Deep homogenization networks for elastic heterogeneous materials with two- and three-dimensional periodicity

Jiajun Wu, Jindong Jiang, Qiang Chen, George Chatzigeorgiou, Fodil Meraghni

Deep homogenization networks for elastic heterogeneous materials with two- and three-dimensional periodicity

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

[a] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China

[b] Arts et Métiers Institute of Technology, PIMM, HESAM University, F-75013 Paris, France

[c]Arts et Métiers Institute of Technology, LCFC, HESAM University, F-57070 Metz, France

[d]Arts et Métiers Institute of Technology, CNRS, Université de Lorraine, LEM3-UMR7239, F-57000 Metz, France

[1]These authors contributed equally to this work.

*Corresponding author. Email: kirkcq@gmail.com; qiang.chen@ensam.eu

Abstract

We present a deep learning framework that leverages computational homogenization expertise to predict the local stress field and homogenized moduli of heterogeneous materials with two- and three-dimensional periodicity, which is named physics-informed Deep Homogenization Networks (DHN). To this end, the displacement field of a repeating unit cell is expressed as two-scale expansion in terms of averaging and fluctuating contributions dependent on the global and local coordinates, respectively, under arbitrary multi-axial loading conditions. The latter is regarded as a mesh-free periodic domain estimated using fully connected neural network layers by minimizing residuals of Navier's displacement equations of anisotropic microstructured materials for specified macroscopic strains with the help of automatic differentiation. Enabled by the novel use of a periodic layer, the boundary conditions are encoded directly in the DHN architecture which ensures exact satisfaction of the periodicity conditions of displacements and tractions without introducing additional penalty terms. To verify the proposed model, the local field variables and homogenized moduli were examined for various composites against the finite-element technique. We also demonstrate the feasibility of the proposed framework for simulating unit cells with

locally irregular fibers via transfer learning and find a significant enhancement in the accuracy of stress field recovery during neural network retraining.

Keywords: Physically informed deep neural network; Computational homogenization; Micromechanics; Composite Materials; Transfer learning

1. Introduction

The development of micromechanical methodologies for heterogeneous materials has gained tremendous attention in the computational mechanics community. Most works concerning micromechanics techniques have been proposed during the past five decades, starting from the early classical micromechanics models such as the self-consistent method (Hill, 1965), the composite cylinders/spheres assemblage (Hashin and Rosen, 1964), the three-phase model (Christensen and Lo, 1979), the Mori-Tanaka scheme (Mori and Tanaka, 1973) and progressing to sophisticated models involving more computationally demanding analytical and numerical treatment such as the fast Fourier transformation approach (Gehrig et al., 2022), asymptotic homogenization method (Cruz-González et al., 2020; He and Pindera, 2021), equivalent inhomogeneity technique (Mogilevskaya et al., 2010), the locally-exact homogenization theory (He and Pindera, 2020), the finite-element/volume based micromechanics schemes (Cavalcante et al., 2011; Chen and Pindera, 2020), and the more recent neural network based techniques (Gajek et al., 2020, 2021; Henkes et al., 2022; Jiang et al., 2023; Kalina et al., 2023; Liu et al., 2019).

The micromechanics approaches may be grouped into two broad categories according to the different geometric representations of the material microstructures and boundary condition implementations based on the concept of representative volume element (RVE) or the repeating unit cell (RUC) (Pindera et al., 2009). Specifically, the former approaches are developed for statistically homogeneous microstructures (nonperiodic assumption) which most often ignore the

microstructural details. They consider that, under either homogeneous displacement or homogeneous traction boundary conditions, the response of an RVE is identical to that of the material-at-large. Alternatively, the RUC-based approaches are based on the periodic material assumption which necessitates the imposition of appropriate periodic boundary conditions. For these approaches, the RUC is viewed as the smallest element of periodic arrays that explicitly take into account the underlying microstructures and their interactions. Hence, they serve as the basic building block for the entire arrays through duplication in relevant directions. The advantages of RUC models relative to the RVE approaches, as well as the advancements in computing power, have sparked greater emphasis on the development of solution techniques for the former, albeit at the cost of greater complexity associated with the solution of the related unit cell boundary value problem. The readers are referred to the recent review articles (Chen et al., 2018; Firooz et al., 2021; Saeb et al., 2016) for additional references in this area.

Despite the rapid advances in micromechanics methodologies, continuous efforts are being made for the exploration of novel micromechanics models with utmost accuracy and efficiency. The deep neural network as a universal function approximator may act as a viable alternative to the conventional analytical or numerical methods for solving the micromechanics problem for composite materials (Dey et al., 2022; Gajek et al., 2020, 2021; Henkes et al., 2022; Jiang et al., 2023; Liu et al., 2019). It has been widely recognized that adopting the blackbox machine learning algorithm to predict the response of microstructural materials without imposing the physics constraints may lack interpretability and generalizability, not to mention the issues of significant demand for training data and the positive definite elastic stiffness tensor of the homogenized materials which cannot be guaranteed (Brodnik et al., 2023). The recent development of physics-informed deep neural networks has unlocked new potential for solving nonlinear partial

differential equations (PDEs) with specified boundary conditions due to its ease of implementation by circumventing directly solving the PDEs (Gu et al., 2023; Haghighat et al., 2021; McClenny and Braga-Neto, 2023; Niu et al., 2023; Raissi et al., 2019; Vahab et al., 2023). The principal idea is to use the neural network layers as a global ansatz function for the PDE with the help of automatic differentiation. Minimization of the cost function constructed in terms of the PDE residuals and the boundary loss evaluated on a set of collocation points in the domain of interest is accomplished using gradient descent optimization, which yields the strong-form solution to the partial differential equations. In the theory of elasticity, the cost function can be also formulated using the total potential energy (Jiang et al., 2022; Nguyen-Thanh et al., 2021; Samaniego et al., 2020), which is given in terms of the sum of the stored strain energy density and the work done by the external force, leading to the weak-form solution to the boundary value problems.

While there is an opportunity to contribute to the computational mechanics research, the application of the physics-informed deep neural networks to the micromechanics analysis of complicated unit cell microstructures remains nontrivial. This is due to the intrinsic difficulty of simultaneously satisfying the displacement and traction periodicity conditions along the unit cell edges. In response to this limitation, in our recent work, we developed a deep homogenization neural network model that solves the strong form of heat-conduction differential equations for thermoelastic periodic materials by introducing a periodic neural network layer (Jiang et al., 2023). This periodic layer, first proposed by Dong and Ni (2021), permits an exact representation of C^{∞} periodic conditions of displacement and traction to the machine's precision by utilizing a set of independent sinusoidal functions with adjustable (namely training) parameters.

Herein, the deep homogenization neural network recently developed by the present authors (Jiang et al., 2023) has been further extended for elastic heterogeneous materials of periodic

microstructures. The principal idea is to adopt a two-scale expansion of the displacement field of a repeating unit cell in terms of averaging and fluctuating/periodic contributions dependent on the global and local coordinates, respectively. Fully connected neural network layers are then utilized to estimate the fluctuating/periodic displacement fields by minimizing residuals of Navier's displacement equations of anisotropic microstructured materials for specified macroscopic strains with the help of automatic differentiation. The present work provides a more general framework for elastic composites and demonstrates the applicability of the DHN framework even in the context of three-dimensional microstructures. This extension is valuable for establishing the versatility and robustness of our recently development DHN approach. It also holds promise for further development to account for inelastic mechanisms, such as plasticity, finite deformation, and damage effects, in our future work.

It should be mentioned that the periodicity boundary conditions can be also enforced by a penalty method, wherein the cost function is represented by the sum of PDE residuals, evaluated at the interior of the domain, and the mean square differences of the displacements and tractions on the mirrored unit cell faces, evaluated at the exterior boundaries. Clearly, the penalty method enforces the periodicity boundary conditions only in an approximate manner, which adversely affects the accuracy of the obtained solutions (Cuomo et al., 2022). Meanwhile, there are competing effects between PDE residuals and boundary loss components, which are detrimental to the training of the neural network model and the ensuing neural network predictions.

An additional challenge with the neural network methods is that they suffer from stability and accuracy issues wherein the solutions are highly inhomogeneous and exhibit sharp spatial transitions, owing to the global optimization nature of the cost function in the gradient-based optimization algorithm (Faroughi et al., 2022; Henkes et al., 2022). This is particularly true in the

case of unit cells with locally irregular multiple inclusions, if not with a single inclusion, wherein significant modulus mismatch and jump are encountered, producing marked deformation/stress gradients at the phase interface. For this, we introduce the transfer learning framework (Weiss et al., 2016) in order to obtain the desired network solution.

2. Preliminary

In this section, we present an overview of the basic equations in the homogenization theory of periodic materials and introduce the relevant notations adopted throughout this work. This sets the stage for the development of the physically informed neural network homogenization theory in the next section.

Following the zeroth-order homogenization theory (Bensoussan et al., 1978; Chen et al., 2023; He et al., 2023; Suquet, 1987), the displacement field in the periodic microstructure can be partitioned into averaging and fluctuating contributions dependent on the global and local coordinates $\mathbf{x}=(x_1,x_2,x_3)$ and $\mathbf{y}=(y_1,y_2,y_3)$ as follows,

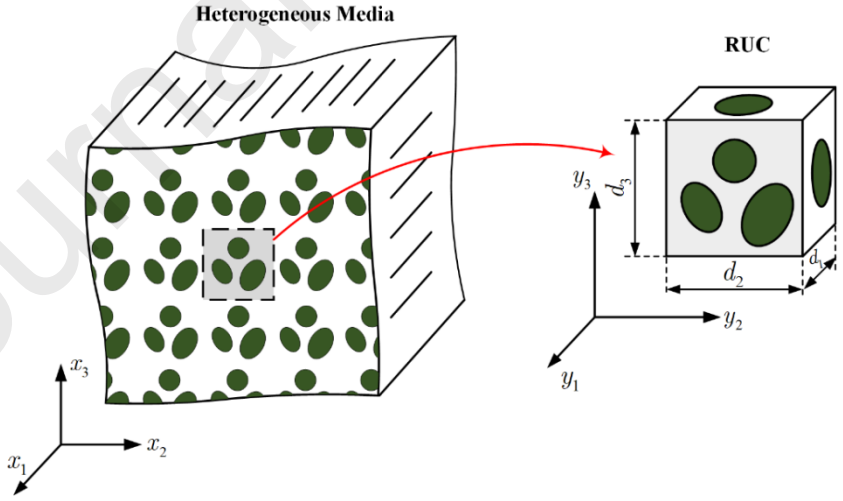


Figure 1 A multiphase periodic array characterized by the smallest building block or the unit cell

$$u_i(\mathbf{x})=\bar{\varepsilon}_{ij}x_j+u_i'(\mathbf{y}) \tag{2}$$

with the fluctuations induced by material heterogeneity and uniform loading defined by the specified average strain $\bar{\varepsilon}_{ij}$. The unit cell problem is then solved subjected to the periodicity displacement and traction conditions (Pindera et al., 2009):

$$u_i\left(\mathbf{x}_\mathrm{o}+\mathbf{d}\right)=u_i\left(\mathbf{x}_\mathrm{o}\right)+\bar{\varepsilon}_{ij}d_j,\quad t_i\left(\mathbf{x}_\mathrm{o}+\mathbf{d}\right)+t_i\left(\mathbf{x}_\mathrm{o}\right)=0 \tag{3}$$

where $i=1,2,3$, $\left(\mathbf{x}_\mathrm{o},\mathbf{x}_\mathrm{o}+\mathbf{d}\right)\in S$, S is the unit cell boundary, $\mathbf{d}=\left(d_1,d_2,d_3\right)$ denotes a characteristic distance that defines the unit cell array microstructure. Tractions t_i are related to stresses from Cauchy's relations $t_i=\sigma_{ji}\cdot n_j$, with $\left(n_1,n_2,n_3\right)$ representing the unit normal vector to the boundary. The stresses are related to strains through Hooke's law of elastic materials $\sigma_{ij}\left(\mathbf{y}\right)=C_{ijkl}\left(\mathbf{y}\right)\varepsilon_{kl}\left(\mathbf{y}\right)$ and strains are related to displacements through the strain-displacement equations:

$$\varepsilon_{ij}\left(\mathbf{y}\right)=\bar{\varepsilon}_{ij}+\varepsilon'_{ij}\left(\mathbf{y}\right)=\bar{\varepsilon}_{ij}+\frac{1}{2}\left(\frac{\partial u'_i}{\partial y_j}+\frac{\partial u'_j}{\partial y_i}\right) \tag{4}$$

where $\varepsilon'_{ij}\left(\mathbf{y}\right)$ denotes fluctuating strain induced by the heterogeneity.

For linearly elastic periodic materials comprised of distinct phases, the governing differential equations, or the so-called Navier's displacement equations, in the absence of body force and inertia effect read:

$$\begin{aligned}
&\frac{\partial}{\partial y_1}\left(C_{11}\left(\bar{\varepsilon}_{11}+\frac{\partial u'_1}{\partial y_1}\right)+C_{12}\left(\bar{\varepsilon}_{22}+\frac{\partial u'_2}{\partial y_2}\right)+C_{13}\left(\bar{\varepsilon}_{33}+\frac{\partial u'_3}{\partial y_3}\right)\right)+\frac{\partial}{\partial y_2}C_{66}\left(2\bar{\varepsilon}_{12}+\frac{\partial u'_1}{\partial y_2}+\frac{\partial u'_2}{\partial y_1}\right)+\frac{\partial}{\partial y_3}C_{55}\left(2\bar{\varepsilon}_{13}+\frac{\partial u'_1}{\partial y_3}+\frac{\partial u'_3}{\partial y_1}\right)=0\\
&\frac{\partial}{\partial y_2}\left(C_{21}\left(\bar{\varepsilon}_{11}+\frac{\partial u'_1}{\partial y_1}\right)+C_{22}\left(\bar{\varepsilon}_{22}+\frac{\partial u'_2}{\partial y_2}\right)+C_{23}\left(\bar{\varepsilon}_{33}+\frac{\partial u'_3}{\partial y_3}\right)\right)+\frac{\partial}{\partial y_1}C_{66}\left(2\bar{\varepsilon}_{12}+\frac{\partial u'_1}{\partial y_2}+\frac{\partial u'_2}{\partial y_1}\right)+\frac{\partial}{\partial y_3}C_{44}\left(2\bar{\varepsilon}_{23}+\frac{\partial u'_2}{\partial y_3}+\frac{\partial u'_3}{\partial y_2}\right)=0\\
&\frac{\partial}{\partial y_3}\left(C_{31}\left(\bar{\varepsilon}_{11}+\frac{\partial u'_1}{\partial y_1}\right)+C_{32}\left(\bar{\varepsilon}_{22}+\frac{\partial u'_2}{\partial y_2}\right)+C_{33}\left(\bar{\varepsilon}_{33}+\frac{\partial u'_3}{\partial y_3}\right)\right)+\frac{\partial}{\partial y_1}C_{55}\left(2\bar{\varepsilon}_{13}+\frac{\partial u'_1}{\partial y_3}+\frac{\partial u'_3}{\partial y_1}\right)+\frac{\partial}{\partial y_2}C_{44}\left(2\bar{\varepsilon}_{23}+\frac{\partial u'_3}{\partial y_3}+\frac{\partial u'_2}{\partial y_2}\right)=0
\end{aligned} \tag{5}$$

Finally, following the classical homogenization theory, the homogenized constitutive equations of the unit cell with a volume $V=d_1d_2d_3$, under small deformation assumption, can be

expressed in terms of effective constants C^*_{ijkl}, designated by an asterisk, Eq. (6). This equation relates the macroscopic stresses $\bar{\sigma}_{ij}$ to the macroscopic strains $\bar{\varepsilon}_{ij}$ evaluated as volume-averaging of their corresponding local quantities as follows:

$$\bar{\sigma}_{ij} = C^*_{ijkl}\bar{\varepsilon}_{ijkl} \tag{6}$$

where

$$\bar{\sigma}_{ij} = \frac{1}{V}\int_V \sigma_{ij}\mathrm{d}V, \quad \bar{\varepsilon}_{ij} = \frac{1}{V}\int_V \varepsilon_{ij}\mathrm{d}V \tag{7}$$

The homogenized stiffness tensor C^*_{ijkl} of the unit cell is obtained by sequentially applying one unit macroscopic strain at a time to solve iteratively the corresponding fluctuating displacements, hence the local stress and strain fields are obtained. The values of C^*_{ijkl} in each column correspond to the macroscopic stresses for the unit cell under a given unit applied strain.

3. Deep homogenization neural network

The principle idea of the physically informed deep homogenization network is to use the neural network as a function approximator (Hornik et al., 1989; Raissi et al., 2019), namely to estimate a continuous function that maps the spatial coordinates $\left(y_1, y_2, y_3\right)$ in the unit cell domain V to the corresponding fluctuating displacements $\left(u_1', u_2', u_3'\right)$ with specified applied average strain $\bar{\boldsymbol{\varepsilon}}$ and the fourth-order stiffness tensors $\mathbf{C}\left(\mathbf{y}\right)$ as follows:

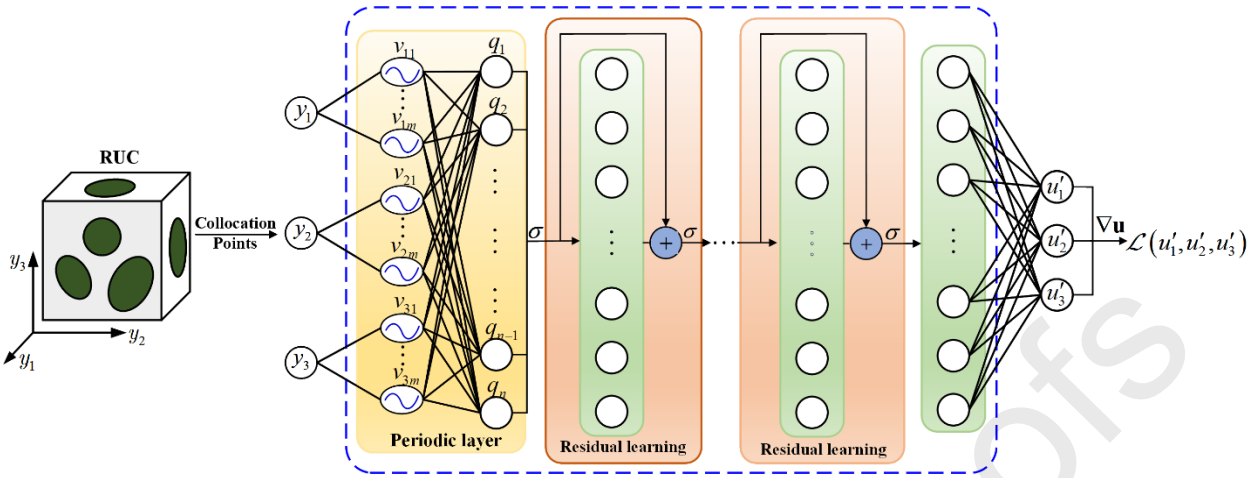


Figure 2 Physically informed deep residual neural network acting as micromechanics model of periodic microstructured materials

$$\left(u_1', u_2', u_3'\right) = \underset{\bar{\boldsymbol{\varepsilon}},\, \mathbf{C}(\mathbf{y})}{\mathrm{DHN}}\left(y_1, y_2, y_3\right), \quad \forall \left(y_1, y_2, y_3\right) \in V \tag{8}$$

i.e., a mapping from $\mathbb{R}^3 \to \mathbb{R}^3$. In the above equation, $\left(u_1', u_2', u_3'\right)$ are periodic functions in the three-dimensional space $\left|y_1\right| \le d_1/2$, $\left|y_2\right| \le d_2/2$, $\left|y_3\right| \le d_3/2$.

Towards this end, a feedforward network with residual learning is constructed, in which the network is fed directly with coordinates of a set of random collocation points sampled over the unit cell domain, as illustrated in Figure 2. The output associated with the jth neuron in the ith layer is formulated as:

$$q_j^l = f\left(\sum_{i=1}^{n} W_{ij}^l q_j^{l-1} + B_j^l + q_j^{l-1}\right) \tag{9}$$

where f indicates nonlinear activation function. q_j^{l-1} is associated with the output of the jth neuron of the $(l-1)$th layer. W_{ij}^l and B_j^l are the weights and biases acting in the lth layer. The parameters of the neural network $\boldsymbol{\theta}^*$ are determined by minimizing a loss function $\mathcal{L}\left(u_1', u_2', u_3'\right)$. This loss function can be expressed as the sum of the residuals of Navier's displacement equations evaluated at a set of collocation points (N_s) :

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\text{ArgMin}}\ \mathcal{L}\left(u_1', u_2', u_3'\right) = \mathcal{L}_{PDE_1} + \mathcal{L}_{PDE_2} + \mathcal{L}_{PDE_3} \tag{10}$$

where

$$\begin{aligned}\mathcal{L}_{PDE_1} = \frac{1}{N_s}\sum_{i=1}^{N_s}\Bigg\{&\frac{\partial}{\partial y_1}\left(C_{11}\left(\overline{\varepsilon}_{11}+\frac{\partial u_1'}{\partial y_1}\right)+C_{12}\left(\overline{\varepsilon}_{22}+\frac{\partial u_2'}{\partial y_2}\right)+C_{13}\left(\overline{\varepsilon}_{33}+\frac{\partial u_3'}{\partial y_3}\right)\right)\\&+\frac{\partial}{\partial y_2}C_{66}\left(2\overline{\varepsilon}_{12}+\frac{\partial u_1'}{\partial y_2}+\frac{\partial u_2'}{\partial y_1}\right)+\frac{\partial}{\partial y_3}C_{55}\left(2\overline{\varepsilon}_{13}+\frac{\partial u_1'}{\partial y_3}+\frac{\partial u_3'}{\partial y_1}\right)\Bigg\}^2\end{aligned} \tag{11}$$

$$\begin{aligned}\mathcal{L}_{PDE_2} = \frac{1}{N_s}\sum_{i=1}^{N_s}\Bigg\{&\frac{\partial}{\partial y_2}\left(C_{21}\left(\overline{\varepsilon}_{11}+\frac{\partial u_1'}{\partial y_1}\right)+C_{22}\left(\overline{\varepsilon}_{22}+\frac{\partial u_2'}{\partial y_2}\right)+C_{23}\left(\overline{\varepsilon}_{33}+\frac{\partial u_3'}{\partial y_3}\right)\right)\\&+\frac{\partial}{\partial y_1}C_{66}\left(2\overline{\varepsilon}_{12}+\frac{\partial u_1'}{\partial y_2}+\frac{\partial u_2'}{\partial y_1}\right)+\frac{\partial}{\partial y_3}C_{44}\left(2\overline{\varepsilon}_{23}+\frac{\partial u_2'}{\partial y_3}+\frac{\partial u_3'}{\partial y_2}\right)\Bigg\}^2\end{aligned} \tag{12}$$

$$\begin{aligned}\mathcal{L}_{PDE_3} = \frac{1}{N_s}\sum_{i=1}^{N_s}\Bigg\{&\frac{\partial}{\partial y_3}\left(C_{31}\left(\overline{\varepsilon}_{11}+\frac{\partial u_1'}{\partial y_1}\right)+C_{32}\left(\overline{\varepsilon}_{22}+\frac{\partial u_2'}{\partial y_2}\right)+C_{33}\left(\overline{\varepsilon}_{33}+\frac{\partial u_3'}{\partial y_3}\right)\right)\\&+\frac{\partial}{\partial y_1}C_{55}\left(2\overline{\varepsilon}_{13}+\frac{\partial u_1'}{\partial y_3}+\frac{\partial u_3'}{\partial y_1}\right)+\frac{\partial}{\partial y_2}C_{44}\left(2\overline{\varepsilon}_{23}+\frac{\partial u_3'}{\partial y_3}+\frac{\partial u_2'}{\partial y_2}\right)\Bigg\}^2\end{aligned} \tag{13}$$

In the above equations, the displacement partial derivatives in the loss components are computed directly using automatic differentiation (Baydin et al., 2018). The optimization process requires the use of a backpropagation algorithm, in which the loss function is minimized iteratively with respect to the weights $\mathbf{W}$ and biases $\mathbf{B}$ in each layer. One of the most common and the most straightforward optimizers utilized in machine learning is gradient descent:

$$\begin{aligned}W_{ij}^{k+1} &= W_{ij}^{k} - \beta\frac{\partial\mathcal{L}}{\partial W_{ij}^{k}}\\ B_{i}^{k+1} &= B_{i}^{k} - \beta\frac{\partial\mathcal{L}}{\partial B_{i}^{k}}\end{aligned} \tag{14}$$

where β denotes the learning rate which is a crucial parameter affecting the neural network training process. The superscripts k and $(k+1)$ indicate the kth and $(k+1)th$ iterations

respectively. The gradients of the loss function are obtained using the chain rule. The weights $\mathbf{W}$ and biases $\mathbf{B}$ are updated during the backpropagation and converge toward a minimized loss value after a number of iterations.

It is widely admitted that periodicity boundary conditions in this specific application play an essential role in yielding an accurate neural network solution. The implementation of the periodicity boundary conditions, however, is far from trivial. Following Dong and Ni (2021), the imposition of the periodicity boundary conditions is accomplished by introducing a periodic neural network layer. This layer permits an exact representation of C^{∞} periodic conditions of displacement and traction by utilizing a set of independent sinusoidal functions with adjustable (namely training) parameters. For a cubic unit cell in the three-dimensional (y_1, y_2, y_3) space with periods of d_1, d_2 and d_3 in y_1, y_2 and y_3 directions, respectively, the periodic layer is defined as (Dong and Ni, 2021):

$$\begin{aligned}
&v_{1i}(y_1) = f\left[A_{1i}\cos(\omega_1 y_1 + \phi_1) + c_{1i}\right],\\
&v_{2i}(y_2) = f\left[A_{2i}\cos(\omega_2 y_2 + \phi_2) + c_{2i}\right],\\
&v_{3i}(y_3) = f\left[A_{3i}\cos(\omega_3 y_3 + \phi_3) + c_{3i}\right],\\
&q_j(y_1, y_2, y_3) = f\left[\sum_{i=1}^{m} v_{1i}(y_1) W_{ij}^{(1)} + \sum_{i=1}^{m} v_{2i}(y_2) W_{ij}^{(2)} + \sum_{i=1}^{m} v_{3i}(y_3) W_{ij}^{(3)} + B_j\right]
\end{aligned} \tag{15}$$

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

$$\omega_1 = \frac{2\pi}{d_1}, \quad \omega_2 = \frac{2\pi}{d_2}, \quad \omega_3 = \frac{2\pi}{d_3} \tag{16}$$

are constants with prescribed periods of d_1, d_2 and d_3, respectively. The periodic layer defined in Eq. (15) ensures the predicted fluctuating displacements and their partial derivatives (namely, strains) inherently adhere to the periodicity conditions, hence the periodicity boundary conditions of tractions are automatically imposed considering the periodic distribution of constituent phases.

The nonlinear activation function f ensures that $v_{1i}(y_1), v_{2i}(y_2)$ and $v_{3i}(y_3)$ contain not only the frequency ω_1, ω_2 and ω_3, but also the 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. In summary, the adjustable (or training) parameters in the periodic layer include:

$$A_{1i}, A_{2i}, A_{3i}, \phi_{1i}, \phi_{2i}, \phi_{3i}, c_{1i}, c_{2i}, c_{3i}, W_{ij}^{(1)}, W_{ij}^{(2)}, W_{ij}^{(3)}, B_j \tag{17}$$

In the case of unidirectional composites reinforced by continuous fibers in the y_1 (namely x_1) direction, the null displacement partial derivatives in the pertinent direction, i.e., $\partial u_i'/\partial y_1 \equiv 0$, hold for any combination of macroscopic strains. The response of such composites can be characterized by periodically repeating material microstructures with two-dimensional periodicity in the $y_2 - y_3$ plane. For this, the remaining Navier's displacement equations are uncoupled for inplane (loaded by $\overline{\boldsymbol{\varepsilon}}_{in} = [\overline{\varepsilon}_{11}, \overline{\varepsilon}_{22}, \overline{\varepsilon}_{33}, 2\overline{\varepsilon}_{23}]^{\mathrm{T}}$) and antiplane (loaded by $\overline{\boldsymbol{\varepsilon}}_{out} = [2\overline{\varepsilon}_{12}, 2\overline{\varepsilon}_{13}]^{\mathrm{T}}$) shear loadings. The corresponding unit cell problems can be solved independently under the generalized plane strain constraint $\varepsilon_{11} = \overline{\varepsilon}_{11}$ as follows:

$$(u_2', u_3') = \underset{\overline{\boldsymbol{\varepsilon}}_{in},\, \mathbf{C}(\mathbf{y})}{\mathrm{DHN}}(y_2, y_3)_{\mathrm{in}}, \quad (u_1') = \underset{\overline{\boldsymbol{\varepsilon}}_{out},\, \mathbf{C}(\mathbf{y})}{\mathrm{DHN}}(y_2, y_3)_{\mathrm{out}} \qquad \forall (y_2, y_3) \in V \tag{18}$$

i.e., mapping from $\mathbb{R}^2 \to \mathbb{R}^2$ for in-plane loading and $\mathbb{R}^2 \to \mathbb{R}^1$ for out-of-plane loading. This significantly reduces the computational effort for computing the whole set of homogenized moduli and local stress fields under arbitrary multi-axial loading conditions without resorting to three-dimensional unit cell simulations. It should be pointed out that the collocation points in heterogeneous microstructures can be generated randomly simply using Monte-Carlo simulation or obtained from real-world μCT -scans (Henkes et al., 2022).

4. Unidirectional composites

The unit cells of the investigated unidirectional boron/aluminum (B/Al) composites with two-dimensional periodicity in a state of generalized plane strain are shown in Figure 3. The Lamé's constants of the fiber and matrix phases (Pindera and Bansal, 2007) are $\lambda_f = 111.11\text{GPa}$, $\mu_f = 166.67\text{GPa}$, and $\lambda_m = 52.84\text{GPa}$, $\mu_m = 27.22\text{GPa}$, respectively. Since the neural network functions are infinitely differentiable, approximating displacement functions for the unit cell problem with stiff stiffness transition is very demanding, and it typically requires a significantly large number of training iterations to obtain a desirable solution. Therefore, as suggested by (Henkes et al., 2022), in an attempt to resolve this issue, the stiffness distributions in the DHN approach are artificially smoothened across the fiber/matrix interface. For an N-fiber composites with elliptical cross-sectional shapes, if the fiber and matrix phases are both isotropic, the distributions of material properties are expressed in terms of Lamé functions $\lambda(\mathbf{y})$ and $\mu(\mathbf{y})$ in the form:

$$\lambda(\mathbf{y}) = \sum_{i=1}^{N} \lambda_1 \left[\lambda_2 + \tanh \frac{1 - \left[\frac{\left(y_2 - o_2^i\right)\cos\alpha^i + \left(y_3 - o_3^i\right)\sin\alpha^i}{b_2^i} \right]^2 - \left[\frac{\left(y_2 - o_2^i\right)\sin\alpha^i - \left(y_3 - o_3^i\right)\cos\alpha^i}{b_3^i} \right]^2}{\delta} \right] + \lambda_3 \tag{19}$$

$$\mu(\mathbf{y}) = \sum_{i=1}^{N} \mu_1 \left[\mu_2 + \tanh \frac{1 - \left[\frac{\left(y_2 - o_2^i\right)\cos\alpha^i + \left(y_3 - o_3^i\right)\sin\alpha^i}{b_2^i} \right]^2 - \left[\frac{\left(y_2 - o_2^i\right)\sin\alpha^i - \left(y_3 - o_3^i\right)\cos\alpha^i}{b_3^i} \right]^2}{\delta} \right] + \mu_3 \tag{20}$$

where $\left(o_2, o_3\right)^{(i)}$, b_2^i and b_3^i denote the origins, major and minor axes of the ith fiber, respectively. α^i denotes the angle inclination between the major axis of the ith fiber and the horizontal axis.

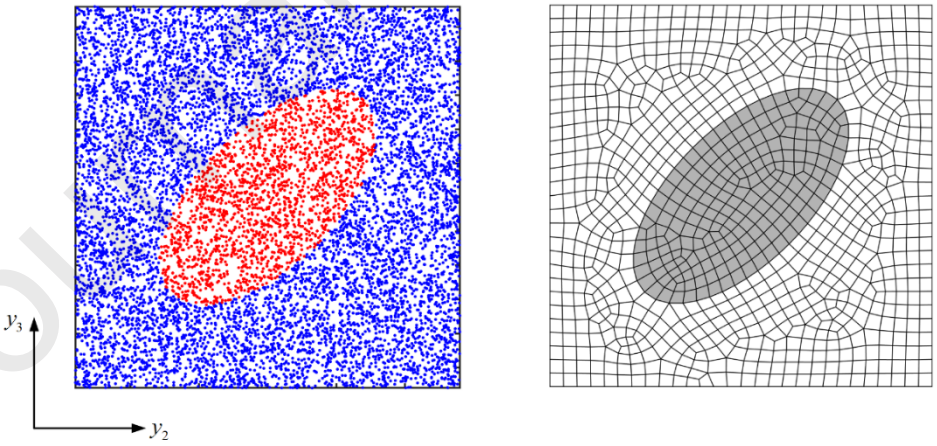
$\tanh(\cdot)$ denotes the hyperbolic tangent function. $\lambda_1, \lambda_2, \lambda_3$ and μ_1, μ_2, μ_3 are materials constants that are related directly to the fiber and matrix Lamé parameters. If λ_2 and μ_2 are both set to 1 GPa, we have

$$\mu_3 = \mu_m, \quad \mu_1 = \left(\mu_f - \mu_m\right)/2, \quad \lambda_3 = \lambda_m, \quad \lambda_1 = \left(\lambda_f - \lambda_m\right)/2, \tag{21}$$

where λ_m, μ_m and λ_f, μ_f the matrix and fiber Lamé's constants, respectively. δ denotes the smoothness of the material transition and it is set to be $\delta = 0.05$ in this manuscript unless otherwise stated.

The accuracy of the DHN relies on the network depth, the number of neurons per layer, and collocation points entering the loss function, which are equivalent to the mesh density in the finite-element method. Thereafter, in order to gain a deeper insight into the effect of these hyperparameters, we perform a number of numerical experiments to understand how they affect neural network performance.

4.1 Effect of network parameters



(a) (b)

Figure 3 (a) Collocation point distributions generated using Monte-Carlo simulation for evaluating PDE residuals; (b) finite-element unit cell discretization for comparison with DHN approach

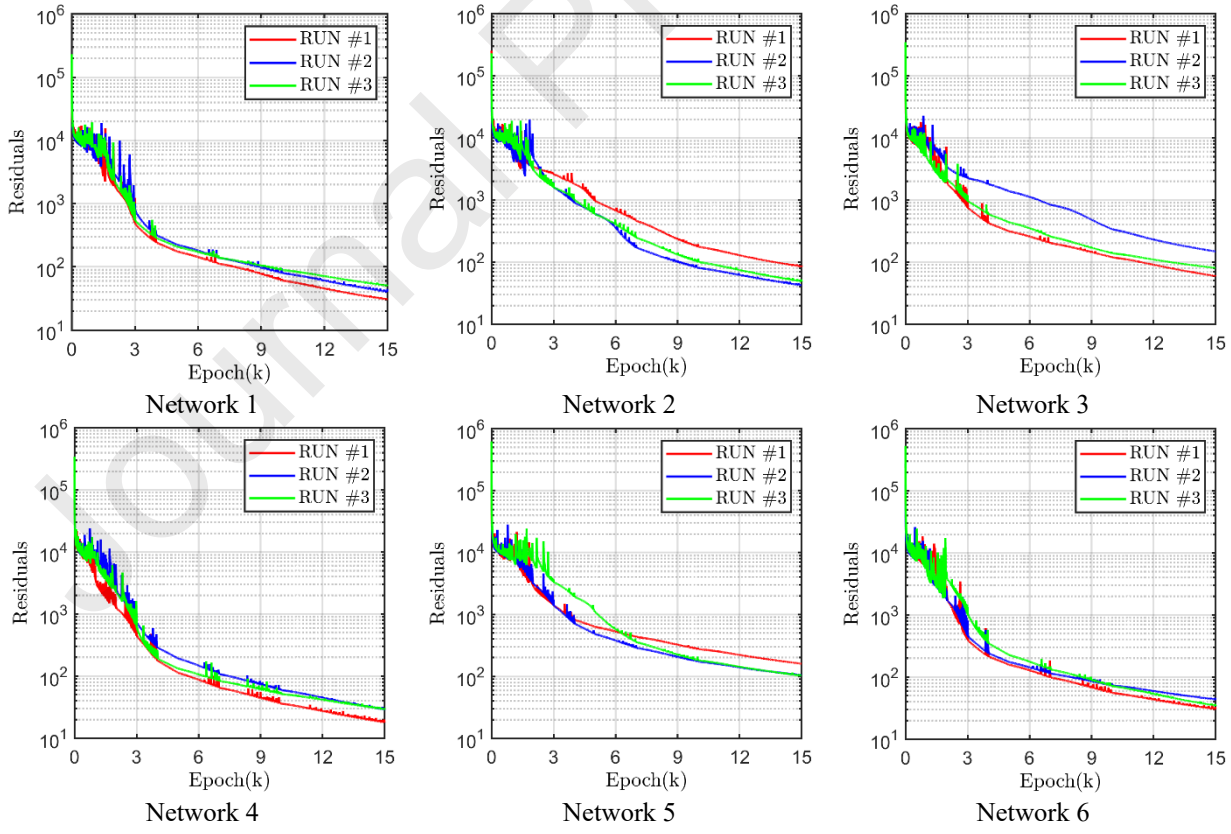
As shown in Figure 3, the first example is characterized by a simple square unit cell $\left(d_2 = d_3\right)$ reinforced by an elliptical cross-sectional fiber with an aspect ratio of $b_2 : b_3 = 2:1$. The major axis of the fiber is rotated by $\alpha = \pi/4$ w.r.t. the horizontal axis. The overall volume fraction of the boron fiber is 20%. Figure 3a shows training collocation point distributions generated randomly using Monte-Carlo simulations. The training dataset contains 10k collocation points which were utilized for evaluating the loss function in the network model. A separate and hold-out validation dataset with 280×280 collocation points was used for plotting PDE residuals, displacement, and stress distributions when the network was trained. Figure 3b shows the finite-element mesh refinement which is discretized into 941 Q4 or Q8 elements for comparison with the DHN predictions. For all the numerical experiments presented in this work, the network solutions were obtained with the PyTorch package on Google's Collaboratory cloud platform with Adam Optimizer. The initial learning rate is 0.01 and decreases progressively as a function of the training epoch to avoid the training loss fluctuation. The hyperbolic tangent function was utilized as the activation function. The finite-element computations were performed with an in-house MATLAB code (Cavalcante et al., 2011; Chen et al., 2022).

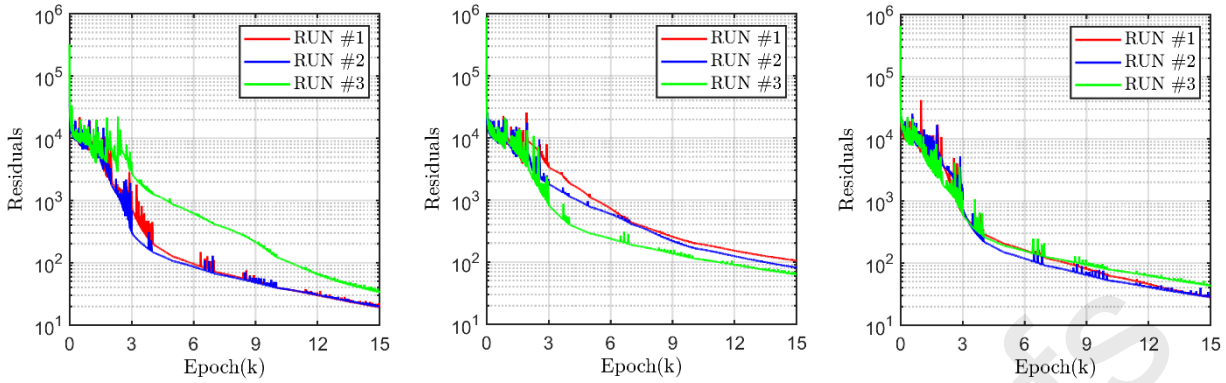
As the first example, a unit transverse normal strain $\bar{\varepsilon}_{22} = 0.1\%$ is prescribed, with the remaining strain components set to zero, since this loading mode is the most demanding. The employed fiber shape and orientation, as well as its dimension relative to the overall unit cell size, give rise to significant fiber-fiber interactions, wherein the effect of periodic boundary conditions on the local displacement and stress field distributions is more important. Hence this system is a good candidate to validate the predictive capability of the network-based homogenization theory relative to the finite-element benchmark solutions.

Nine neural network architectures (Table 1) that cover a wide range of network parameters were used to run the simulations, enabling direct assessment of neural network architecture (i.e., layer and neuron number) on the network performance. Three training runs were performed for each network architecture to illustrate the effect of variations in the seed values of the weights and biases on the neural network convergence.

Table 1 Network parameters for identifying the best network model

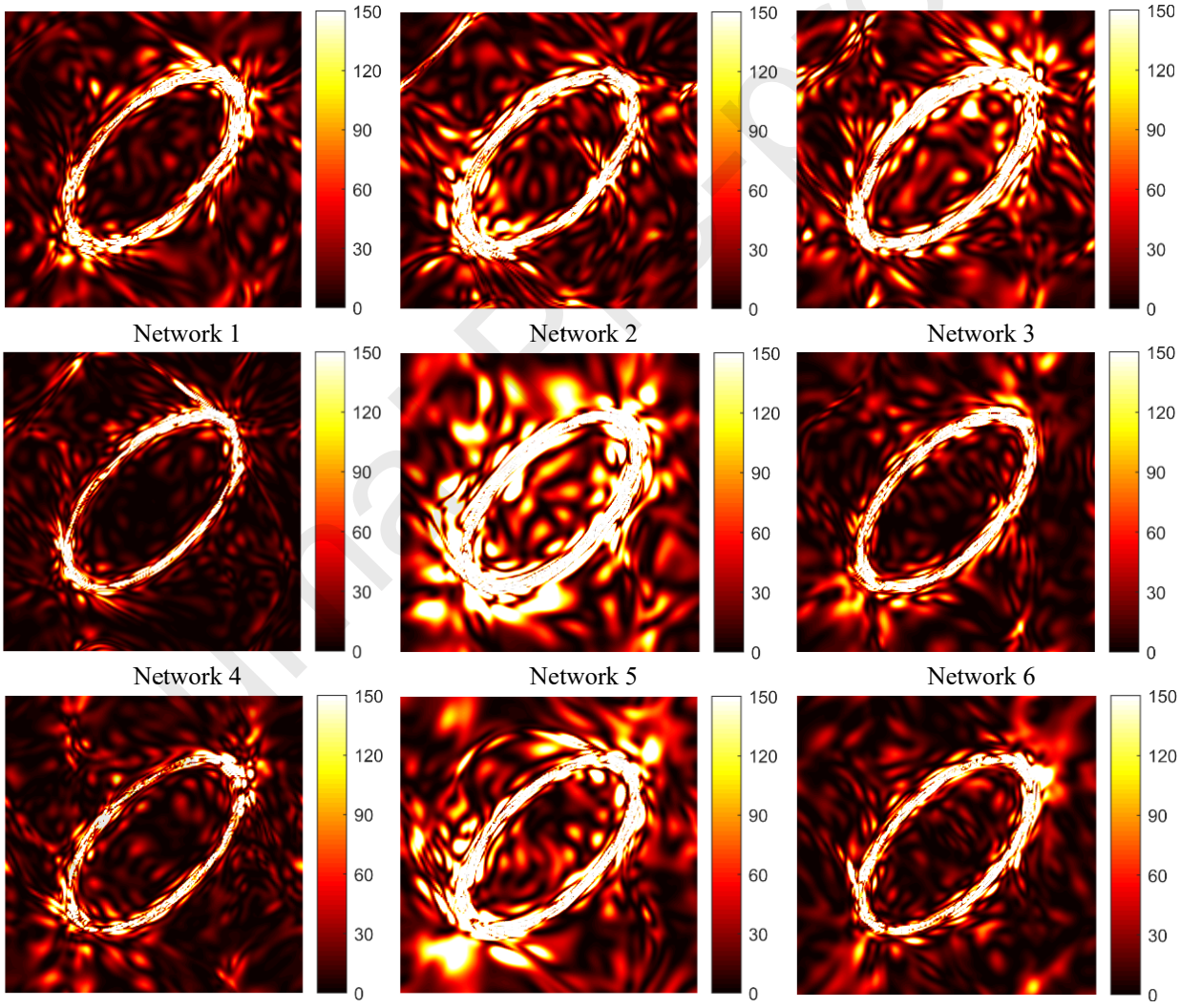
Network	Neurons in the periodic layer	Hidden layers	Hidden neurons	Training parameters
1	10	3	30	3542
2	30	3	30	4862
3	20	2	30	3272
4	20	3	40	6762
5	20	3	20	2242
6	20	3	30	4202
7	20	5	40	10042
8	20	5	20	3082
9	20	5	30	6062





Network 7 Network 8 Network 9

Figure 4 Effect of network depth and hidden neurons on the network performance



Network 1 Network 2 Network 3

Network 4 Network 5 Network 6

Network 7 Network 8 Network 9

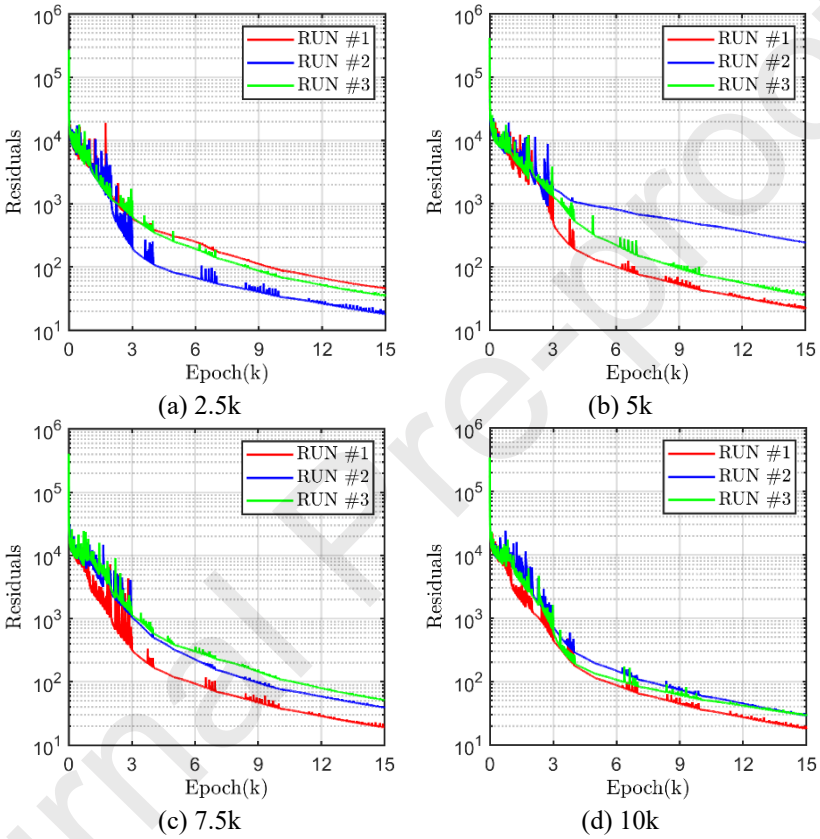
Figure 5 Comparison of pointwise loss distribution (PDE residuals) for the best run of the nine neural network models

Figure 4 presents comparison of the total loss value (Eq. (10)) as a function of the training epoch predicted by the nine neural network architectures. Figure 5 illustrates the pointwise loss distributions after 15k training epochs of the best run for each network architecture evaluated on the separate validation dataset. A cursory look at the loss evolutions for multiple restarts reveals that networks 2, 3, 7, and 8 are more sensitive to the variations in the seeds values for the weights and biases, indicating poorer robustness than that of networks 1, 4, 5, 6, and 9. In addition, the pointwise loss distributions of networks 2, 3, 5, and 8 are markedly higher and more widespread than those of networks 1, 4, 7, 6, and 9. Since the overall minimal loss value for network 6 is an order of magnitude higher than those of networks 1, 4, and 9, the latter networks retain the most promising candidates for the micromechanical analysis of the composites. Network 4 has a higher number of training parameters than networks 1 and 9, hence a better generalization ability may also be expected. In what follows, we chose network 4 for all the numerical experiments in the sequel.

4.2 Effect of dataset size

In this section, we proceed to demonstrate the number of collocation points on the network's accuracy. We re-examine the numerical experiment performed in Section 4.1 but with increasing collocation points, namely 2.5k, 5k, 7.5k, and 10k generated via the Monte Carlo simulations. As before, three training runs were conducted to test the influence of seed values on the network training performance. The corresponding loss functions are shown in Figure 6 and the plots of local stress field σ_{22} distributions for the best runs are shown in Figure 7. It is observed that approximately the same low training loss values are attained after 15k epochs (for the best run), regardless of the size of the collocation points. However, the neural network trained on 2.5k collocation points failed to predict the accurate stress distributions even though the training loss

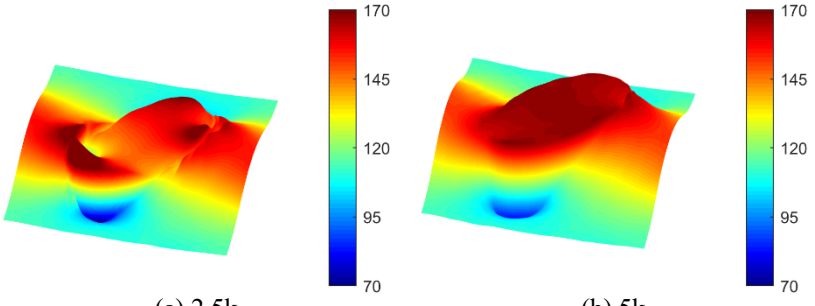
value is small enough. Moreover, the neural network robustness is strongly affected by the size of collocation points, namely, the training loss functions are generally less sensitive to the variations in the seed values of weights and biases with increasing size of collocation points. Based on the good performance of the 10k dataset, it will be employed in generating the results in the sequel unless otherwise stated.



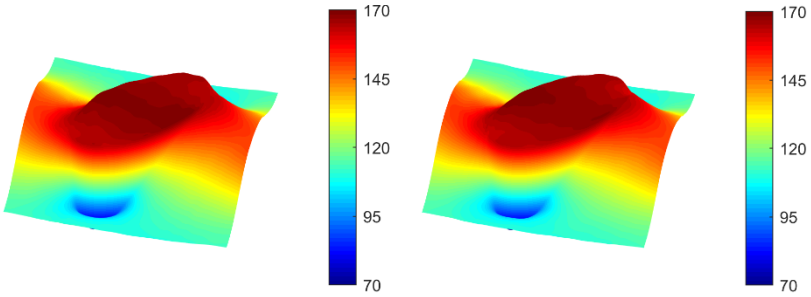
(a) 2.5k (b) 5k

(c) 7.5k (d) 10k

Figure 6 Effect of training dataset size on the network performance



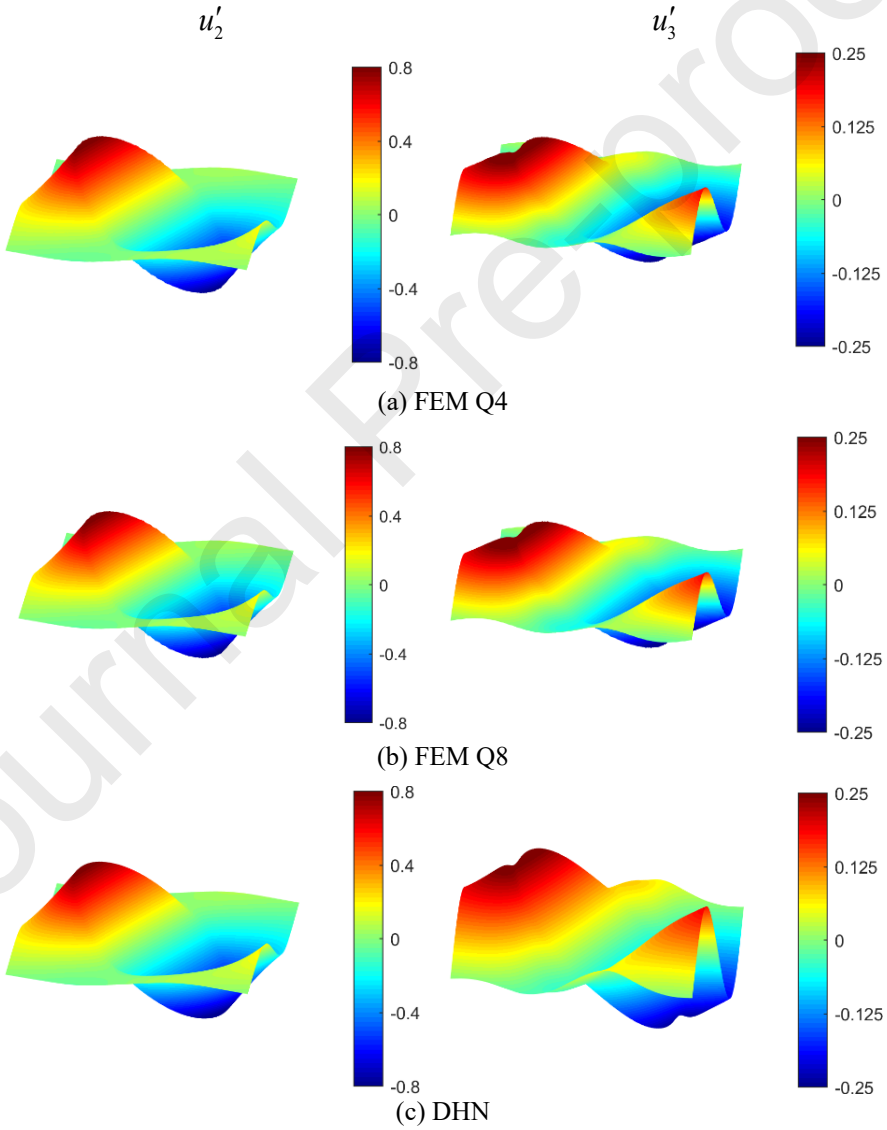
(a) 2.5k (b) 5k



(c) 7.5k (d) 10k

Figure 7 Effect of training dataset size on the predicted local stress σ_{22} distribution

4.3 Verifications



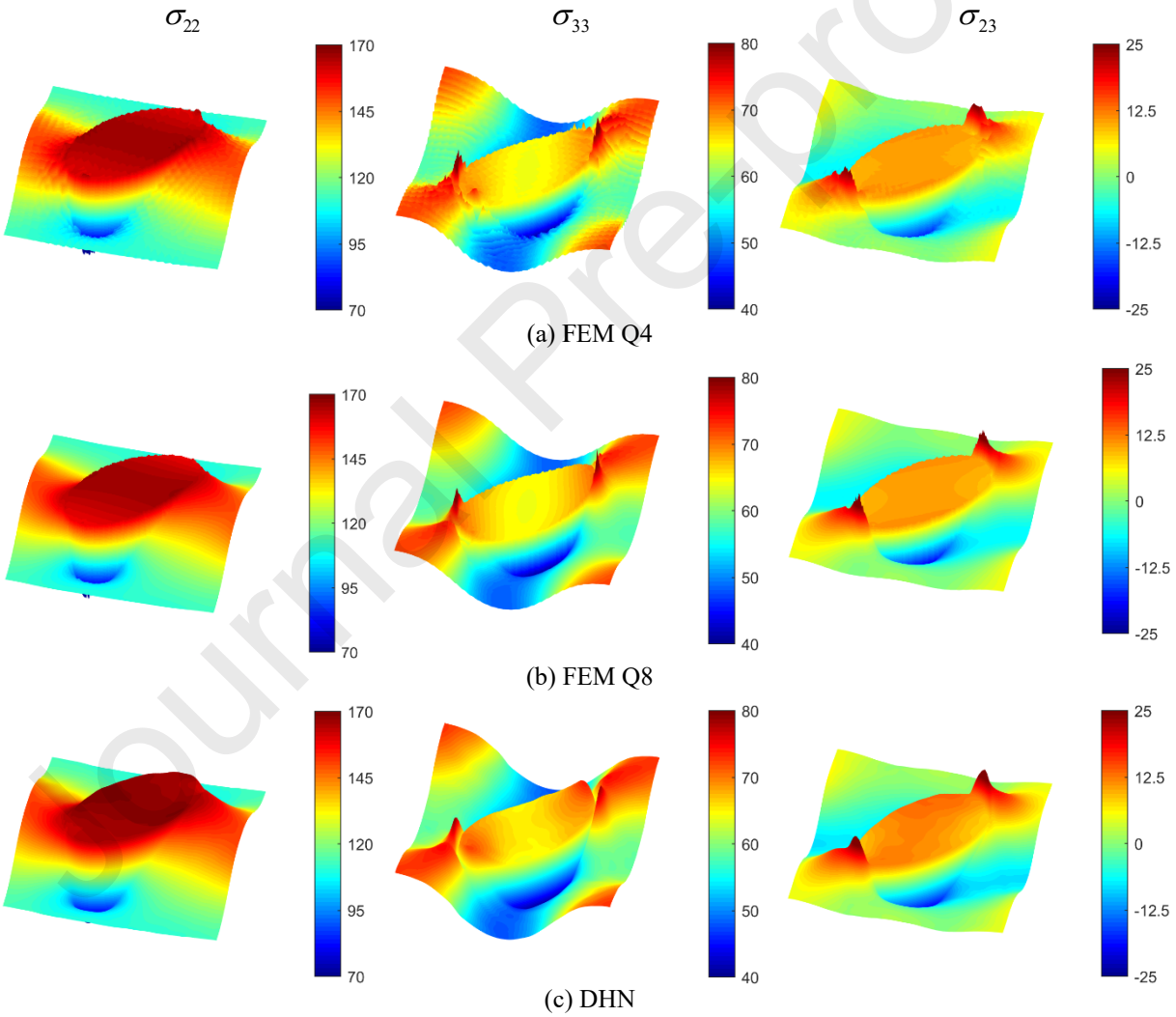
(a) FEM Q4

(b) FEM Q8

(c) DHN

Figure 8 Comparison of u_2' and u_3' displacement field distributions ($10^{-3}\mu m$) predicted by Q4 and Q8 FEM and DHN approaches with the imposition of only one nonzero macroscopic strain component $\bar{\varepsilon}_{22} = 0.1\%$

This section presents verification of the DHN theory's predictive capability to accurately capture both the homogenized moduli and local displacement and stress field distributions. First of all, a comparison of the displacement and stress fields is illustrated, predicted by the DHN with in-house finite-element results which serve as the gold standard. It should be highlighted that the solution methods employed in the network theory (strong-form solution) and the finite-element approach (weak-form solution) are totally different, leading credence to the network theory's rigorous validation and the ensuing conclusions.



(a) FEM Q4

(b) FEM Q8

(c) DHN

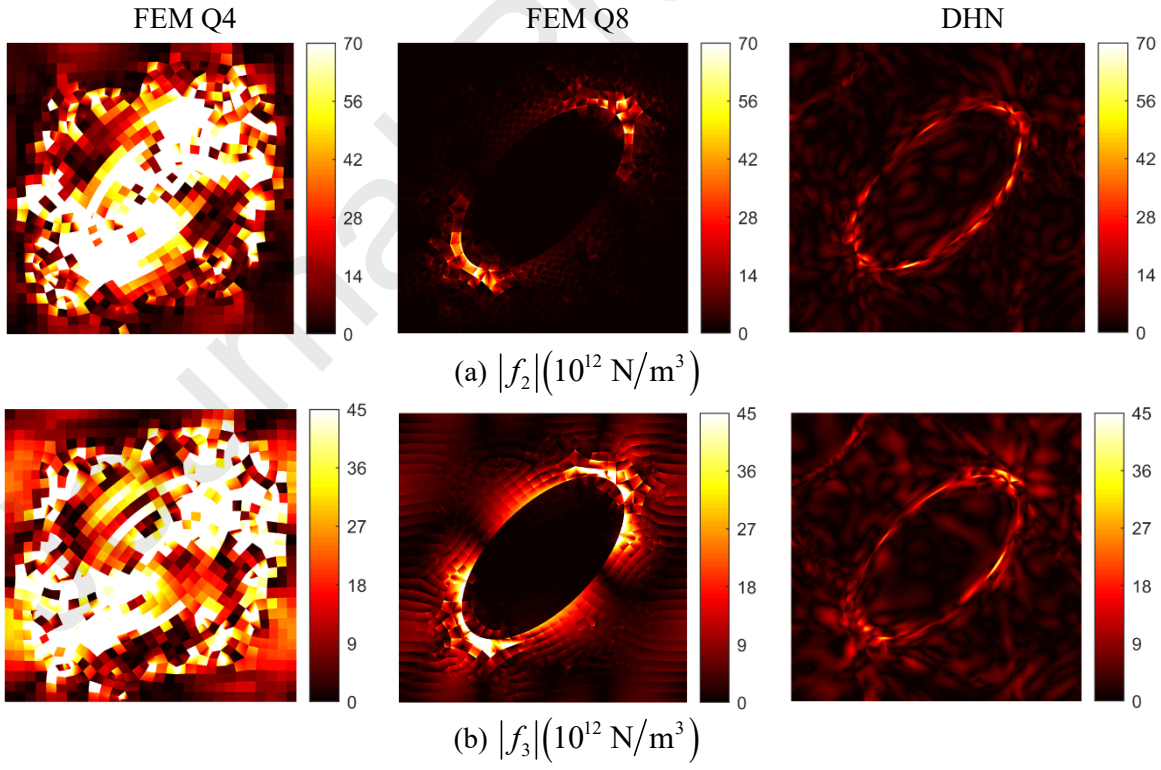
Figure 9 Comparison of local stress field distributions generated by Q4 and Q8 FEM and DHN approaches with the imposition of only one nonzero macroscopic strain component $\bar{\varepsilon}_{22} = 0.1\%$

Under transverse normal strain loading by $\bar{\varepsilon}_{22} = 0.1\%$ with all other strain components remaining null, the fluctuating displacement u_2' field generated by network 4 trained on 10k collocation points is virtually identical to the Q4 and Q8 finite-element predictions. A close look at the displacement u_3' field indicates that the magnitude of u_3' predicted by the DHN scheme is slightly higher than those of Q4 and Q8 finite-element results. Yet, the main characteristics of the displacement distribution are well-captured. We note that the displacement fields for all the approaches are smoothly varying even in the vicinity of or at the fiber/matrix interface.

Figure 9 compares the differences in concomitant σ_{22}, σ_{33} and σ_{23} stress field distributions predicted by the DHN and Q4 and Q8 finite-element simulations. A cursory examination of the three stress components reveals an excellent correlation between the network and finite-element predictions. However, the stress fields computed by the FEM Q4 exhibit apparent nonsmooth spatial variations. These stress discontinuities are even more remarkable in the vicinity of the fiber/matrix interface wherein the highest deformation and stress gradients typically appear. These stress discontinuities arise because the classical finite element method satisfies only the C^0 continuity of the displacements, while the calculated displacement derivatives are not necessarily continuous from one element to another, even with finer meshes. FEM Q8 predicts smoother stress fields relative to the Q4 element predictions due to a higher order of displacement field representation. Nevertheless, the stress discontinuities associated with the finite-element method do not vanish even with higher-order displacement representation. In contrast, the displacements predicted by the neural network model are indefinitely differentiable. Therefore, the obtained stress and stress gradients in the network approach smoothly vary in the entire unit cell domain, without apparent discontinuities observed in the non-traction stress components associated with the pertinent directions.

It should be noted that small differences are observed between the network and finite-element approaches which occur near the vicinity of the fiber/matrix interface. They are attributed to the fact that the DHN utilized smoothed stiffness coefficient distributions across the fiber/matrix interface, Eqs. (19) and (20), leading to the less abrupt stress transition in the interfacial area. The network approach captures the small effect in stiffness distributions with sufficient fidelity.

Figure 10 presents comparison of DHN's absolute residual body forces for the in-plane Navier's equations with Q4 and Q8 finite-element approaches subject to transverse macroscopic strain $\bar{\varepsilon}_{22} = 0.1\%$. As anticipated, the Q4 element exhibits significantly high residual body forces due to its low degree of freedom. The DHN and Q8 elements exhibit comparable accuracy, resulting in reduced residual body forces thanks to their enhanced precision when representing the unit cell solution.

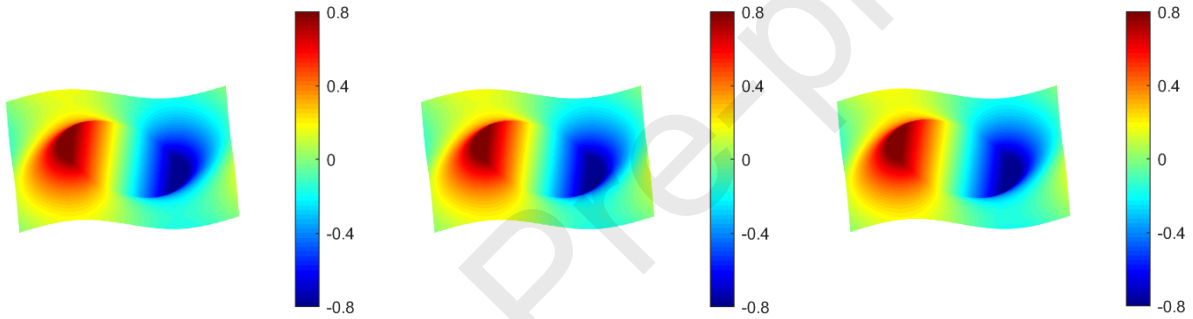


(a) $|f_2|\left(10^{12}\ \mathrm{N/m^3}\right)$

(b) $|f_3|\left(10^{12}\ \mathrm{N/m^3}\right)$

Figure 10 Comparison of DHN's residual body forces with Q4 and Q8 finite-element predictions

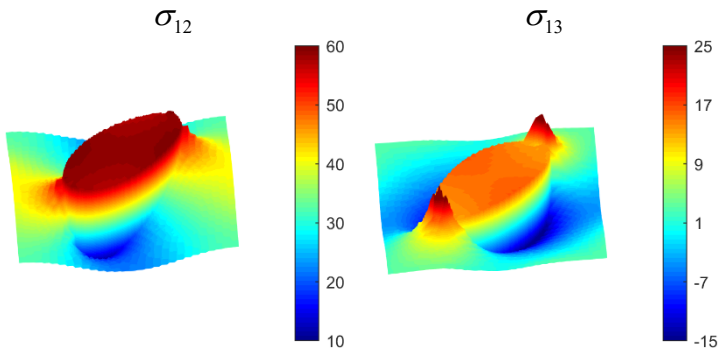
Figures 11 and 12 present comparison of the out-of-plane fluctuating displacement u_1' and axial shear stresses σ_{12} and σ_{13} distributions under axial shear loading by $2\bar{\varepsilon}_{12} = 0.1\%$, generated by the network and finite-element approaches. As before, the DHN and finite-element techniques predict virtually the same displacement distributions, which are smoothly varying in the entire unit cell domain. The stress fields predicted by the two approaches also show a good level of correlation. As demonstrated in the previous example, while the stress fields predicted by the DHN exhibit smooth spatial variation even in the vicinity of the fiber/matrix interface, notable stress discontinuities can be observed in the finite-element approaches under this loading configuration.



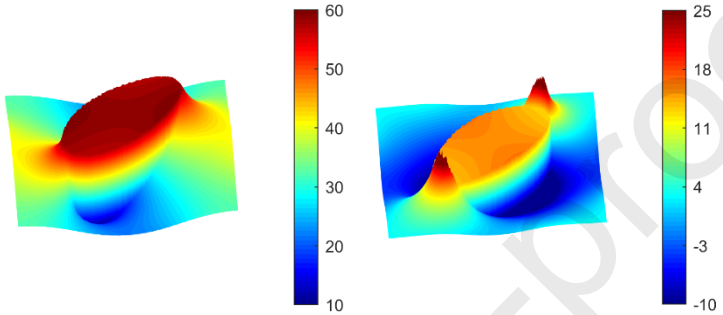
(a) FEM Q4 (b) FEM Q8 (c) DHN

Figure 11 Comparison of u_1' displacement field distributions ($10^{-3}\mu m$) predicted by Q4 and Q8 FEM and DHN under uniaxial shear loading by $2\bar{\varepsilon}_{12} = 0.1\%$

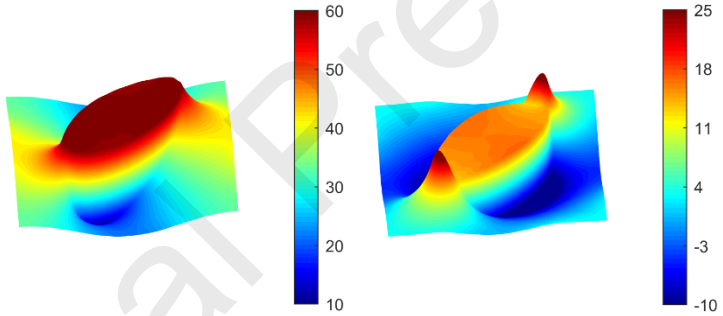
The full set of homogenized properties can be obtained by sequentially applying only one nonzero unit macroscopic strain component at a time and solving the corresponding boundary value problems six times. The values in each column of the homogenized stiffness tensor correspond to the homogenized stresses for each loading case. The DHN- and FEM-computed homogenized properties for the microstructure shown in Figure 3 are given in Eq. (22). It is observed a good agreement between the network and finite-element approaches, with relative differences of less than 1%. The small differences are due to the interfacial stiffness distributions in the DHN technique, further indicating the network method is sufficiently sensitive to correctly capture these small effects.



(a) FEM Q4



(b) FEM Q8



(c) DHN

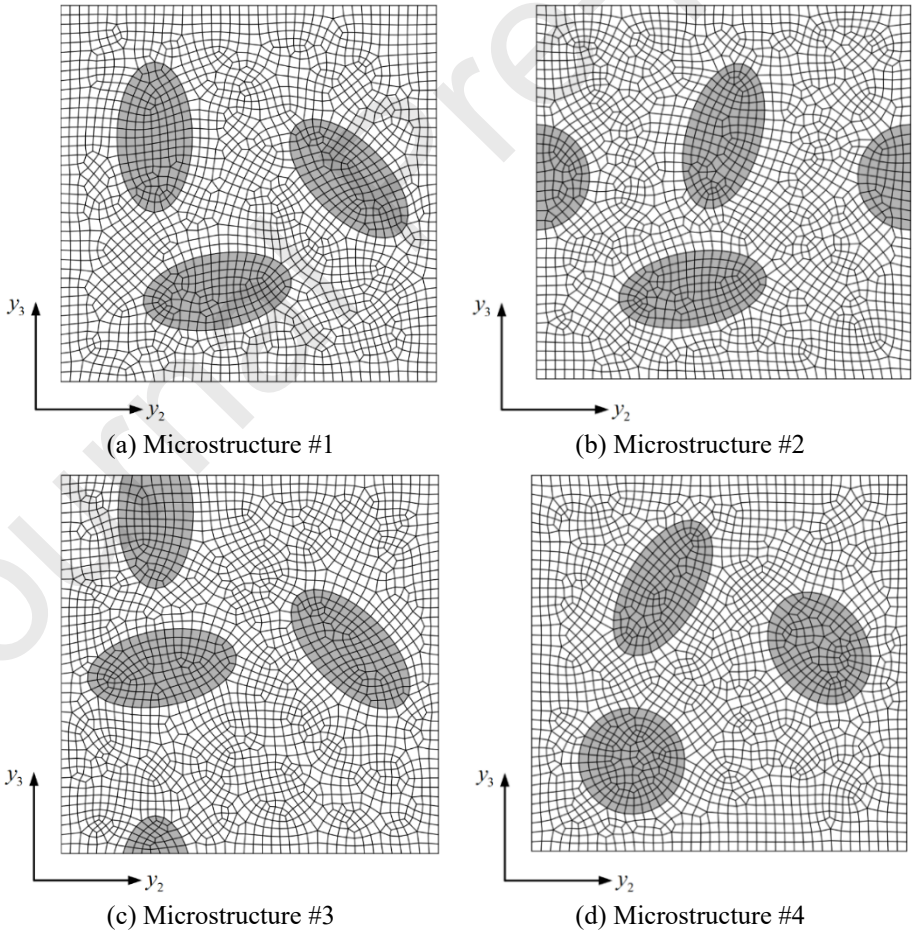
Figure 12 Comparison of axial shear stress distributions predicted by Q4 and Q8 FEM and DHN under uniaxial shear loading by $2\bar{\varepsilon}_{12} = 0.1\%$

$$\mathbf{C}^{*}_{DHN} = \begin{bmatrix} 172.1 & 57.8 & 57.9 & 0.5 & 0 & 0 \\ 57.8 & 132.7 & 61.3 & 1.8 & 0 & 0 \\ 57.8 & 61.2 & 132.7 & 1.8 & 0 & 0 \\ 0.7 & 1.9 & 1.9 & 35.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 37.4 & 2.6 \\ 0 & 0 & 0 & 0 & 2.6 & 37.4 \end{bmatrix} \text{GPa}, \quad \mathbf{C}^{*}_{FEM} = \begin{bmatrix} 172.7 & 57.7 & 57.7 & 0.5 & 0 & 0 \\ 57.7 & 131.8 & 61.0 & 1.7 & 0 & 0 \\ 57.7 & 61.0 & 131.8 & 1.7 & 0 & 0 \\ 0.5 & 1.7 & 1.7 & 34.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 37.0 & 2.5 \\ 0 & 0 & 0 & 0 & 2.5 & 37.0 \end{bmatrix} \text{GPa} \tag{22}$$

5. Multi-inclusion composites

In this section, the modelling and predictive capabilities of the DHN theory are further studied for simulating unidirectional composites with locally-irregular fibers. As shown in Figure 13, while

the overall fiber volume fractions have been kept constant at 18.85%, four random microstructural realizations with different cross-sectional fiber shapes were generated using Eqs. (19) and (20). In order to qualify as an RUC, the fibers near the unit cell boundary are cropped and moved apart, as shown in Figure 13b and Figure 13c. They are precisely added on the opposite side of the unit cell edge such that the periodic positioning of fibers is formed. For the unit cells labeled as microstructure #1~ #4, the finite-element unit cells were constructed with periodic meshes with 1970, 2004, 2084, and 2012 four and eight-noded elements, respectively. For each microstructural realization, we employed 20k randomly sampled collocation points (not shown) for training the neural network model.



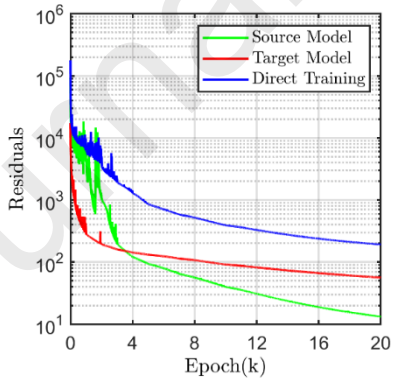
(a) Microstructure #1 (b) Microstructure #2

(c) Microstructure #3 (d) Microstructure #4

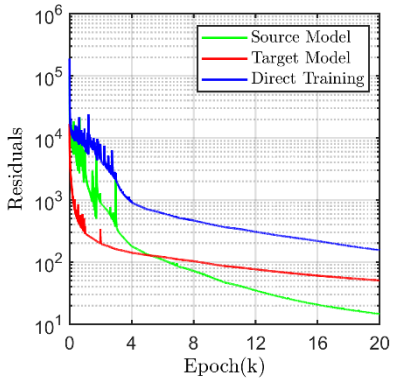
Figure 13 Finite-element mesh discretizations of unit cells with locally irregular fiber distributions and different fiber shapes

We note that the neural network methods as function approximators are known to suffer from stability and accuracy issues where the solution has sharp spatial transitions. The fiber-fiber interactions in random fiber realizations are more important than in the single fiber case, which leads to pronounced deformation and stress gradients inside the unit cell. These interactions pose a marked challenge to the convergence of the neural network loss function, leading to inaccurate neural network predictions. Therefore, the random fiber unit cells provide a more demanding test of the DHN's accuracy and efficiency. Trials with different network architectures have been extensively conducted but have not yielded a satisfactory improvement on this issue.

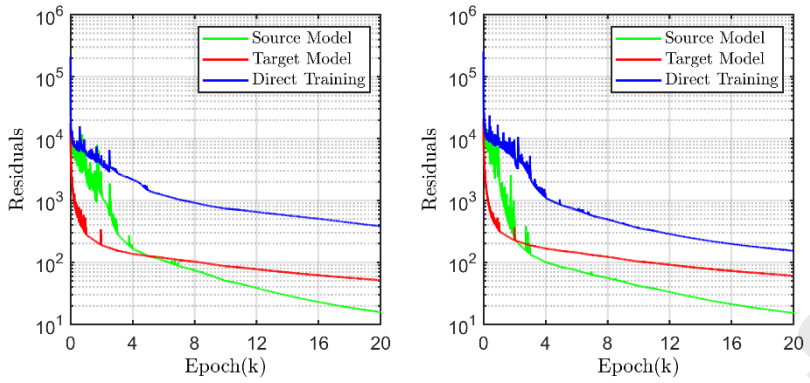
In response to this challenge, a transfer-learning enhanced DHN strategy for simulating the unit cells reinforced with locally irregular fibers is proposed. This is motivated by the observation that when increasing the material transition parameter δ from $\delta = 0.05$ to $\delta = 0.1$ (namely, decreasing the stiffness of material transition in different phases), the spatial transition of the stress in the vicinity of fiber/matrix becomes less abrupt. The overall displacement and stress field distributions, however, are not fundamentally altered.



(a) Microstructure #1



(b) Microstructure #2



(c) Microstructure #3 (d) Microstructure #4

Figure 14 Finite-element mesh discretizations of unit cells with locally irregular fiber distributions

Therefore, instead of directly building the neural network model for a specific random fiber realization with the more demanding parameter $\delta = 0.05$ (denoted as target model) from scratch, we first train an identical neural network model with smoother stiffness transition $\delta = 0.1$ (denoted as source model). When the loss value of the source model is minimized, the target model copies all the pre-trained weights and biases from the source model. Then we assume that the latter has learned sufficiently the common knowledge of displacement and stress fields which are useful for training the target model. In the second step, these weights and biases are fine-tuned during the re-training of the target model. It should be noted that a smaller learning rate should be utilized in the transfer learning procedure to avoid significant updating of the parameters copied from the trained source model.

To provide a direct assessment of the effectiveness of the transfer learning technique, Figure 14 presents the evolution of the loss value as a function of the training epoch for both the source and the target models. For comparison purpose, direct training of the target model from scratch (indicated as direct training in the figures) has also been included in the figure. For all four microstructural realizations, it is clear that the training of the source models always attains the lowest loss values since they are the least demanding because of the smaller stress concentrations around the fiber/matrix interface relative to the target model. On the contrary, the direct training

of the target models from scratch yields the highest ones. Another observation is that the rate of convergence for the neural network loss values with transfer learning is the most rapid at the first three thousand training epochs as their weights are more effective. The loss values of the target models enhanced with transfer learning at every training epoch are remarkably lower than those trained from scratch, indicating that the transfer learning is capable of attaining a better network solution for the multi-inclusion unit cell albeit at the cost of a pre-training process.

Figure 15 presents comparison between the u_2' and u_3' displacement distributions predicted by the various network models and the eight-noded finite-element solution under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ for the selected microstructure #1. As observed, both the source model and the target model with transfer learning provide very good estimates of the displacement fields against the finite-element solutions. Direct training of the target model from scratch fails to provide accurate displacement estimations relative to the baseline solutions.

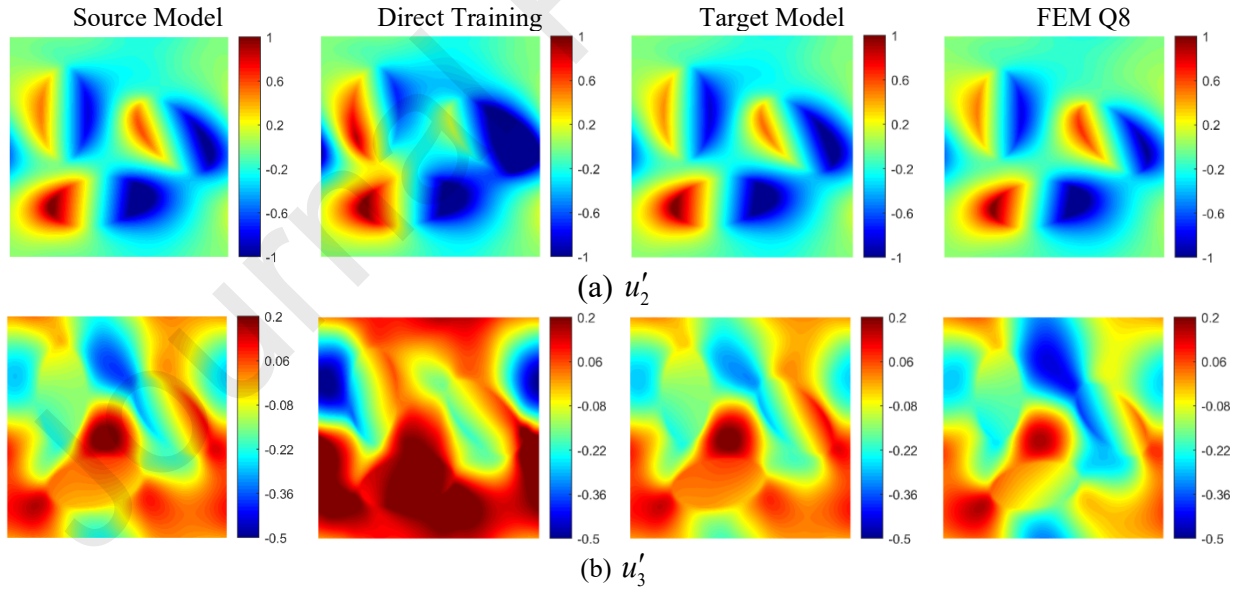


Figure 15 Comparison of u_2' and u_3' fluctuating displacements ($10^{-3}\mu m$) in microstructure #1 predicted by the source model, direct training, target model, and FEM simulations with the imposition of only one nonzero macroscopic strain component $\bar{\varepsilon}_{22} = 0.1\%$

Figures 16 and 17 show the differences between the σ_{22} and σ_{23} stress components produced by the four-noded and eight-noded finite-element models and the target model simulations for the four microstructural realizations. As demonstrated earlier, the DHN predictions show a good level of accordance with the finite-element benchmark solutions. However, the stress distributions of the former approach are smoother than the latter ones.

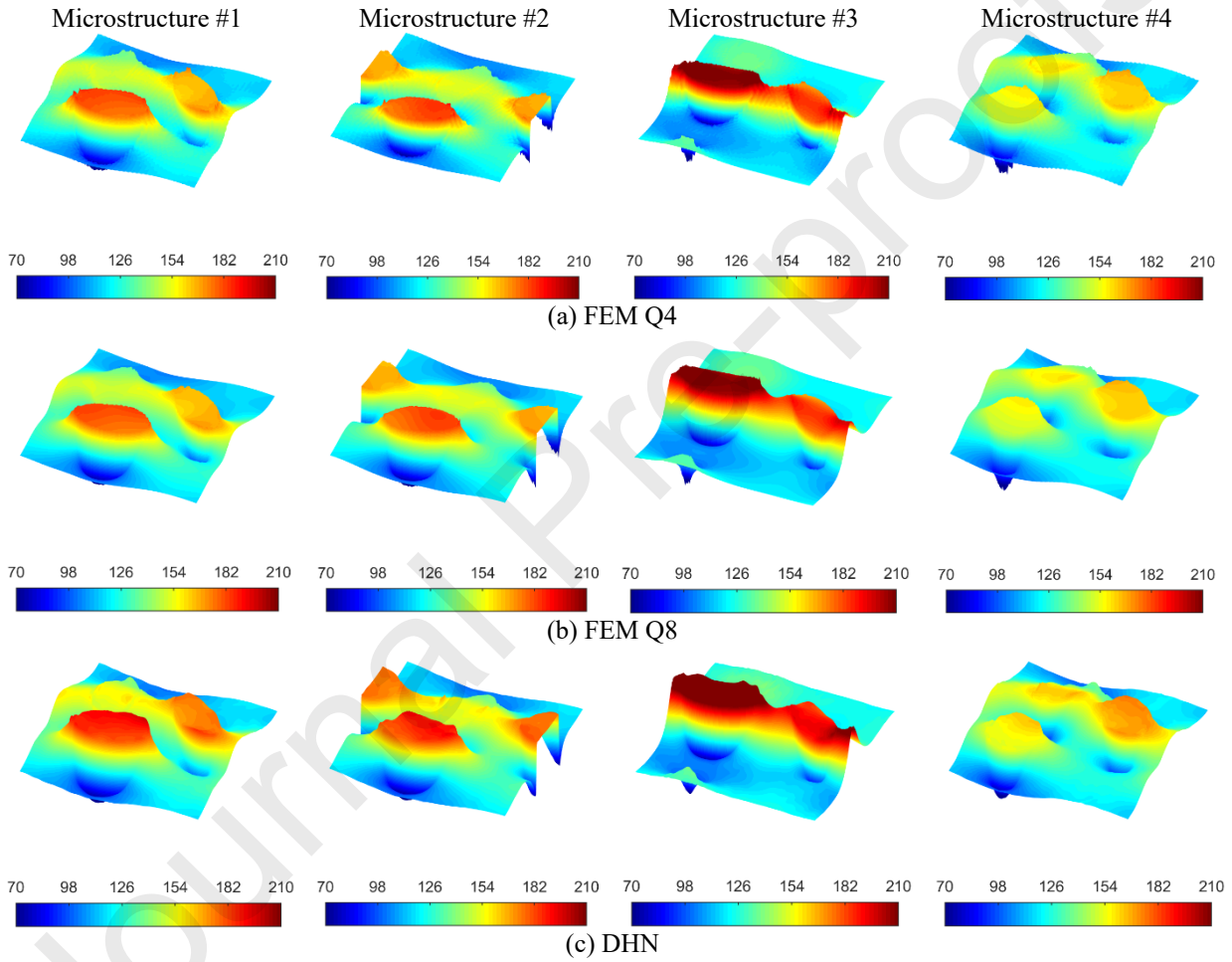
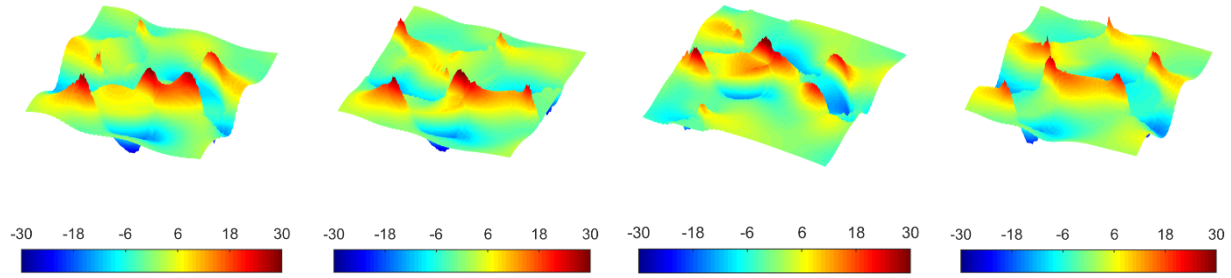
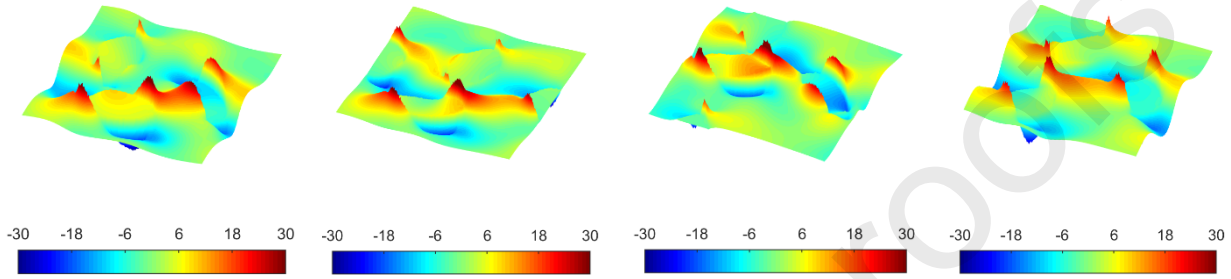


Figure 16 Comparison of σ_{22} [MPa] stress predicted by the target model and Q4 and Q8 FEM simulations the imposition of only one nonzero macroscopic strain component $\bar{\varepsilon}_{22} = 0.1\%$ with other strain components kept as zero

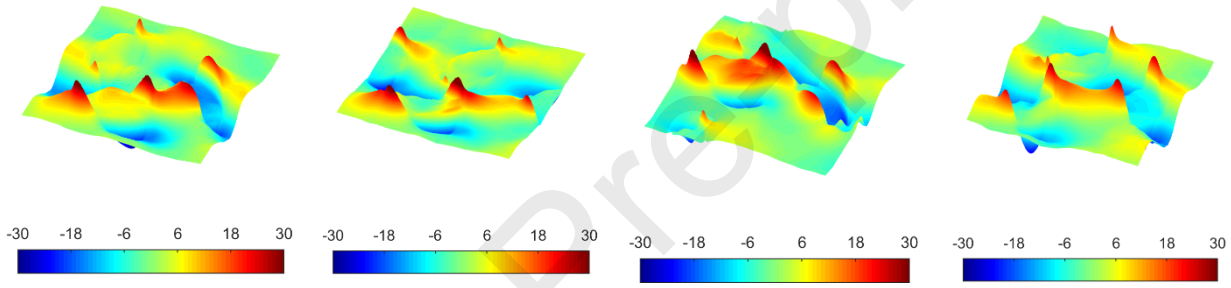
Microstructure #1 Microstructure #2 Microstructure #3 Microstructure #4



(a) FEM Q4



(b) FEM Q8



(a) DHN

Figure 17 Comparison of σ_{23} [MPa] stress predicted by the target model and Q4 and Q8 FEM simulations with the imposition of only one nonzero macroscopic strain component $\bar{\varepsilon}_{22} = 0.1\%$ with other strain components kept as zero

6. Three-dimensional composites

In this section, we present numerical experiments to further demonstrate the DHN's performance on composites reinforced by ellipsoidal inclusions (boron/aluminum system) or weakened by porosity (pore/aluminum) with periodicity in the space vis-à-vis the finite-element benchmark solutions. We limit the comparison to the single ellipsoidal inclusion/porosity since the generation of the three-dimensional periodic mesh for random composites is particularly cumbersome and computationally expensive in the finite-element method. In the case of network simulation, the

distributions of material properties are expressed in terms of Lamé functions $\lambda(\mathbf{y})$ and $\mu(\mathbf{y})$ in the form:

$$\lambda(\mathbf{y}) = \lambda_1 \left[\lambda_2 + \tanh \frac{1 - \left(\frac{y_1}{b_1}\right)^2 - \left(\frac{y_2}{b_2}\right)^2 - \left(\frac{y_3}{b_3}\right)^2}{\delta} \right] + \lambda_3 \tag{23}$$

$$\mu(\mathbf{y}) = \mu_1 \left[\mu_2 + \tanh \frac{1 - \left(\frac{y_1}{b_1}\right)^2 - \left(\frac{y_2}{b_2}\right)^2 - \left(\frac{y_3}{b_3}\right)^2}{\delta} \right] + \mu_3 \tag{24}$$

where (y_1, y_2, y_3) are the coordinates of the collocation points and b_1, b_2 and b_3 are the length of semi-axes of an ellipsoid in the y_1, y_2 and y_3 direction, respectively.

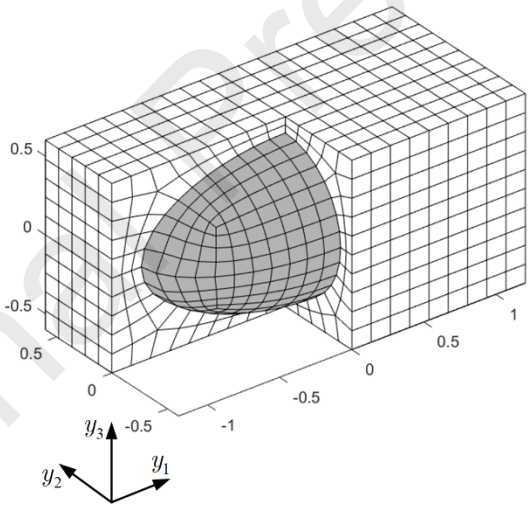
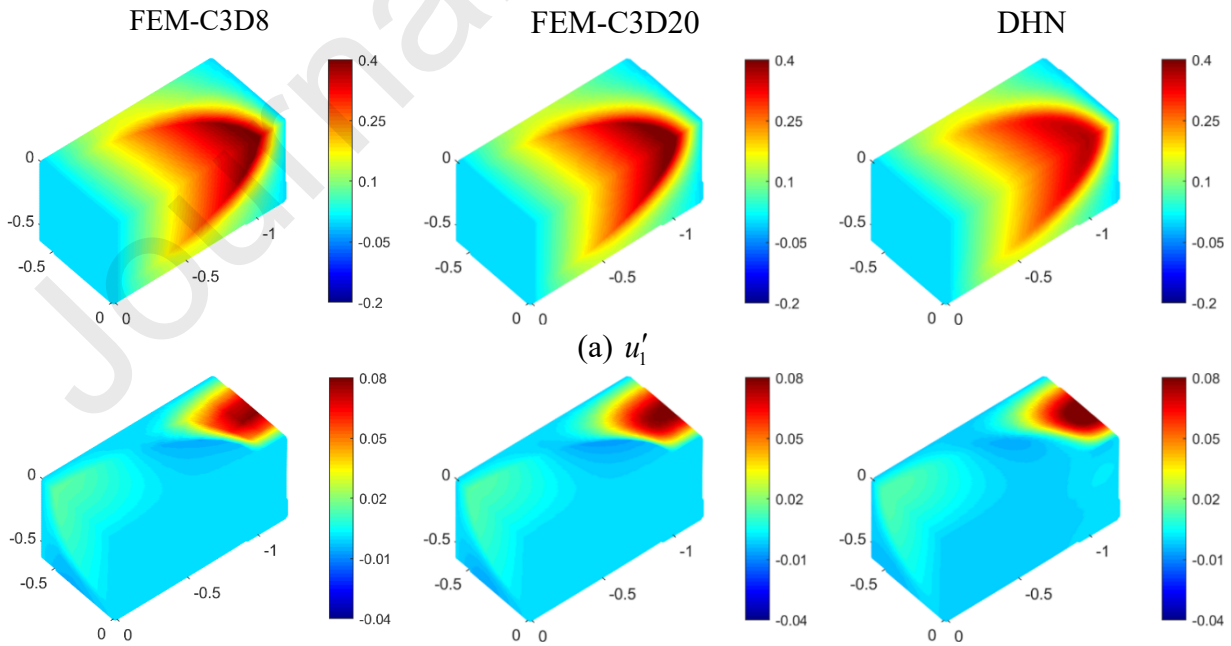


Figure 18 Finite element mesh refinement of the unit cell with ellipsoidal inclusion

Figure 18 shows the finite-element mesh refinement for composites reinforced with an ellipsoidal inclusion, which is discretized into 2952 C3D8 or C3D20 brick elements (ABAQUS notations). The volume fraction of the inclusion over the unit cell is $V_p = 0.3$ and the length of the semi-axes is $b_1 = 2b_2 = 2b_3 = 1mm$. We randomly generate 20k collocation points in the space (not

shown) for evaluating the PDE residuals. During the training of the neural network model, the learning rate is prescribed as 0.01 and decays by a factor of 0.5 per 3k epochs.

Figure 19 presents comparison of the u_1', u_2' and u_3' fluctuating displacement field distributions at one-eighth of the unit cell predicted by the C3D8 and C3D20 finite-element and neural network theories trained with 20k collocation points after 30k training epochs. Overall, the DHN-predicted fluctuating displacements agree well with the finite-element results and the differences are within the acceptance range. Figure 20 summarizes the comparison of concomitant σ_{11}, σ_{22} and σ_{23} stresses predicted by the network and finite-element approaches. As expected, the DHN and C3D20 finite elements predict similar local stress distributions, providing additional evidence for the developed approach in the three-dimensional domain. However, the C3D8 finite elements capture the essential characteristics of the stress distributions but the stresses in the soft matrix exhibit remarkable discontinuities despite that the associated displacements are smoothly varying.



(a) u_1'

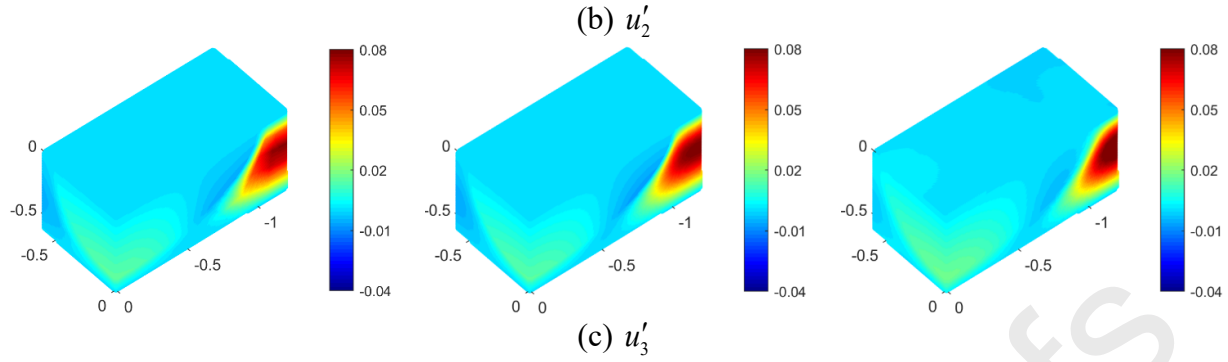


Figure 19 Comparison of displacement field distributions ($10^{-3}\mu m$) predicted by FEM and DHM with the imposition of only one macroscopic strain of 0.1%.

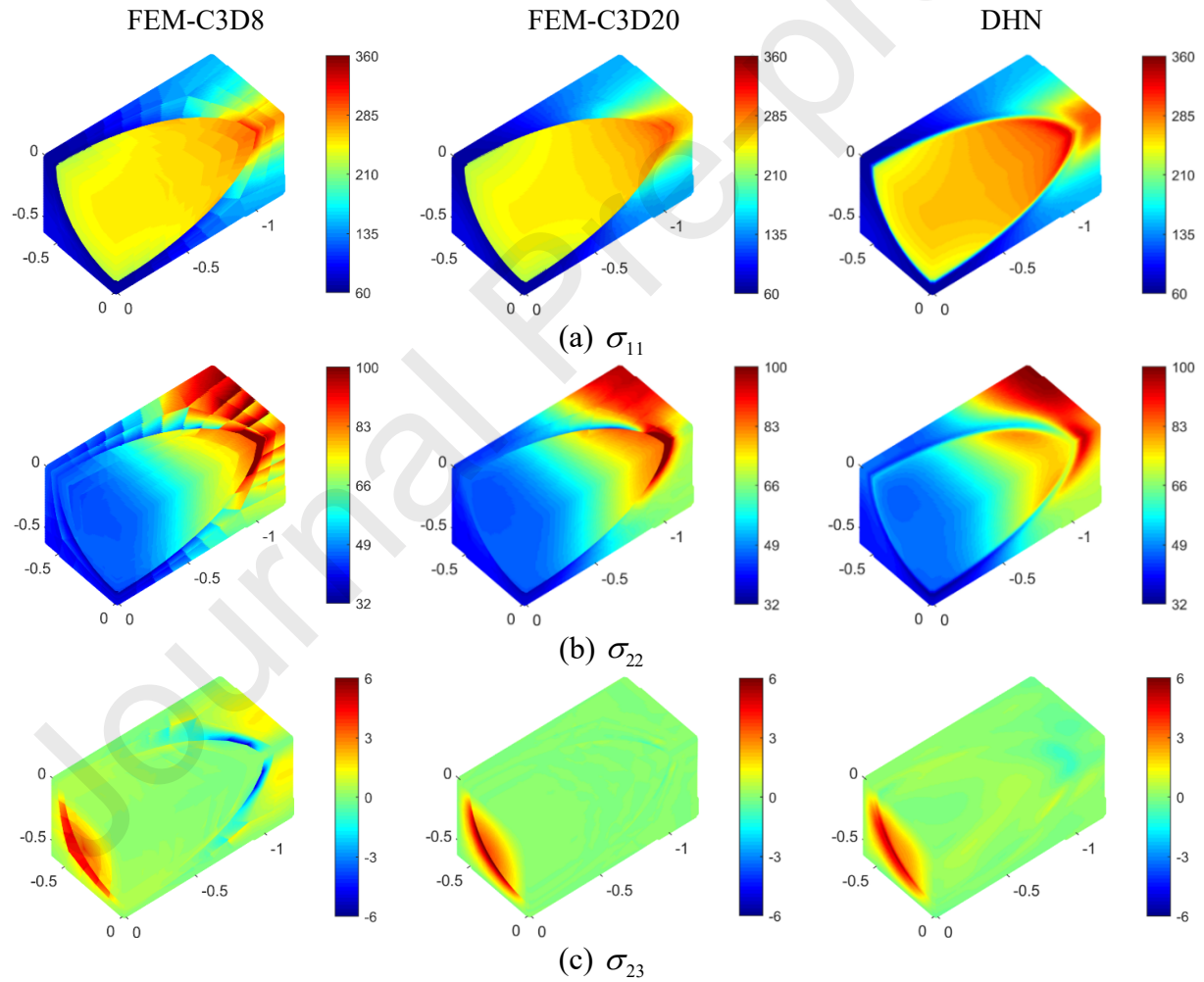


Figure 20 Comparison of local field distributions (MPa) predicted by FEM and DHM with the imposition of only one macroscopic strain of $\bar{\varepsilon}_{11} = 0.1\%$.

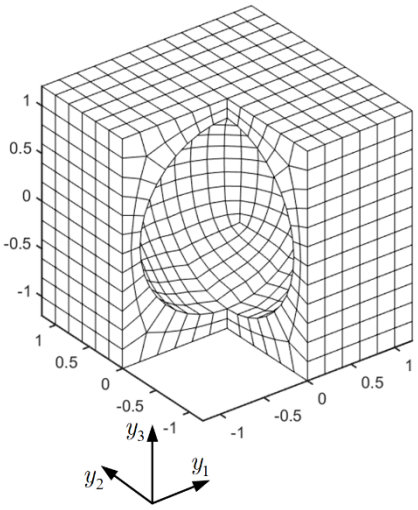
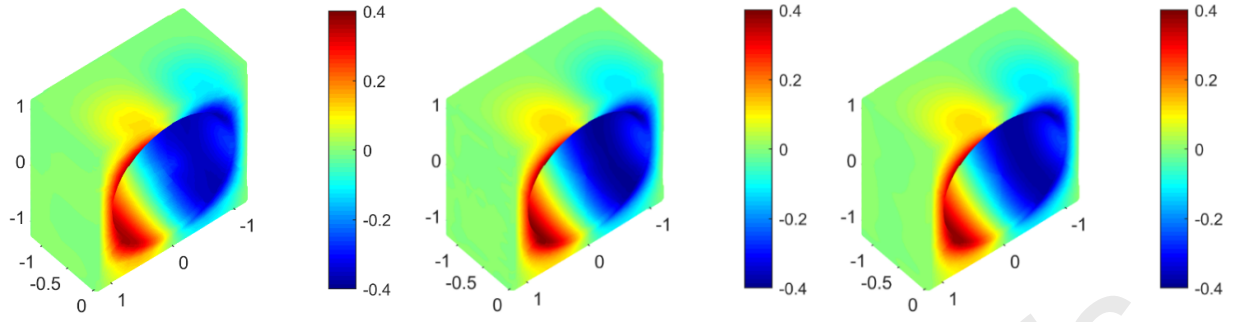


Figure 21 Finite element mesh refinement of the unit cell with spherical porosity

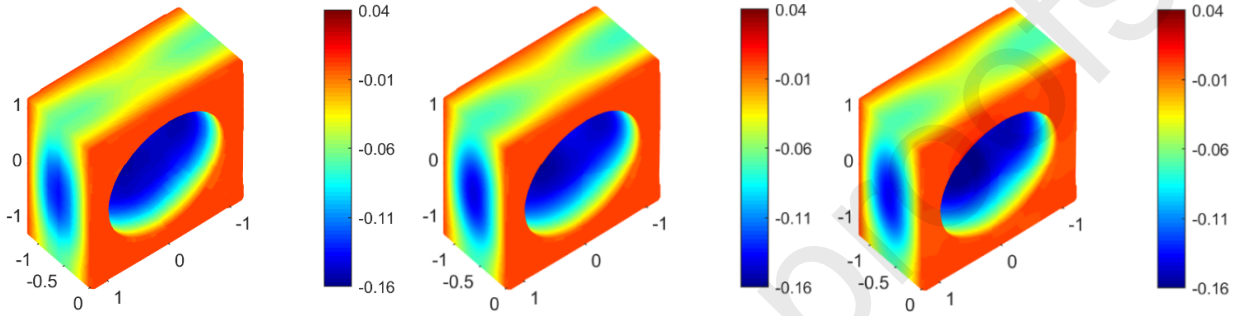
Similar simulations can be easily conducted for porous microstructured materials by removing the hard inclusion phase from the unit cell microstructure. For this, we consider a spherical porosity with a radius of $1mm$ embedded in a cubic unit cell. The porosity volume fraction is 30%. Figure 21 shows the finite-element mesh refinement for the unit cell that is discretized into 1382 C3D8 and C3D20 solid elements. Correspondingly, the neural network approach employs 14k collocation points to evaluate the PDE residuals.

Figure 22 presents comparison of the u_1', u_2' and u_3' fluctuating displacement predicted by the DHN trained with 14k collocation points after 20k training epochs and finite-element benchmark predictions. The differences between σ_{11}, σ_{22} and σ_{23} distributions predicted by the two theories are shown in Figure 23. It is remarkable how well the neural network approach is capable of mimicking the finite-element predictions of displacement and stress distributions in bulk materials.

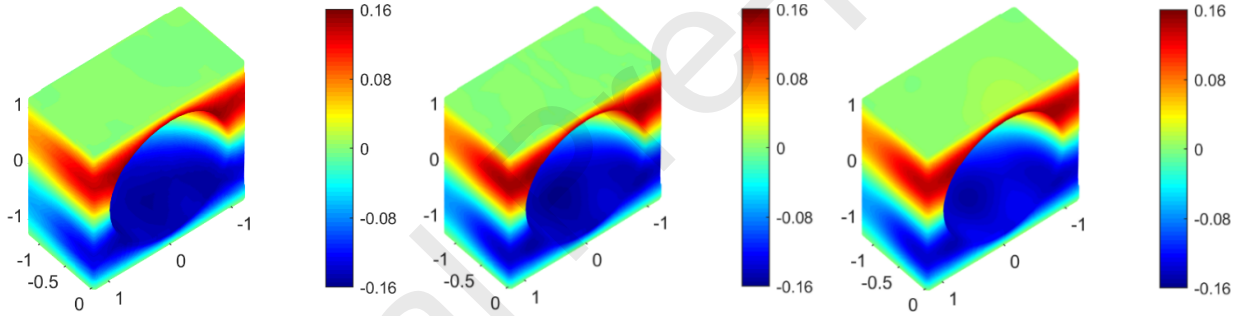
FEM-C3D8 FEM-C3D20 DHN



(a) u_1'



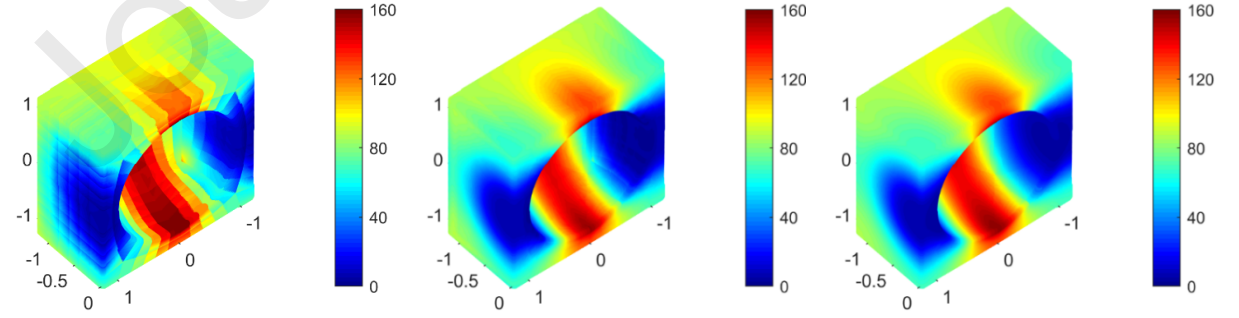
(b) u_2'



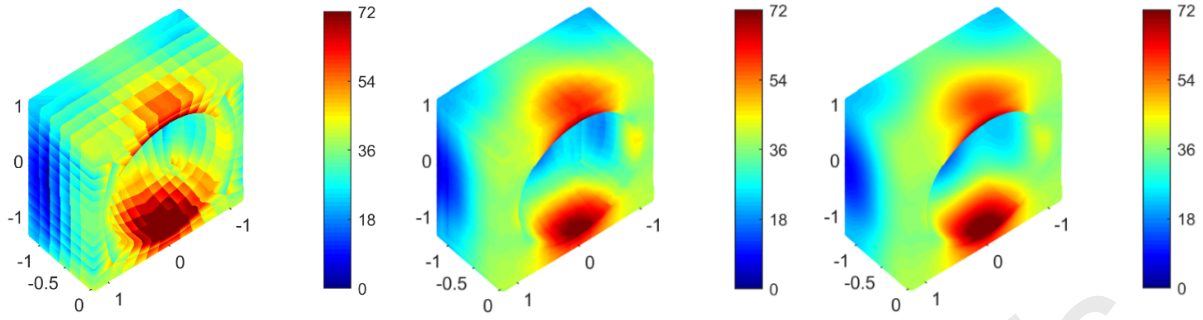
(c) u_3'

Figure 22 Comparison of displacement field distributions ($10^{-3}\mu m$) predicted by FEM and DHM with the imposition of only one macroscopic strain of $\bar{\varepsilon}_{11} = 0.1\%$.

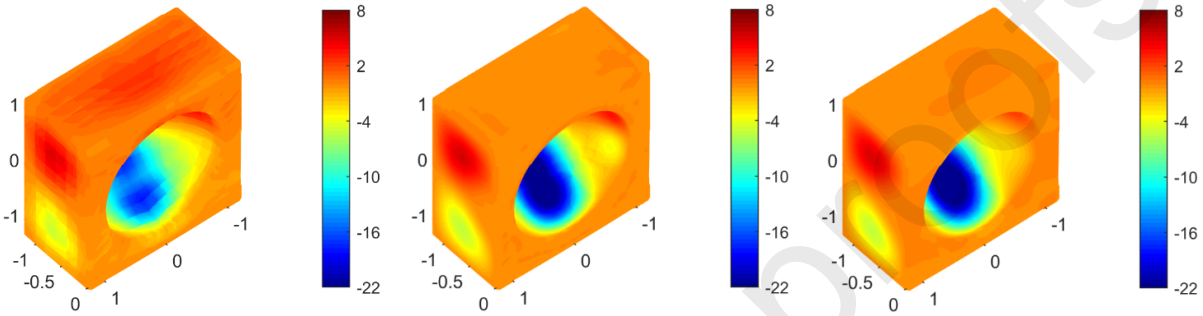
FEM-C3D8 FEM-C3D20 DHN



(a) σ_{11}



(b) σ_{22}



(c) σ_{23}

Figure 23 Comparison of local field distributions (MPa) predicted by FEM and DHM with the imposition of only one macroscopic strain of $\bar{\varepsilon}_{11} = 0.1\%$.

7 Discussion

The present contribution provides a general physically informed deep neural network framework for predicting the local displacement and stress fields, as well as homogenized moduli for elastic composites. The applicability of the DHN framework has been demonstrated in the context of both two- and three-dimensional microstructures. Thus, the DHN model has its place in the area of micromechanics of heterogeneous materials and holds promise for further development, such as incorporating plasticity and damage mechanisms. The advantages of the proposed techniques over the conventional finite-element techniques are several-fold. First of all, the periodicity boundary conditions of both displacements and tractions are automatically and exactly satisfied along the unit cell edges, while only the displacement periodicity boundary condition is fulfilled in the conventional finite-element method. Secondly, the network-based micromechanics theory is a mesh-free technique. The preparation of the input data (namely, the collocation points) for

complicated microstructures is much simpler than the finite-element method whose solution requires extensive mesh refinements. Thirdly, both the network-predicted stress and displacement fields are continuous and vary smoothly within the fiber/matrix domains without the discontinuities observed in the conventional finite-element technique.

While it offers certain benefits, the DHN framework also comes with inherent limitations. The infinitely differentiable nature of the neural network indicates that the predicted displacements fulfill C^{∞} continuity. However, in heterogeneous materials, the continuity of derivatives of displacements (namely the strains) does not hold at the fiber/matrix interface. Therefore, to obtain a desirable neural network solution, the DHN necessitates the incorporation of an empirical smoothing parameter in the vicinity of the fiber/matrix interface to allow a gradual transition of material stiffness. Secondly, the neural network approaches encounter significant challenges involving multiple random inclusions where the displacements exhibit sharp spatial transitions within the composite microstructures. While the transfer learning strategy is capable of yielding a better neural network solution, it requires an additional pre-training process. Thirdly, the DHN transforms a linear unit cell problem into a nonlinear optimization problem. Hence the computational efficiency of the DHN cannot compete with the conventional numerical methods. Nevertheless, the advances in computational power with parallel computing have truly unlocked the potential for employing neural networks to tackle these challenges within a feasible timeframe (Jiang et al., 2022).

8 Conclusions

A novel deep learning framework has been proposed for continuum micromechanical analysis of elastic composites with two- or three-dimensional periodicity by seamlessly combing the elements of zeroth-order homogenization and physics-informed neural networks. To this end, we adopt a

two-scale expansion of the displacement field in terms of macroscopic and microscopic contributions dependent on the global and the local coordinate systems, respectively. The latter (microscopic displacement) is obtained from neural networks by minimizing the residual of Navier's displacement equations of anisotropic microstructured materials for specified macroscopic strains with the help of automatic differentiation over a set of randomly generated collocation points. The successful application of the proposed technique relies on the simultaneous and exact satisfaction of the periodicity boundary conditions of displacements and tractions, required by the unit cell solutions, through the novel use of a periodic layer involving a set of sinusoidal functions and nonlinear activation functions. This circumvents the need for incorporating additional penalty terms representing the boundary conditions. In addition, the transfer learning technique plays a crucial role in obtaining accurate neural network solutions when dealing with multiple random inclusions. The effectiveness of this technique is highlighted by the improved accuracy achieved, which, in most cases, is challenging to attain through direct network training alone. The predictive capabilities of the DHN theory were assessed extensively against the finite-element benchmark solutions reinforced by unidirectional fiber/ellipsoidal inclusions or weakened by spherical porosity and good agreement was obtained for all configurations.

References

Baydin, A.G., Pearlmutter, B.A., Radul, A.A., Siskind, J.M., 2018. Automatic differentiation in machine learning: a survey. Journal of Machine Learning Research 18, 1-43.

Bensoussan, A., Lions, J.-L., Papanicolaou, G., 1978. Asymptotic analysis for periodic structures. North Holland, Amsterdam,Netherlands.

Brodnik, N.R., Muir, C., Tulshibagwale, N., Rossin, J., Echlin, M.P., Hamel, C.M., Kramer, S.L.B., Pollock, T.M., Kiser, J.D., Smith, C., Daly, S.H., 2023. Perspective: Machine learning in experimental solid mechanics. Journal of the Mechanics and Physics of Solids 173, 105231.

Cavalcante, M.A.A., Khatam, H., Pindera, M.-J., 2011. Homogenization of elastic–plastic periodic materials by FVDAM and FEM approaches – An assessment. Composites Part B: Engineering 42, 1713-1730.

Chen, Q., Chatzigeorgiou, G., Meraghni, F., Javili, A., 2022. Homogenization of size-dependent multiphysics behavior of nanostructured piezoelectric composites with energetic surfaces. European Journal of Mechanics - A/Solids 96, 104731.

Chen, Q., Du, X., Wang, W., Chatzigeorgiou, G., Meraghni, F., Zhao, G., 2023. Isogeometric homogenization of viscoelastic polymer composites via correspondence principle. Composite Structures 323, 117475.

Chen, Q., Pindera, M.-J., 2020. Homogenization and localization of elastic-plastic nanoporous materials with Gurtin-Murdoch interfaces: An assessment of computational approaches. International Journal of Plasticity 124, 42-70.

Chen, Q., Wang, G., Pindera, M.-J., 2018. Homogenization and localization of nanoporous composites-A critical review and new developments. Composites Part B: Engineering 155, 329-368.

Christensen, R., Lo, K., 1979. Solutions for effective shear properties in three phase sphere and cylinder models. Journal of the Mechanics and Physics of Solids 27, 315-330.

Cruz-González, O.L., Rodríguez-Ramos, R., Otero, J.A., Ramírez-Torres, A., Penta, R., Lebon, F., 2020. On the effective behavior of viscoelastic composites in three dimensions. International Journal of Engineering Science 157, 103377.

Cuomo, S., Di Cola, V.S., Giampaolo, F., Rozza, G., Raissi, M., Piccialli, F., 2022. Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next. Journal of Scientific Computing 92, 88.

Dey, A.P., Welschinger, F., Schneider, M., Gajek, S., Böhlke, T., 2022. Rapid inverse calibration of a multiscale model for the viscoplastic and creep behavior of short fiber-reinforced thermoplastics based on Deep Material Networks. International Journal of Plasticity, 103484.

Dong, S., Ni, N., 2021. A method for representing periodic functions and enforcing exactly periodic boundary conditions with deep neural networks. Journal of Computational Physics 435, 110242.

Faroughi, S.A., Pawar, N., Fernandes, C., Das, S., Kalantari, N.K., Mahjour, S.K., 2022. Physics-Guided, Physics-Informed, and Physics-Encoded Neural Networks in Scientific Computing. arXiv preprint arXiv:2211.07377.

Firooz, S., Steinmann, P., Javili, A., 2021. Homogenization of Composites With Extended General Interfaces: Comprehensive Review and Unified Modeling. Applied Mechanics Reviews 73.

Gajek, S., Schneider, M., Böhlke, T., 2020. On the micromechanics of deep material networks. Journal of the Mechanics and Physics of Solids 142, 103984.

Gajek, S., Schneider, M., Böhlke, T., 2021. An FE–DMN method for the multiscale analysis of short fiber reinforced plastic components. Computer Methods in Applied Mechanics and Engineering 384, 113952.

Gehrig, F., Wicht, D., Krause, M., Böhlke, T., 2022. FFT-based investigation of the shear stress distribution in face-centered cubic polycrystals. International Journal of Plasticity 157, 103369.

Gu, Y., Zhang, C., Zhang, P., Golub, M.V., Yu, B., 2023. Enriched physics-informed neural networks for 2D in-plane crack analysis: Theory and MATLAB code. International Journal of Solids and Structures 276, 112321.

Haghighat, E., Raissi, M., Moure, A., Gomez, H., Juanes, R., 2021. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Computer Methods in Applied Mechanics and Engineering 379, 113741.

Hashin, Z., Rosen, B.W., 1964. The Elastic Moduli of Fiber-Reinforced Materials. Journal of Applied Mechanics 31, 223-232.

He, Z., Liu, J., Chen, Q., 2023. Higher-order asymptotic homogenization for piezoelectric composites. International Journal of Solids and Structures 264, 112092.

He, Z., Pindera, M.-J., 2020. Locally exact asymptotic homogenization of periodic materials under anti-plane shear loading. European Journal of Mechanics - A/Solids 81, 103972.

He, Z., Pindera, M.-J., 2021. Locally exact asymptotic homogenization of viscoelastic composites under anti-plane shear loading. Mechanics of Materials 155, 103752.

Henkes, A., Wessels, H., Mahnken, R., 2022. Physics informed neural networks for continuum micromechanics. Computer Methods in Applied Mechanics and Engineering 393, 114790.

Hill, R., 1965. Theory of mechanical properties of fibre-strengthened materials—III. self-consistent model. Journal of the Mechanics and Physics of Solids 13, 189-198.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359-366.

Jiang, J., Wu, J., Chen, Q., Chatzigeorgiou, G., Meraghni, F., 2023. Physically informed deep homogenization neural network for unidirectional multiphase/multi-inclusion thermoconductive composites. Computer Methods in Applied Mechanics and Engineering 409, 115972.

Jiang, J., Zhao, J., Pang, S., Meraghni, F., Siadat, A., Chen, Q., 2022. Physics-informed deep neural network enabled discovery of size-dependent deformation mechanisms in nanostructures. International Journal of Solids and Structures 236-237, 111320.

Kalina, K.A., Linden, L., Brummund, J., Kästner, M., 2023. FE$${}^\textrm{ANN}$$: an efficient data-driven multiscale approach based on physics-constrained neural networks and automated data mining. Computational Mechanics 71, 827-851.

Liu, Z., Wu, C.T., Koishi, M., 2019. A deep material network for multiscale topology learning and accelerated nonlinear modeling of heterogeneous materials. Computer Methods in Applied Mechanics and Engineering 345, 1138-1168.

McClenny, L.D., Braga-Neto, U.M., 2023. Self-adaptive physics-informed neural networks. Journal of Computational Physics 474, 111722.

Mogilevskaya, S.G., Crouch, S.L., Stolarski, H.K., Benusiglio, A., 2010. Equivalent inhomogeneity method for evaluating the effective elastic properties of unidirectional multi-phase composites with surface/interface effects. International Journal of Solids and Structures 47, 407-418.

Mori, T., Tanaka, K., 1973. Average stress in matrix and average elastic energy of materials with misfitting inclusions. Acta metallurgica 21, 571-574.

Nguyen-Thanh, V.M., Anitescu, C., Alajlan, N., Rabczuk, T., Zhuang, X., 2021. Parametric deep energy approach for elasticity accounting for strain gradient effects. Computer Methods in Applied Mechanics and Engineering 386, 114096.

Niu, S., Zhang, E., Bazilevs, Y., Srivastava, V., 2023. Modeling finite-strain plasticity using physics-informed neural network and assessment of the network performance. Journal of the Mechanics and Physics of Solids 172, 105177.

Pindera, M.-J., Bansal, Y., 2007. On the Micromechanics-Based Simulation of Metal Matrix Composite Response. Journal of Engineering Materials and Technology 129, 468.

Pindera, M.-J., Khatam, H., Drago, A.S., Bansal, Y., 2009. Micromechanics of spatially uniform heterogeneous media: A critical review and emerging approaches. Composites Part B: Engineering 40, 349-378.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686-707.

Saeb, S., Steinmann, P., Javili, A., 2016. Aspects of Computational Homogenization at Finite Deformations: A Unifying Review From Reuss' to Voigt's Bound. Applied Mechanics Reviews 68.

Samaniego, E., Anitescu, C., Goswami, S., Nguyen-Thanh, V.M., Guo, H., Hamdia, K., Zhuang, X., Rabczuk, T., 2020. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. Computer Methods in Applied Mechanics and Engineering 362, 112790.

Suquet, P.M., 1987. Introduction, in: Sanchez-Palencia, E., Zaoui, A. (Eds.), Homogenization Techniques for Composite Media. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 193-198.

Vahab, M., Shahbodagh, B., Haghighat, E., Khalili, N., 2023. Application of Physics-Informed Neural Networks for forward and inverse analysis of pile–soil interaction. International Journal of Solids and Structures 277-278, 112319.

Weiss, K., Khoshgoftaar, T.M., Wang, D., 2016. A survey of transfer learning. Journal of Big data 3, 1-40.

Highlights

- •A physically informed deep homogenization network was developed for elastic heterogeneous composites
- The neural network outputs exactly satisfy both periodic displacement and traction boundary conditions
- The proposed framework is verified by comparison with finite-element predictions for both unidirectional composites and particulate-reinforced composites