradient $\widetilde{H}_2$ distributions for the best neural network model with different dataset sizes. As observed, the differences between the network results with different dataset sizes are invisible. It is remarkable how well the DHN approach is capable of capturing the local field concentration with no signs of overfitting even when using only 2k training data points. Therefore, the small dataset can be used for fast check of unit cell solution of the heterogeneous microstructures while the large dataset can be used when a reliable unit cell solution is pursued, albeit at the cost of computational resources. It should be noted that overfitting is a common issue in the data-driven approach where the model memorizes the data patterns in the training data but fails to generalize to unseen data. As pointed out by Bajaj et al. [53], the DHN may also encounter an overfitting issue when the number of parameters for a DHN architecture is much larger than the number of training points. The present study shows that consistent and reliable results can be obtained even when only 2k collocation points were employed in the training process. Therefore, for the network with 3061 trainable parameters, the results obtained from 10k training collocation points, which are sufficiently large to avoid high variance and bias in the results generated, can be considered free from overfitting issues.

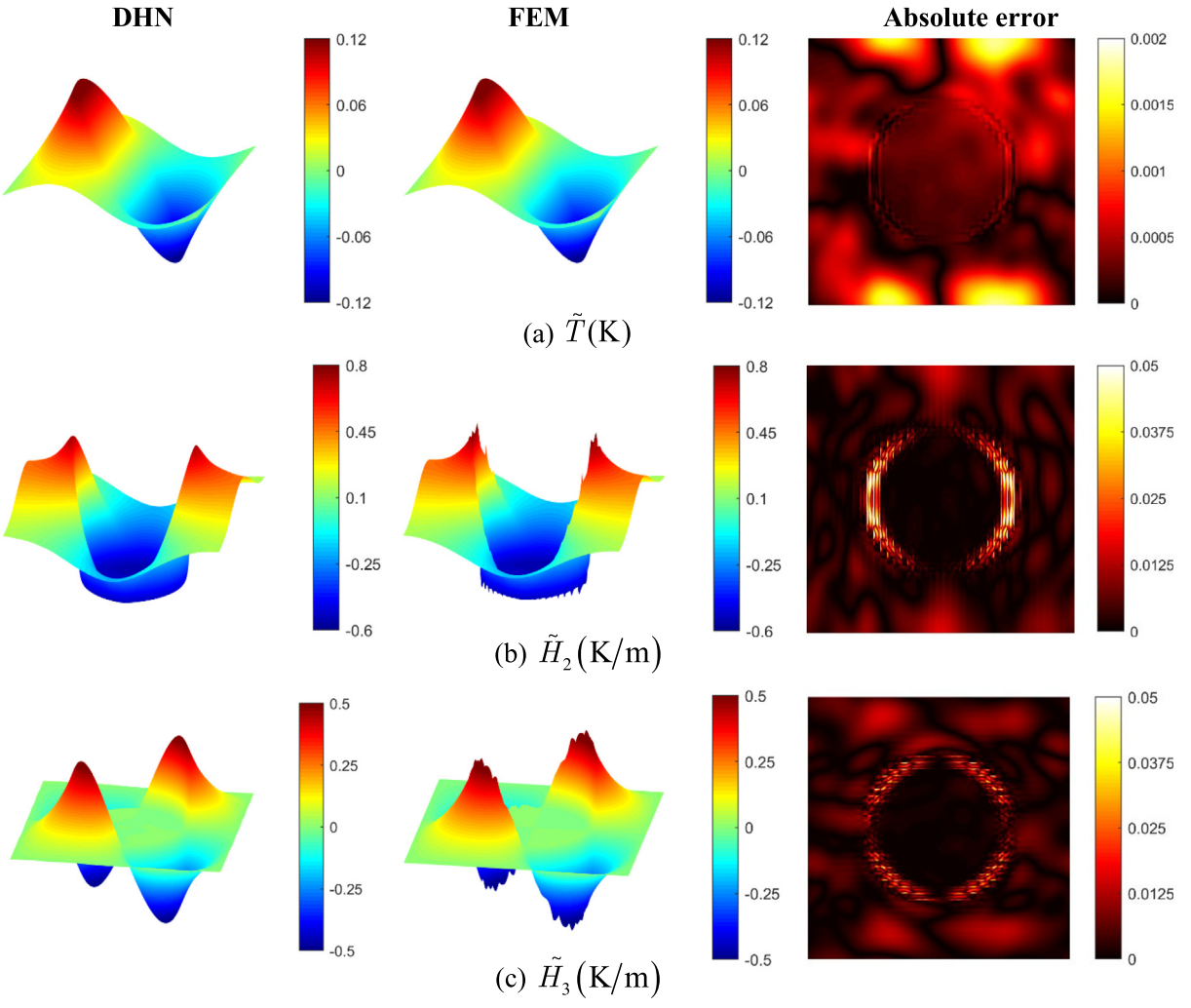


Fig. 6. Comparison of fluctuating temperature and its gradient distributions generated by the DHN and FEM reference solution. It should be noted that the classical finite element method satisfies only C^0 continuities across adjacent elements, hence does not guarantee the continuity of the temperature gradients.

4.3. Porous media

Next, we compare the predictive capability of the DHN theory with the finite-element homogenization predictions of the local and homogenized response of porous composites. Fig. 9 illustrates unit cells containing 5% and 50% of porosity volume fractions, with unidirectionally oriented porosity arranged in a square array. Two datasets with 9.5k and 5k collocation points were generated via Monte-Carlo simulation respectively for 5% and 50% volume fraction cases by excluding collocation points from the porous phase. What is more, to model the porous inclusion, $c_1 = -c_3/2$ in Eq. (26).

Fig. 10 shows the loss functions for the two porous microstructures with the best network model identified in the previous subsection, under transverse temperature gradient loading by $\overline{H}_2 = 1$ K/m. For both cases, the loss functions reach relatively low values after 2000 epochs and remain almost constant afterward, indicating the fast convergence of the neural network models. Comparison of the fluctuating temperature gradient distributions generated by the DHN and finite-element predictions is compared in Fig. 11. It is observed good agreement between the proposed theory and the finite-element reference solutions.

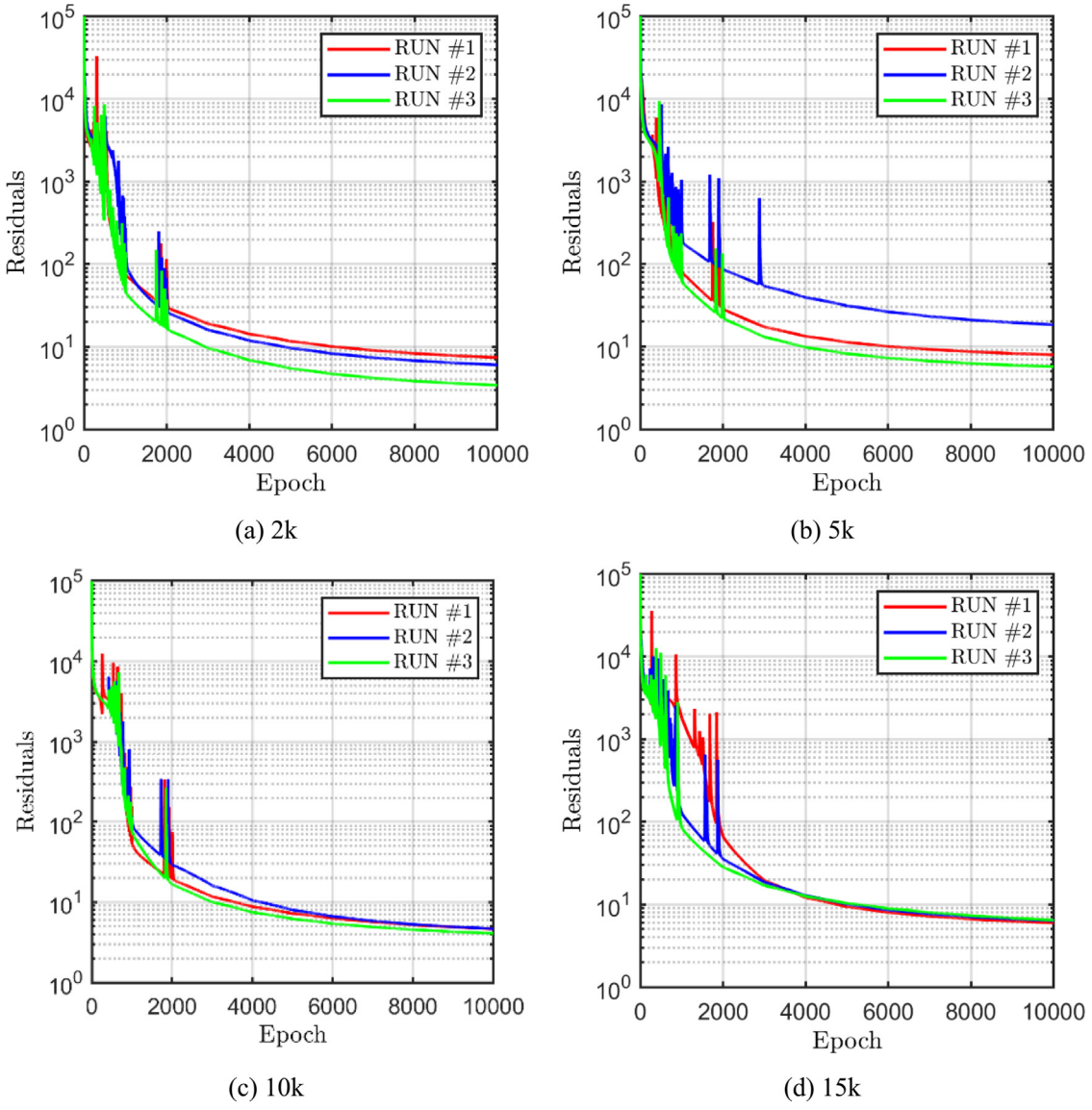


Fig. 7. Effect of dataset size on the neural network performance with three training attempts.

The effective thermal conductivities, normalized by the corresponding matrix coefficient, were generated for the fiber or pore volume fractions $V_f = 0.05 - 0.5$ in increments of $\Delta V_f = 0.05$. The plots of the normalized effective k_{22}^*/k^m as a function of the inclusion volume fraction for the two combinations of fiber/matrix or pore/matrix properties are shown in Fig. 12, which are generated by the developed DHN theory and finite-element simulations. The classical Hashin–Shtrikman bounds have also been included in the figure for comparison [54]. As expected, effective thermal conductivity tends to the fiber/pore property in a nonlinear manner with increasing fiber/pore volume fraction. The DHN and finite-element techniques predict virtually the same results at the entire volume fraction range, which fall within the Hashin–Shtrikman bounds, providing additional evidence of the accuracy of the developed approach.

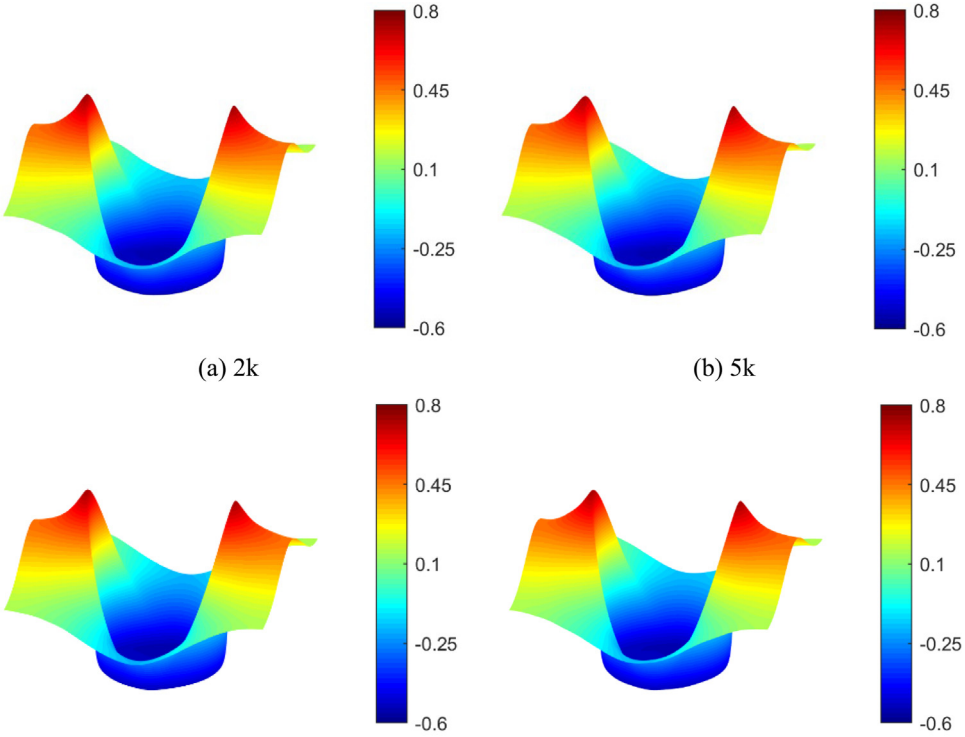


Fig. 8. Effect of training dataset size on predicted transverse temperature gradient distribution $\widetilde{H}_2$ (K/m) with imposition of the macroscopic temperature gradient $\overline{H}_2 = 1$ K/m.

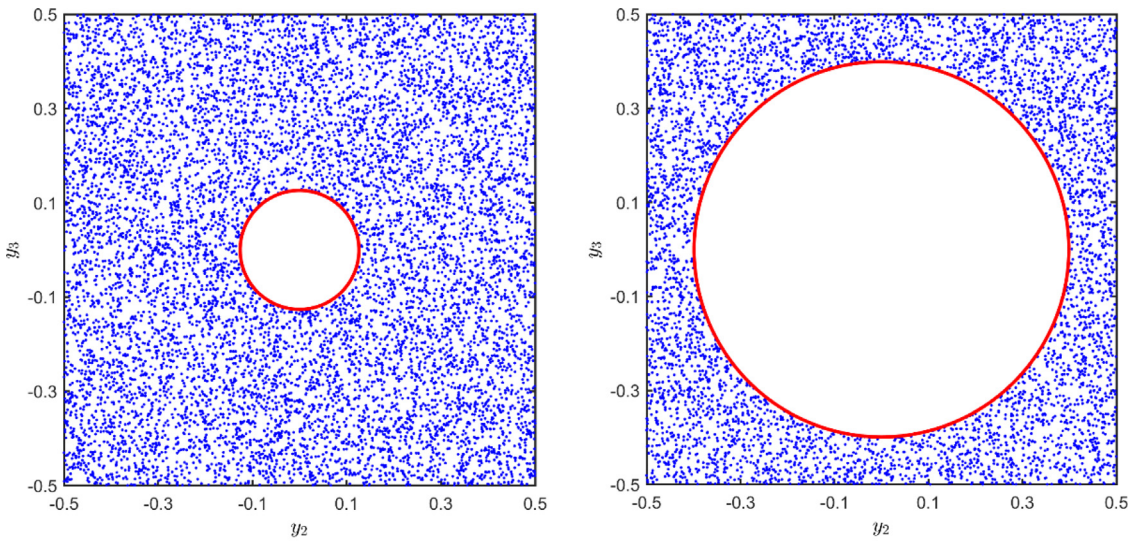


Fig. 9. Unit cells containing (a) 5% and (b) 50% porosity volume fractions, respectively.

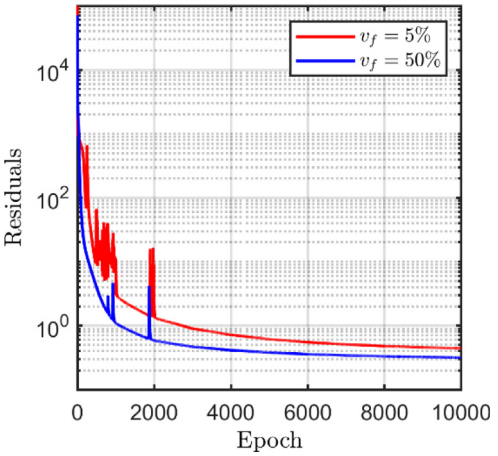


Fig. 10. Comparison of network performance for the porous unit cells with 5% and 50% volume fractions, respectively.

5. Transfer-learning assisted DHN

We further employ the deep homogenization network to conduct numerical experiments to demonstrate the accuracy and capability of the proposed approach to analyze periodic arrays with locally-irregular fiber distributions, as well as to understand fiber–fiber interactions in the local field concentration and homogenized behavior. Such investigation is important in order to interpret accurately the discrepancy of homogenized properties obtained from experimental data, as well as to design/optimize composite structural components in a multiscale analysis setting. While simulating random multi-inclusion periodic arrays is typically outside of elasticity-based approaches' ability, multi-inclusion unit cells with locally-irregular fiber distributions require extensive mesh refinement in the case of the finite-element method, which is necessary to capture the local field distributions. Hence the numerical experiments considered herein highlight the DHN's special strength among the available methods.

Fig. 13 illustrates four microstructural realizations of multi-inclusion periodic arrays containing 20% total fiber volume fractions characterized by regular (Fig. 13a) and locally irregular (Fig. 13b, c, d) fiber distributions. For the regularly spaced fiber arrays, four fibers were created and displaced at the center in a 2×2 subdomain. The random fiber distributions were produced by displacing the fiber centers randomly using the Monte-Carlo simulation and constraints were imposed to prevent fiber from overlapping or being displaced too close in order to avoid numerical issues. Moreover, the fibers at the unit cell edges were cut precisely and moved to the opposite unit cell faces so that complete fibers can be obtained that strictly satisfy periodicity constraints. All fibers have the same radius and the unit cell's overall dimensions have been normalized as in the single inclusion case in Section 4. A dataset with 40k data points was generated such that the data point density for each of the 2×2 subdomains remains the same as in Section 4. As before, a macroscopic temperature gradient by $\overline{H}_2 = 1$ K/m is imposed in the transverse direction.

The small fiber diameter relative to the unit cell's overall dimension and the complex microstructures pose new challenges to the DHN convergence due to the important fiber–fiber interactions which lead to more significant variations of temperature field and temperature gradient distributions. To address this issue, the transfer learning technique was introduced in this section to facilitate neural network convergence, as well as to ensure correct local temperature and gradient field can be obtained. The central idea of the transfer learning assisted DHN is motivated by the fact that increasing smoothness of the fiber/matrix thermal conductivity transition promotes lower values of loss values with the same training epochs but without altering the fundamental characters and magnitudes of the temperature and temperature gradient distributions. Fig. 14(a) depicts the thermal conductivity distributions for Microstructure #2 with the smoothness parameter $\delta = 0.1R$, which is employed as a source model to train the periodic network. Fig. 14(b) illustrates the thermal conductivity distributions for Microstructure #2 with the smoothness parameter $\delta = 0.05R$, which is the target model whose solution is closer to the real scenario. Two identical neural network models, namely the source and target models, are built. The DHN is first employed to train the source model to obtain the response of periodic arrays. Then, it is assumed that the trained source model contains the common knowledge of the temperature field and gradients that can be applied to the target model

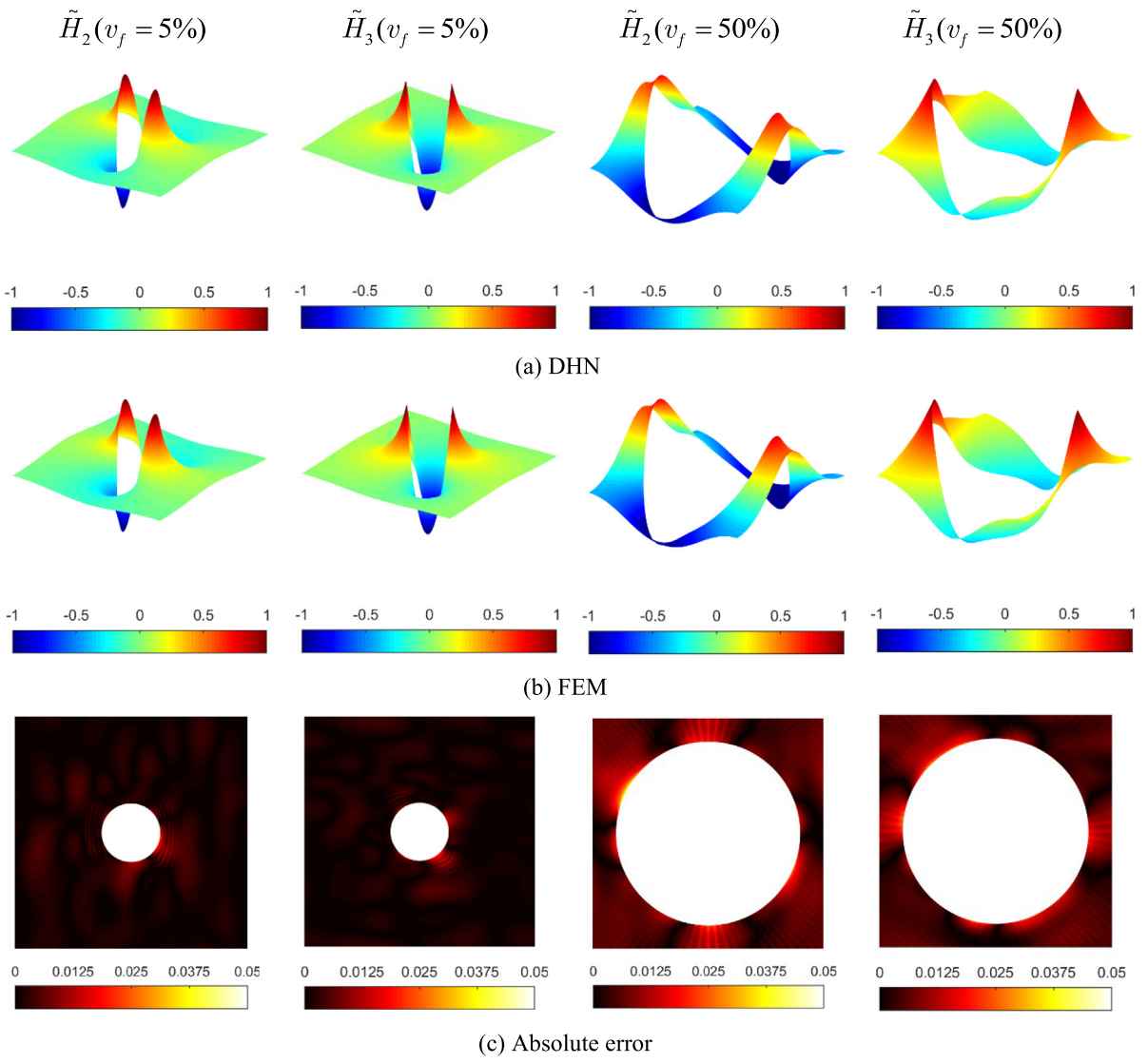
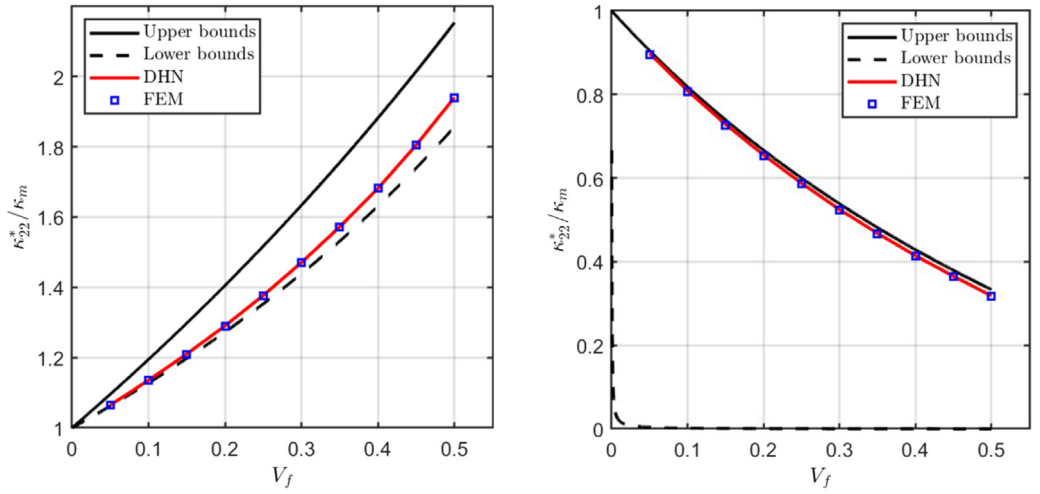


Fig. 11. Comparison of fluctuating temperature gradient distribution (K/m) generated by the DHN and FEM reference solution for two porous unit cells with 5% and 50% volume fractions, respectively.

(since the temperature field and gradients are not fundamentally altered in the case of different material transition parameters). Therefore, the weights and biases of the pre-trained source models can be utilized as prior knowledge to initialize the network for the target models. As will be shown in the sequel, the rate of convergence of the neural network model with transfer learning is much faster than that trained directly from scratch.

Comparison of the training performance of the four microstructural realizations for the source and target models is shown in Fig. 15. The loss functions by training directly the target model with random seed values are plotted in each subfigure for comparison. It is evident that the direct training on the target models over 15 000 epochs for Microstructures #1, #2, and #3 yield significantly high loss values, indicating that reliable unit cell solutions may have not been obtained. In contrast, the training on the source models always retains the lowest loss values for each microstructural realization because these cases are less demanding due to the smoother material transition. The subsequent training on the target models with the help of transfer learning demonstrates the effectiveness of the developed framework, that is, the loss values of the transfer learning target models are always lower than training neural networks from scratch because the weights and biases of the former models are more effective.



(a) Fiber/matrix system (b) Pore/matrix system

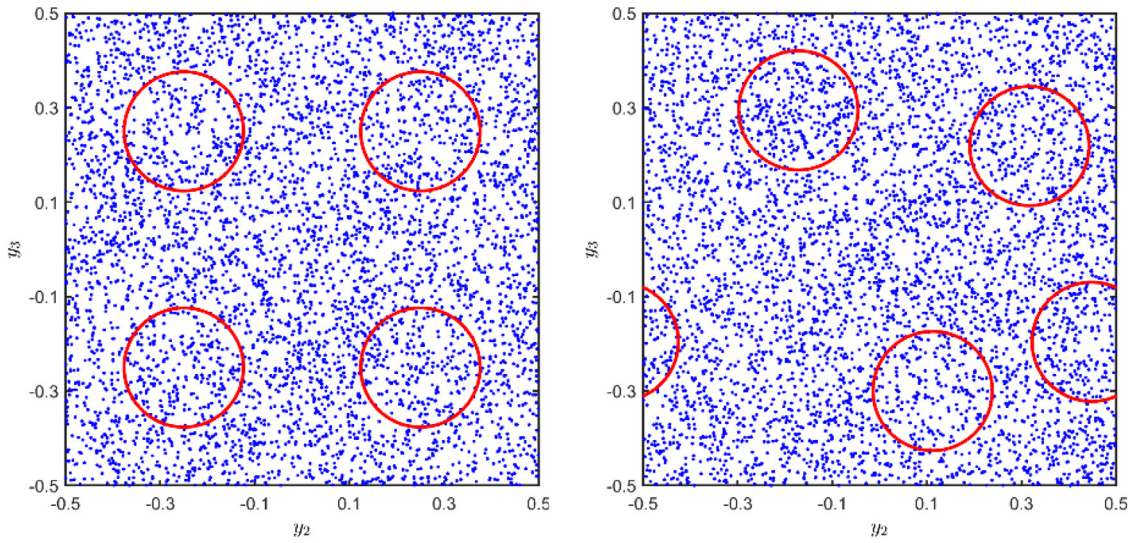
Fig. 12. Comparison of the homogenized thermoconductive coefficient as a function of fiber or porosity volume fraction predicted by the DHN method against the FEM and the Hashin–Shtrikman bounds of Willis [54].

Fig. 16 compares the transverse fluctuating temperature gradient $\widetilde{H}_2$ distributions generated by training the source model, the target model with transfer learning, the target model from scratch, and the finite-element gold solution for the selected periodic array (Microstructure #2). As anticipated, the overall temperature distributions predicted by the source model and the target model with transfer learning remain virtually the same over the entire unit cell domain. The small differences between the source model and target model with transfer learning predictions along the fiber/matrix interface are due to the differences in the smoothness of the material transition, illustrating that the DHN is sufficiently sensitive to correctly capture this small effect. The result obtained from the target model with transfer learning matches very well the finite-element reference solution, and the distributions themselves are nearly identical, providing concrete support for the developed transfer-learning DHN approach. In contrast, training the source model from scratch produces a drastically different local temperature gradient distribution from the finite-element reference solution, indicating that the training of DHN from scratch failed in this loading scenario.

Fig. 17 presents comparison of local fluctuating temperature distributions $\widetilde{T}$ (K) for all the considered microstructures generated by the DHN with transfer learning and the finite-element referenced results with the imposition of a unit transverse temperature gradient $\overline{H}_2 = 1$ K/m. The correlations between the two approaches are remarkable. Particularly worthy of mentioning is that the temperature field is smoothly varying without any discontinuities even at the fiber/matrix interface. Figs. 18 and 19 show differences in resulting fluctuating temperature gradients $\widetilde{H}_2$ (K/m) and $\widetilde{H}_3$ (K/m) distributions obtained by the DHN and finite-element reference results, respectively. As before, the DHN produces markedly smoother temperature gradient distributions along the fiber/matrix interface than the FEM, but the characters remain the same for all the microstructures. Away from the interfaces, the differences in temperature gradients predicted by the two approaches vanish in the presented figures.

We note that an important advantage of the proposed DHN method over the finite-element homogenization technique lies in the manner of satisfying the periodicity boundary conditions. In the former case, the periodic layer in the DHN framework ensures that the periodicity boundary conditions of both temperature and temperature gradient (or heat flux) are satisfied exactly (or to the machine precision) along the unit cell opposite edges, as presented in Fig. 20(a) and (b), respectively. In the latter case, the periodicity conditions are applied only to the external nodal temperatures of the unit cell.

We end this section by summarizing the hyperparameters for reproducing various results presented in the present work, including the network architecture, activation function, optimizer type, training dataset size, and the number of epochs, etc., as shown in Table 2. We emphasize that the same network architecture identified in Section 4.1 has been utilized for generating all the numerical cases in the subsequent sections.



(a) Microstructure #1

(b) Microstructure #2

(c) Microstructure #3

(d) Microstructure #4

Fig. 13. Unit cells containing 20% fiber volume fraction characterized by regular and locally-irregular fiber distributions (only 5k collocation points were plotted for visualization purpose).

6. Discussion

Despite the rapid advance in the micromechanics and homogenization theories of heterogeneous media in the past two decades, the search for ideal micromechanics models continues. Elements of the zeroth-order homogenization theory and physically informed deep neural network were seamlessly connected to construct a new micromechanics technique for the micromechanical analysis of thermoconductive composites with regular or locally-irregular fiber

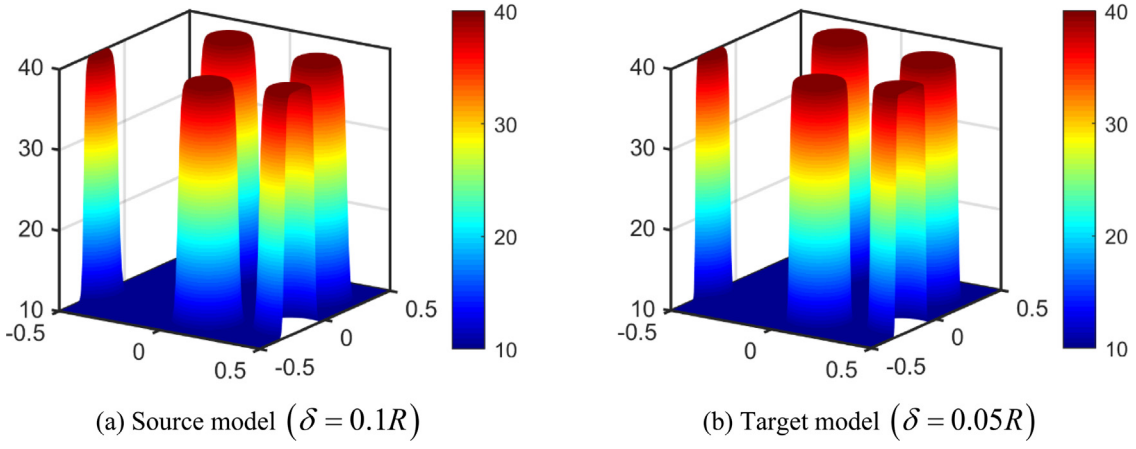


Fig. 14. Comparison of thermal conductivity distributions for the selected microstructure #2: (a) source model and (b) target model.

Table 2
Summary of the hyperparameters for reproducing the various results presented in this work.

	Fig. 6	Fig. 11	Fig. 17
Neurons in the periodic layer	20	20	20
Hidden layers	5	5	5
Hidden neurons	20	20	20
Activation function	$\sigma(x) = \tanh(x)$	$\sigma(x) = \tanh(x)$	$\sigma(x) = \tanh(x)$
Learning rate	0.01 and decaying by a factor of 0.5 for every 1000 epoch	0.01 and decaying by a factor of 0.5 for every 1000 epoch	0.01 and decaying by a factor of 0.5 for every 1000 epoch
Optimizer	Adam	Adam	Adam
Training dataset size	10k	9.5k ($v_f = 5\%$) 5k ($v_f = 50\%$)	40k
Testing dataset size	280 × 280	280 × 280	280 × 280
Number of epochs	10k	10k	15k

distributions. Novel use of a periodic layer makes possible the exact and automatic imposition of periodicity boundary conditions to infinite order in a pointwise sense, circumventing incorporating a penalty term that represents the residual norm of the boundary conditions into the loss function. In contrast, in the finite-element approach, the periodicity conditions are only enforced on the external nodal temperature of the unit cell.

The second advantage lies in the mesh-free nature of the neural network approach. Specifically, the spatial coordinates for training a network model can be collected randomly using Monte-Carlo simulation over the unit cell domain. Hence, the conventional unit cell mesh discretization, which may be computationally extensive and cumbersome in the case of multiple random fiber distribution, is eliminated.

Another important advantage of the proposed theory is that, intrinsically, the neural network solution is infinitely differentiable. Therefore, both fluctuating temperature and its gradients are continuous and smoothly varying over the unit cell domain. The primal finite-element method, however, can only satisfy the temperature continuity condition

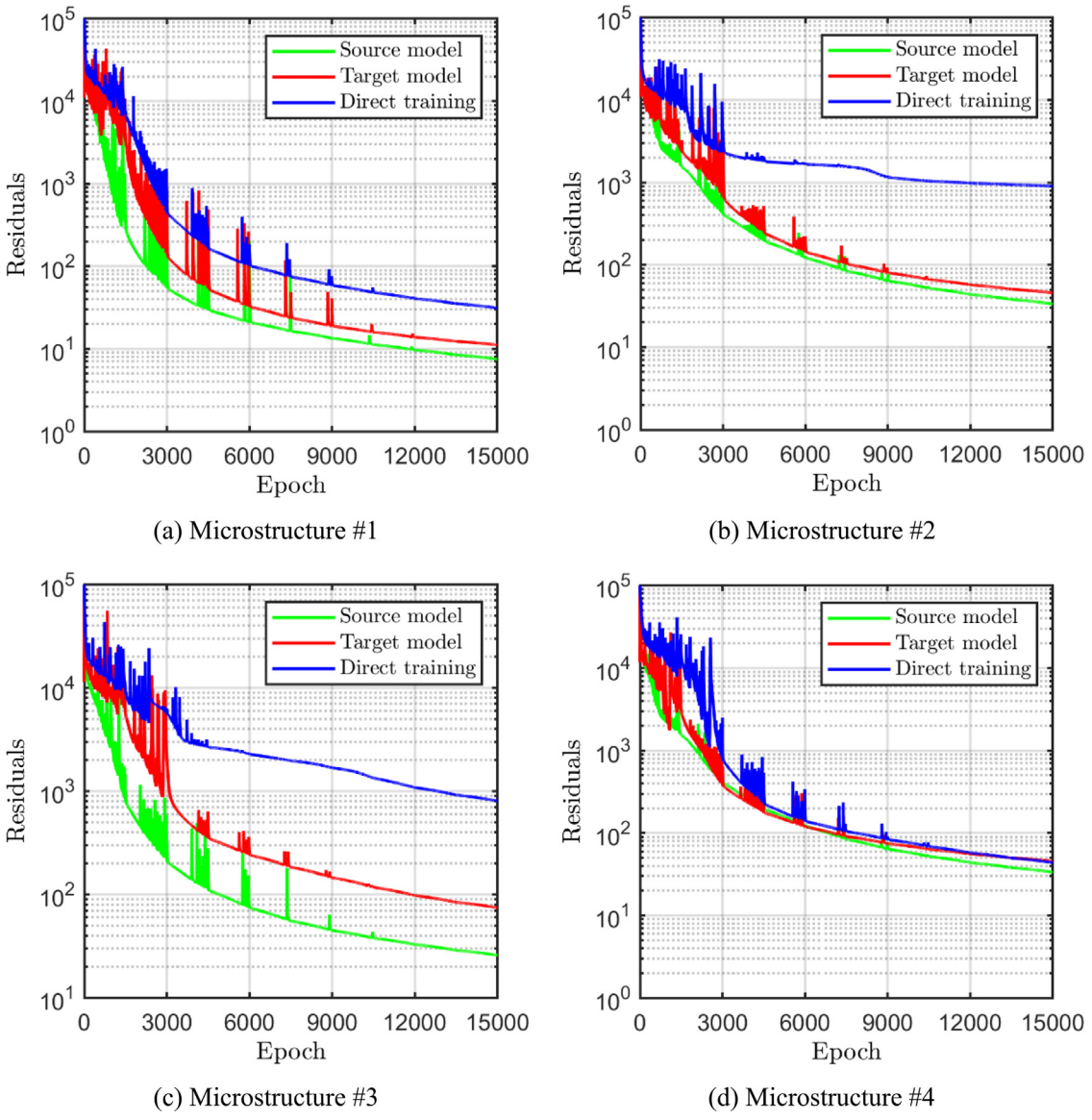


Fig. 15. Comparison of the loss functions for four different microstructural realizations for the source model, target model and direct training model.

between the adjacent elements while the calculated temperature gradients are not necessarily continuous from one element to another.

7. Summary and conclusions

A novel physics-informed neural network-based homogenization theory was developed within the zeroth-order homogenization framework for continuum micromechanical analysis of thermoconductive composites with

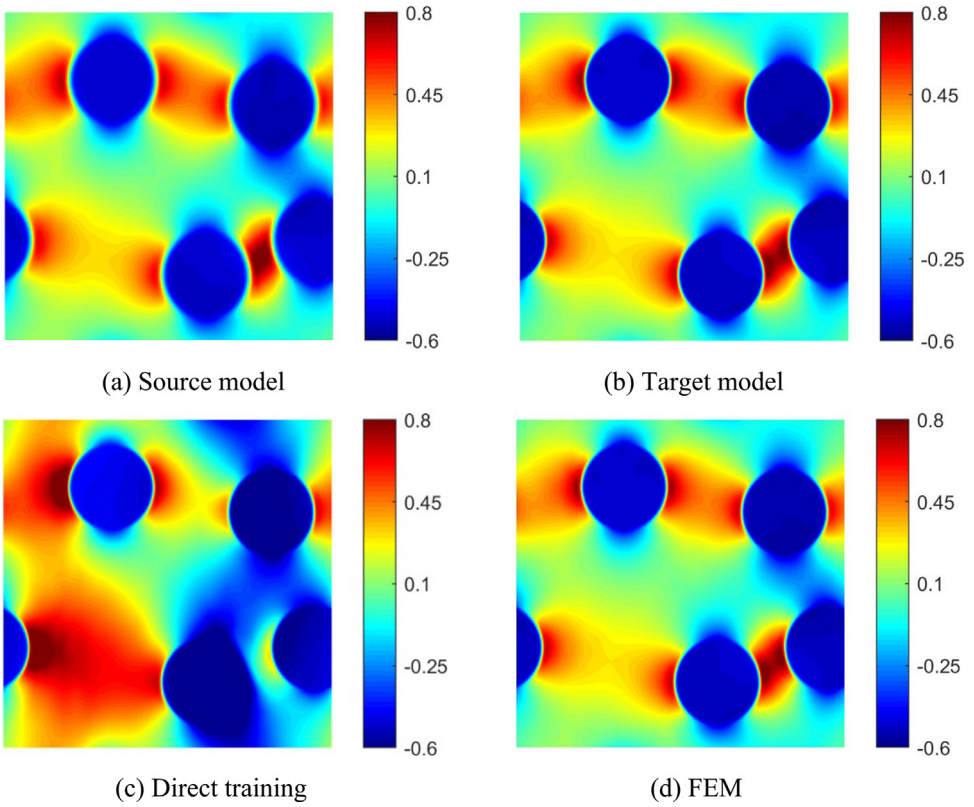


Fig. 16. Comparison of fluctuating temperature gradients $\widetilde{H}_2$ (K/m) predicted by (a) source model; (b) target model; (c) direct training; (d) FEM model.

unidirectional fiber-reinforced or porous composites with random fiber distributions. This theory employs a two-scale expansion of the temperature field of spatially uniform composites in terms of macroscopic and fluctuating contributions. The latter is solved using the feed-forward neural network layers by minimizing a cost function given by the residuals of the steady-state heat conduction governing equation evaluated over a set of training data points. The periodicity boundary conditions are automatically and exactly satisfied in a novel manner with the help of a periodic layer involving a set of periodic functions and nonlinear activation functions.

Extensive numerical experiments were conducted to illustrate the effect of neural network hyper-parameters and dataset size on neural network performance. The accuracy of the homogenized properties and local temperature and gradient distributions of unit cells reinforced by a unidirectional fiber or weakened by a porosity generated by the proposed theory is demonstrated by comparison with the finite-element simulations. The network-based homogenization approach is further enhanced with the transfer learning technique, enabling the efficient and accurate simulation of random fiber distributions where periodicity boundary conditions are more important.

The marriage of the physics-informed deep neural network model and the homogenization theory offers many possibilities for further extension and investigation. For instance, it remains to be an open question if the developed approach is also applicable to solving mechanical problems involving multiple PDEs (hence multiple loss terms). The developed approach also motivates the extension to three-dimensional periodic composites with arbitrarily shaped and oriented inclusions.

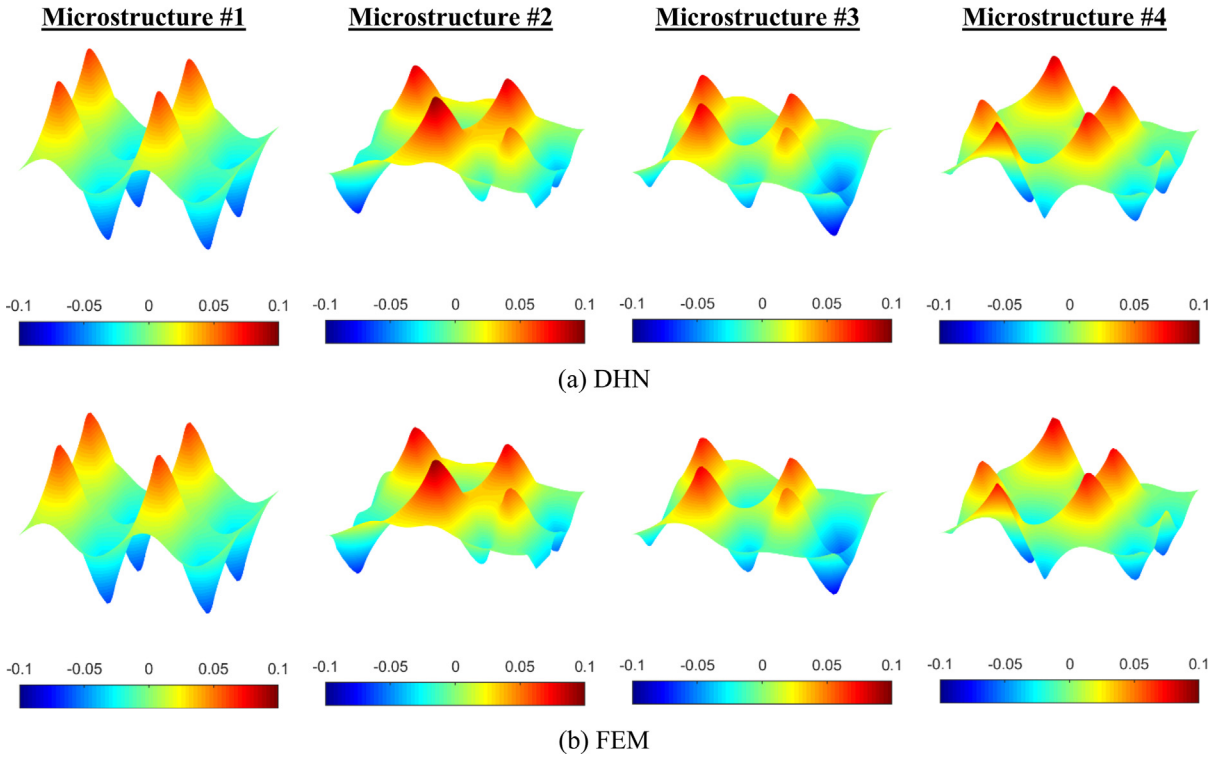


Fig. 17. Comparison of fluctuating temperature distribution $\widetilde{T}$ (K) generated by the DHN and FEM reference solution for different microstructures.

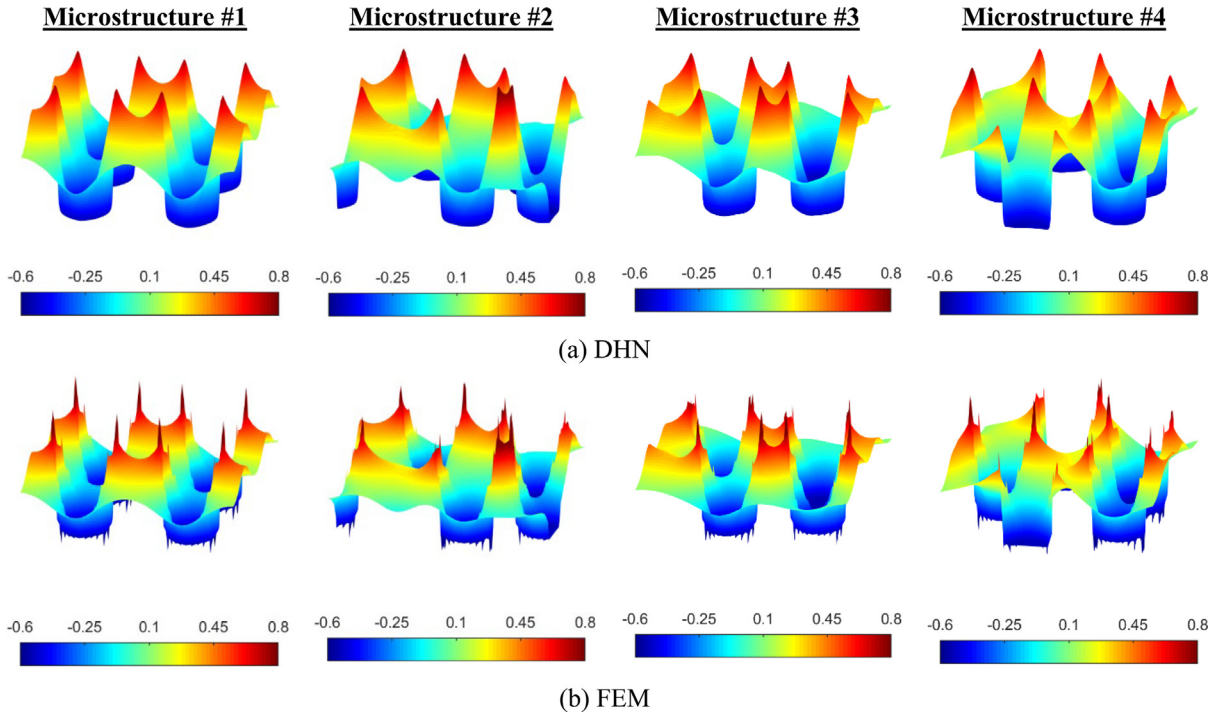


Fig. 18. Comparison of fluctuating temperature gradient $\widetilde{H}_2$ (K/m) distribution generated by the DHN and FEM reference solution for uniform fiber arrangement.

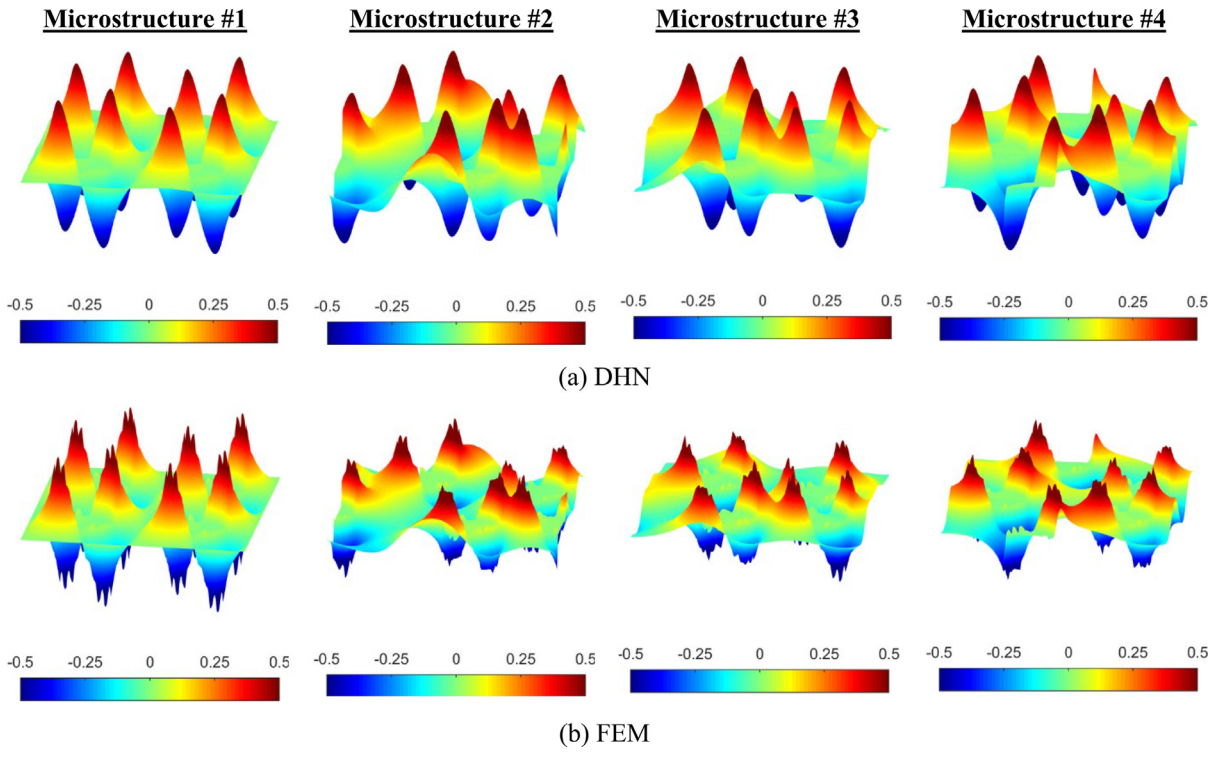


Fig. 19. Comparison of fluctuating temperature gradient $\tilde{H}_3$ (K/m) distribution generated by the DHN and FEM reference solution for uniform fiber arrangement.

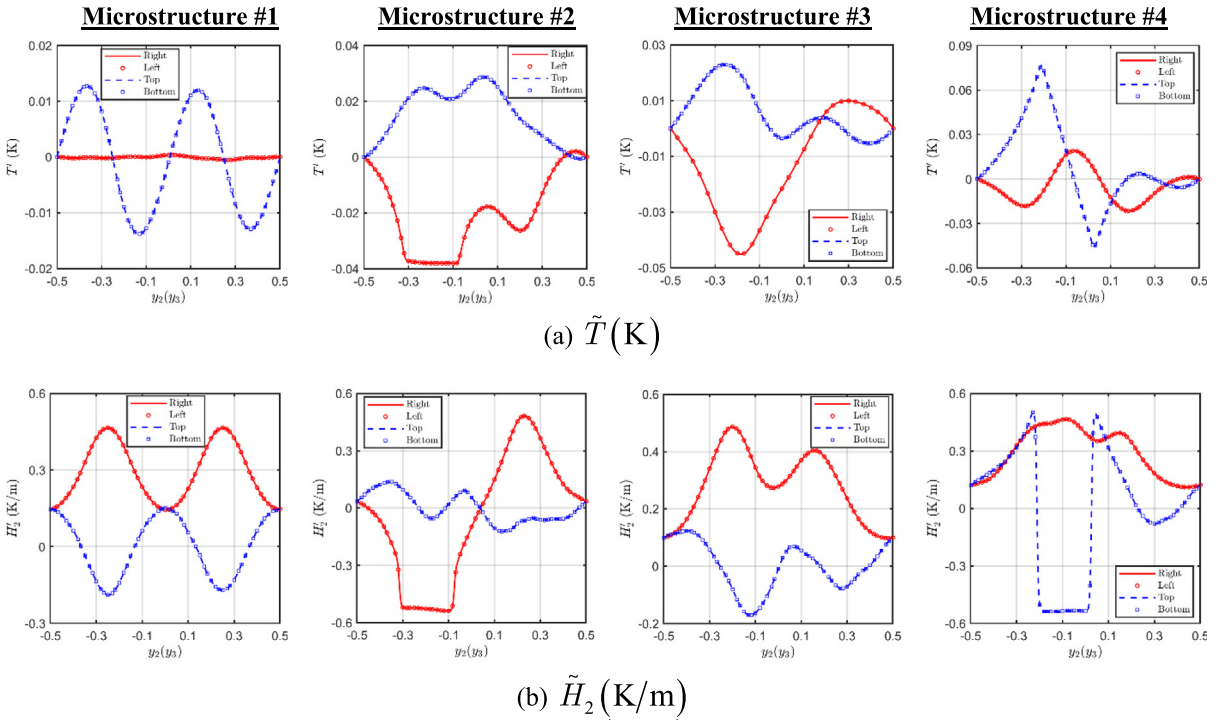


Fig. 20. Comparison of fluctuating temperature and its gradients along the unit cell boundaries, illustrating the exact satisfaction of periodicity boundary conditions.

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.

Data availability

Data will be made available on request.

References

[1] F. Gehrig, D. Wicht, M. Krause, T. Böhlke, FFT-based investigation of the shear stress distribution in face-centered cubic polycrystals, Int. J. Plast. 157 (2022) 103369.

[2] L. Wu, L. Noels, L. Adam, I. Doghri, An implicit-gradient-enhanced incremental-secant mean-field homogenization scheme for elasto-plastic composites with damage, Int. J. Solids Struct. 50 (24) (2013) 3843–3860.

[3] Q. Chen, G. Chatzigeorgiou, F. Meraghni, Hybrid hierarchical homogenization theory for unidirectional CNTs-coated fuzzy fiber composites undergoing inelastic deformations, Compos. Sci. Technol. 215 (2021) 109012.

[4] F. Praud, G. Chatzigeorgiou, F. Meraghni, Fully integrated multi-scale modelling of damage and time-dependency in thermoplastic-based woven composites, Int. J. Damage Mech. 30 (2) (2020) 163–195.

[5] J. Zhi, K. Raju, T.-E. Tay, V.B.C. Tan, Transient multi-scale analysis with micro-inertia effects using direct FE2 method, Comput. Mech. 67 (6) (2021) 1645–1660.

[6] Z. He, M.-J. Pindera, Locally exact asymptotic homogenization of periodic materials under anti-plane shear loading, Eur. J. Mech. A Solids 81 (2020) 103972.

[7] M.-J. Pindera, H. Khatam, A.S. Drago, Y. Bansal, Micromechanics of spatially uniform heterogeneous media: a critical review and emerging approaches, Composites B 40 (5) (2009) 349–378.

[8] G. Chatzigeorgiou, N. Charalambakis, Y. Chemisky, F. Meraghni, 6 - Composites with random structure, in: G. Chatzigeorgiou, N. Charalambakis, Y. Chemisky, F. Meraghni (Eds.), Thermomechanical Behavior of Dissipative Composite Materials, Elsevier, 2017, pp. 195–236.

[9] Q. Chen, G. Chatzigeorgiou, F. Meraghni, Extended mean-field homogenization of viscoelastic-viscoplastic polymer composites undergoing hybrid progressive degradation induced by interface debonding and matrix ductile damage, Int. J. Solids Struct. 210–211 (2021) 1–17.

[10] R. Hill, Theory of mechanical properties of fibre-strengthened materials—III. Self-consistent model, J. Mech. Phys. Solids 13 (4) (1965) 189–198.

[11] R.M. Christensen, K.H. Lo, Solutions for effective shear properties in three phase sphere and cylinder models, J. Mech. Phys. Solids 27 (4) (1979) 315–330.

[12] T. Mori, K. Tanaka, Average stress in matrix and average elastic energy of materials with misfitting inclusions, Acta Metall. 21 (5) (1973) 571–574.

[13] Q. Chen, G. Chatzigeorgiou, F. Meraghni, Extended mean-field homogenization of unidirectional piezoelectric nanocomposites with generalized Gurtin–Murdoch interfaces, Compos. Struct. 307 (2023) 116639.

[14] Z. Hashin, B.W. Rosen, The elastic moduli of fiber-reinforced materials, J. Appl. Mech. 31 (2) (1964) 223–232.

[15] S.G. Mogilevskaya, S.L. Crouch, H.K. Stolarski, A. Benusiglio, Equivalent inhomogeneity method for evaluating the effective elastic properties of unidirectional multi-phase composites with surface/interface effects, Int. J. Solids Struct. 47 (3) (2010) 407–418.

[16] E. Tikarrouchine, A. Benaarbia, G. Chatzigeorgiou, F. Meraghni, Non-linear FE2 multiscale simulation of damage, micro and macroscopic strains in polyamide 66-woven composite structures: Analysis and experimental validation, Compos. Struct. 255 (2021) 112926.

[17] E. Tikarrouchine, G. Chatzigeorgiou, Y. Chemisky, F. Meraghni, Fully coupled thermo-viscoplastic analysis of composite structures by means of multi-scale three-dimensional finite element computations, Int. J. Solids Struct. 164 (2019) 120–140.

[18] W. Tu, Q. Chen, Homogenization and localization of unidirectional fiber-reinforced composites with evolving damage by FVDAM and FEM approaches: A critical assessment, Eng. Fract. Mech. 239 (2020) 107280.

[19] Q. Chen, M.-J. Pindera, Homogenization and localization of elastic–plastic nanoporous materials with Gurtin–Murdoch interfaces: An assessment of computational approaches, Int. J. Plast. 124 (2020) 42–70.

[20] P. Cardiff, I. Demirdžić, Thirty years of the finite volume method for solid mechanics, Arch. Comput. Methods Eng. 28 (5) (2021) 3721–3780.

[21] M.A.A. Cavalcante, M.-J. Pindera, Generalized FVDAM theory for elastic–plastic periodic materials, Int. J. Plast. 77 (2016) 90–117.

[22] A.S. Drago, M.-J. Pindera, A locally exact homogenization theory for periodic microstructures with isotropic phases, Trans. ASME, J. Appl. Mech. 75 (5) (2008).

[23] Z. He, M.-J. Pindera, Locally exact asymptotic homogenization of viscoelastic composites under anti-plane shear loading, Mech. Mater. 155 (2021) 103752.

[24] S. Yin, Z. He, M.-J. Pindera, A new hybrid homogenization theory for periodic composites with random fiber distributions, Compos. Struct. 269 (2021) 113997.

[25] E.N. Lages, S.P. Cavalcanti Marques, Prediction of effective thermal conductivity of multiphase composites with periodic microstructures using an expanded micromechanical model, Int. J. Therm. Sci. 171 (2022) 107226.

[26] E.N. Lages, S.P.C. Marques, Thermoelastic homogenization of periodic composites using an eigenstrain-based micromechanical model, Appl. Math. Model. 85 (2020) 1–18.
[27] S. Saeb, P. Steinmann, A. Javili, Aspects of computational homogenization at finite deformations: A unifying review from Reuss' to Voigt's bound, Appl. Mech. Rev. 68 (5) (2016).
[28] Q. Chen, G. Wang, M.-J. Pindera, Homogenization and localization of nanoporous composites - A critical review and new developments, Composites B 155 (2018) 329–368.
[29] F. Masi, I. Stefanou, Multiscale modeling of inelastic materials with thermodynamics-based artificial neural networks (TANN), Comput. Methods Appl. Mech. Engrg. 398 (2022) 115190.
[30] A. Henkes, H. Wessels, R. Mahnken, Physics informed neural networks for continuum micromechanics, Comput. Methods Appl. Mech. Engrg. 393 (2022) 114790.
[31] S. Gajek, M. Schneider, T. Böhlke, On the micromechanics of deep material networks, J. Mech. Phys. Solids 142 (2020) 103984.
[32] S. Gajek, M. Schneider, T. Böhlke, An FE–DMN method for the multiscale analysis of short fiber reinforced plastic components, Comput. Methods Appl. Mech. Engrg. 384 (2021) 113952.
[33] S. Haggag, L. Nasrat, H. Ismail, ANN approaches to determine the dielectric strength improvement of MgO based low density polyethylene nanocomposite, J. Adv. Dielectr. 11 (04) (2021) 2150016.
[34] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707.
[35] E. Samaniego, C. Anitescu, S. Goswami, V.M. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, T. Rabczuk, An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications, Comput. Methods Appl. Mech. Engrg. 362 (2020) 112790.
[36] V.M. Nguyen-Thanh, C. Anitescu, N. Alajlan, T. Rabczuk, X. Zhuang, Parametric deep energy approach for elasticity accounting for strain gradient effects, Comput. Methods Appl. Mech. Engrg. 386 (2021) 114096.
[37] H. Guo, X. Zhuang, T. Rabczuk, A deep collocation method for the bending analysis of Kirchhoff plate, 2021, arXiv preprint arXiv:2102.02617.
[38] J. Bai, T. Rabczuk, A. Gupta, L. Alzubaidi, Y. Gu, A physics-informed neural network technique based on a modified loss function for computational 2D and 3D solid mechanics, Comput. Mech. (2022).
[39] J. Jiang, J. Zhao, S. Pang, F. Meraghni, A. Siadat, Q. Chen, Physics-informed deep neural network enabled discovery of size-dependent deformation mechanisms in nanostructures, Int. J. Solids Struct. 236–237 (2022) 111320.
[40] S. Dong, J. Yang, On computing the hyperparameter of extreme learning machines: Algorithm and application to computational PDEs, and comparison with classical and high-order finite elements, J. Comput. Phys. 463 (2022) 111290.
[41] S. Dong, J. Yang, Numerical approximation of partial differential equations by a variable projection method with artificial neural networks, Comput. Methods Appl. Mech. Engrg. 398 (2022) 115284.
[42] W. Li, M.Z. Bazant, J. Zhu, A physics-guided neural network framework for elastic plates: Comparison of governing equations-based and energy-based approaches, Comput. Methods Appl. Mech. Engrg. 383 (2021) 113933.
[43] J. Gasick, X. Qian, Isogeometric neural networks: A new deep learning approach for solving parameterized partial differential equations, Comput. Methods Appl. Mech. Engrg. 405 (2023) 115839.
[44] S. Dong, N. Ni, A method for representing periodic functions and enforcing exactly periodic boundary conditions with deep neural networks, J. Comput. Phys. 435 (2021) 110242.
[45] S. Metoui, E. Pruliere, A. Ammar, F. Dau, I. Iordanoff, A multiscale separated representation to compute the mechanical behavior of composites with periodic microstructure, Math. Comput. Simulation 144 (2018) 162–181.
[46] Q. Chen, W. Tu, M. Ma, Deep learning in heterogeneous materials: Targeting the thermo-mechanical response of unidirectional composites, J. Appl. Phys. 127 (17) (2020) 175101.
[47] G. Wang, M. Gao, B. Yang, Q. Chen, The morphological effect of carbon fibers on the thermal conductive composites, Int. J. Heat Mass Transfer 152 (2020) 119477.
[48] A. Bensoussan, J.-L. Lions, G. Papanicolaou, Asymptotic Analysis for Periodic Structures, North Holland, Amsterdam, Netherlands, 1978.
[49] Z. He, J. Liu, Q. Chen, Higher-order asymptotic homogenization for piezoelectric composites, Int. J. Solids Struct. 264 (2023) 112092.
[50] M.A.A. Cavalcante, H. Khatam, M.-J. Pindera, Homogenization of elastic–plastic periodic materials by FVDAM and FEM approaches – An assessment, Composites B 42 (6) (2011) 1713–1730.
[51] Q. Chen, G. Chatzigeorgiou, F. Meraghni, A. Javili, Homogenization of size-dependent multiphysics behavior of nanostructured piezoelectric composites with energetic surfaces, Eur. J. Mech. A Solids 96 (2022) 104731.
[52] A. Zhang, D. Mohr, Using neural networks to represent von mises plasticity with isotropic hardening, Int. J. Plast. 132 (2020) 102732.
[53] C. Bajaj, L. McLennan, T. Andeen, A. Roy, Robust learning of physics informed neural networks, 2021, arXiv preprint arXiv:2110.13330.
[54] J.R. Willis, Bounds and self-consistent estimates for the overall properties of anisotropic composites, J. Mech. Phys. Solids 25 (3) (1977) 185–202.