Adaptive deep homogenization theory for periodic heterogeneous materials

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

Adaptive deep homogenization theory for periodic heterogeneous materials

Jiajun Wu [b], Qiang Chen [a,*], Jindong Jiang [c], 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

A B S T R A C T

We present an adaptive physics-informed deep homogenization neural network (DHN) approach to formulate a full-field micromechanics model for elastic and thermoelastic periodic arrays with different microstructures. The unit cell solution is approximated by fully connected multilayers via minimizing a loss function formulated in terms of the sum of residuals from the stress equilibrium and heat conduction partial differential equations (PDEs), together with interfacial traction-free or adiabatic boundary conditions. In comparison, periodicity boundary conditions are directly satisfied by introducing a network layer with sinusoidal functions. Fully trainable weights are applied on all collocation points, which are simultaneously trained alongside the network weights. Hence, the network automatically assigns higher weights to the collocation points in the vicinity of the interface (particularly challenging regions of the unit cell solution) in the loss function. This compels the neural networks to enhance their performance at these specific points. The accuracy of adaptive DHN is verified against the finite element and the elasticity solution respectively for elliptical and circular cylindrical pores/fibers. The advantage of the adaptive DHN over the original DHN technique is justified by considering locally irregular porous architecture where pore–pore interaction makes training the network particularly slow and hard to optimize.

1. Introduction

Engineered microstructured materials with man-made porosity/fiber are prevalent in a diverse array of technologically significant applications, including aerospace engineering, building construction, and the electronic industry, among others [1–5,31]. They are utilized to obtain better strength-to-weight ratios, acoustic suppression, and filtration. The porosity/fiber in heterogeneous materials produces local stress field concentrations which significantly affects the overall stiffness and strength [6]. Typically, these microstructured materials are fabricated with a uniform dispersion of diverse hole/fiber shapes and distribution patterns. These characteristics are defined by a repeating unit cell (RUC), serving as the fundamental building block for the entire array.

The porosity/fiber architecture (such as porosity/fiber volume fraction, shape, and arrangement) holds a crucial role in determining the overall behavior and local stress field distributions, and thus the failure mechanisms in this type of material. In order to design microstructured materials with desired properties, micromechanics and homogenization approaches have been extensively utilized for analyzing their response to different loading types. To predict the elastic properties of such material, we refer to the early contributions of Koiter and Langer [7], Grigolyuk et al. [8], and Mejiers [9], with the work of O'Donnell and coworkers [10] leading the way. More recent work in the literature utilizes the finite element, finite volume, or finite difference methods for solutions to the unit cell problem [2,11–13]. These numerical techniques are capable of handling more complicated pore/fiber shapes and arrangements, as well as inelastic mechanisms that are taking place in each phase. Exact solutions of microstructural arrays of different patterns and shapes, subjected to periodic traction and displacement boundary conditions, are scarce with the exception of the locally exact homogenization theory (LEHT) pioneered by Pindera and his coworkers [14,15].

Recently, physically informed deep neural networks (PINN) have become a compelling alternative to conventional numerical and analytical methods for solving partial differential equations (PDE) [16]. In contrast to the purely data-driven approaches that do not take into account the relevant physics, the PINN incorporates explicitly the physics law described by a set of PDEs as a part of the neural network loss function over a set of collocation points, such that the network

* Corresponding author.
E-mail address: qiangchen@xjtu.edu.cn (Q. Chen).

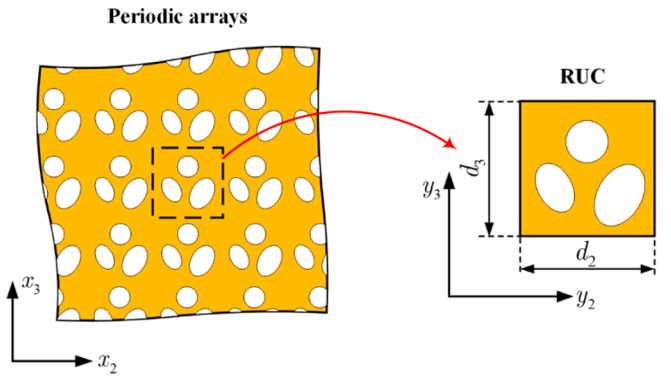

Fig. 1. A multiphase periodic porous array characterized by the smallest building block or the unit cell.

outputs are physics-law consistent. A range of elastic and inelastic boundary-value problems, in both the infinitesimal and finite-deformation domain, have been addressed by different variants of PINN [17–21]. The demonstrated advantage of PINN motivated us to develop a new micromechanics theory called deep homogenization neural nets for predicting the local temperature, displacement, temperature gradient, and heat flux distributions of thermoconductive and elastic unidirectional or particulate composites [22,23]. This method adopts a double-scale expansion of the temperature/displacement field into averaged and microscopic contributions. The incorporation of a periodic layer with sinusoidal functions ensures that the periodic boundary conditions of displacement/traction and temperature/heat flux are exactly fulfilled [24]. Nevertheless, when dealing with traction-free interfacial boundary conditions, the DHN must incorporate extra loss terms, posing a challenge in determining the appropriate loss weights. What's more, the loss landscape for the microstructural materials is quite complex and non-convex and is harder to optimize due to the microstructure-induced stress concentrations. In many scenarios, the DHN may face challenges in training or optimization, often necessitating a considerably large number of layers and neurons and training epochs to achieve the desired solutions.

Herein, we incorporate an adaptive weighting technique into our recently proposed DHN theory and extend the DHN to the micromechanical analysis of the elastic and conductive periodic heterogeneous arrays with different porosity/fiber architectures. The new theory applies fully trainable weights on each collocation point associated with the corresponding loss term [25]. These weights are simultaneously trained alongside the network weights. The significance of introducing trainable weights in the new DHN technique is that it automatically assigns higher weights to the collocation points in the vicinity of the interface (particularly challenging regions of the unit cell solution) in the loss function. This compels the neural networks to enhance their performance at these specific points. The accuracy of the new DHN technique is verified extensively against the finite element and the elasticity solution of Drago and Pindera [14]. The advantage of the adaptive DHN is justified by comparison with the original DHN technique for reproducing the local displacements and stress fields of the unit cell with locally irregular inclusion distributions.

2. Governing equations for homogenization of periodic arrays

This section aims to offer a comprehensive overview of homogenization theories applied to thermoconductive and elastic porous/fibrous composites. Additionally, it introduces the pertinent notations adopted in this work, laying the groundwork for the subsequent development of the physically informed machine learning theory in the following section.

For a heterogeneous solid containing periodically dispersed cylindrical porosities/fibers, the homogenized response of such heterogeneous media mirrors that of a Repeating Unit Cell (RUC) under periodic boundary conditions, as illustrated in Fig. 1. This representation can be duplicated in two dimensions to constitute the entirety of the heterogeneous material.

In the framework of asymptotic expansion homogenization [26–28], the temperature scalar variable T in the RUC of thermoelastic periodic microstructures can be partitioned into averaging and fluctuating contributions dependent on the global and local coordinates $\boldsymbol{x} = (x_1, x_2, x_3)$ and $\boldsymbol{y} = (y_1, y_2, y_3)$ as follows,

$$T(\boldsymbol{x}, \boldsymbol{y}) = -\overline{\boldsymbol{H}} \cdot \boldsymbol{x} + T'(\boldsymbol{y}) \tag{1}$$

where $\overline{\boldsymbol{H}} = [H_1, H_2, H_3]$ indicates the homogenized (or applied) temperature gradient and $T'(\boldsymbol{y})$ is the fluctuating temperature scalar variable represented as a periodic function. The local temperature gradients are related to the temperature as follows:

$$\boldsymbol{H} = \overline{\boldsymbol{H}} - \nabla T' \tag{2}$$

where ∇ indicates the gradient operator. The heat flux $\boldsymbol{\Phi} = [\Phi_1, \Phi_2, \Phi_3]^{\mathrm{T}}$ is expressed in terms of temperature gradient $\boldsymbol{H}$ via Fourier's law:

$$\boldsymbol{\Phi} = -\boldsymbol{k} \cdot \boldsymbol{H} \tag{3}$$

where $\boldsymbol{k}$ denotes the second-order thermal conductivity tensor. The fluctuating temperature $T'(\boldsymbol{y})$ is obtained by satisfying the periodicity boundary conditions, and interfacial adiabatic condition together with the steady-state heat conduction partial differential equation. In the case of infinitely-long cylindrical porosity/fiber, noting that none of the temperature gradient and heat flux depends on the out-of-plane coordinate y_1, the local steady-state heat conduction partial differential equation (PDE) expressed in terms of fluctuating temperature reads:

$$k_{22}\frac{\partial^2 T'}{\partial y_2^2} + k_{33}\frac{\partial^2 T'}{\partial y_3^2} = 0, \qquad \forall (y_2, y_3) \in \Omega \tag{4}$$

subjected to:

$$\begin{aligned} &T(\boldsymbol{x}_{\mathrm{o}} + \boldsymbol{d}) = T(\boldsymbol{x}_{\mathrm{o}}) + \overline{H}_i d_i, \\ &\Phi_n(\boldsymbol{x}_{\mathrm{o}} + \boldsymbol{d}) + \Phi_n(\boldsymbol{x}_{\mathrm{o}}) = 0, \qquad (\boldsymbol{x}_{\mathrm{o}}, \boldsymbol{x}_{\mathrm{o}} + \boldsymbol{d}) \in S \end{aligned} \tag{5}$$

$$\Phi_n(\boldsymbol{y}_{\mathrm{o}}) = 0, \quad \boldsymbol{y}_{\mathrm{o}} \in S_r \tag{6}$$

where $i = 1, 2, 3$. Ω is the unit cell bulk domain, S is the unit cell boundary, and S_r denotes the pore boundary. $\boldsymbol{d} = (d_2, d_3)$ denotes a characteristic distance that defines the periods of the repeating unit cell in relevant directions. $\Phi_n = \Phi_2 n_2 + \Phi_3 n_3$ indicates the heat flux passing through an inclined surface and $[n_2, n_3]$ denotes the outer normal vector of the surface.

In the same spirit, the displacement field $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}) = [u_1, u_2, u_3]^{\mathrm{T}}$ in the RUC of elastic periodic arrays can be written as:

$$\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{y}) = \overline{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x} + \boldsymbol{u}'(\boldsymbol{y}) \tag{7}$$

where $\overline{\boldsymbol{\varepsilon}}$ denotes the macroscopic strains and $\boldsymbol{u}'(\boldsymbol{y}) = [u_1', u_2', u_3']^{\mathrm{T}}$ indicates the fluctuating displacements induced by the heterogeneities. The strains are expressed in terms of displacements through the strain–displacement relation:

$$\boldsymbol{\varepsilon} = \overline{\boldsymbol{\varepsilon}} + \frac{1}{2}\left(\nabla \boldsymbol{u}' + \nabla \boldsymbol{u}'^{\mathrm{T}}\right) \tag{8}$$

and the stresses are expressed in terms of the strains through Hooke's law:

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon} \tag{9}$$

where $\boldsymbol{C}$ represents the fourth-order elasticity tensor.

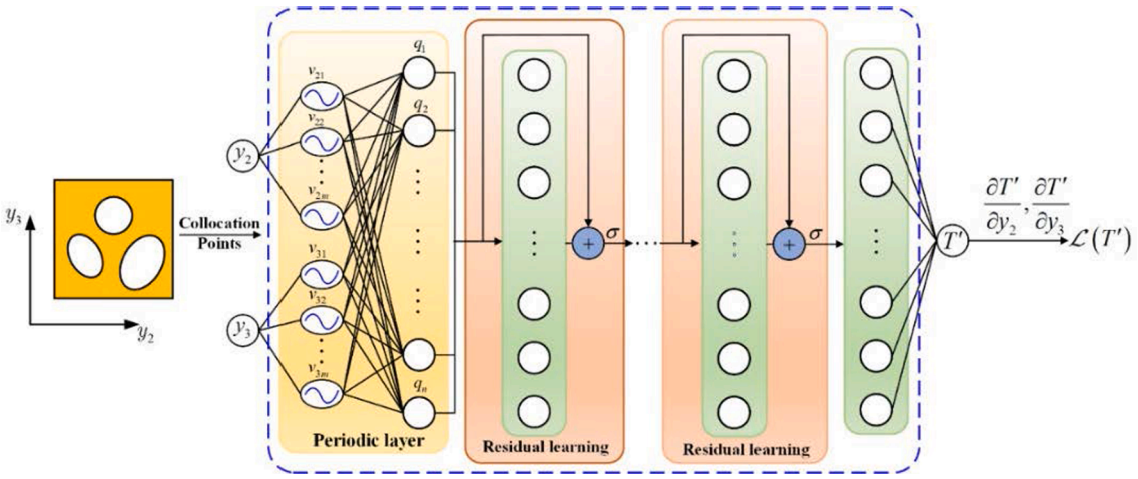

Fig. 2. Physically informed deep residual neural network acting as a surrogate model for homogenization of periodic thermoelastic microstructured materials.

For infinitely long cylindrical porosities/fibers, enforcing the periodicity boundary condition and the interfacial traction-free condition (in the case of porous microstructures) together with the local equilibrium equations yields the solution for fluctuating displacements $\boldsymbol{u}'(\boldsymbol{y})$ under specified macroscopic strains:

$$\begin{aligned} &C_{66}\frac{\partial^2 u_1'}{\partial y_2^2}+C_{55}\frac{\partial^2 u_1'}{\partial y_3^2}=0,\\ &C_{22}\frac{\partial^2 u_2'}{\partial y_2^2}+C_{44}\frac{\partial^2 u_2'}{\partial y_3^2}+(C_{23}+C_{44})\frac{\partial^2 u_3'}{\partial y_2 \partial y_3}=0, \qquad \forall (y_2,y_3)\in\Omega\\ &(C_{23}+C_{44})\frac{\partial^2 u_2'}{\partial y_2^2}+C_{44}\frac{\partial^2 u_3'}{\partial y_3^2}+C_{33}\frac{\partial^2 u_3'}{\partial y_3^2}=0 \end{aligned} \tag{10}$$

subjected to:

$$\begin{aligned} &u_i(\boldsymbol{x}_o+\boldsymbol{d})=u_i(\boldsymbol{x}_o)+\overline{\varepsilon}_{ij}d_j,\\ &t_i(\boldsymbol{x}_o+\boldsymbol{d})+t_i(\boldsymbol{x}_o)=0, \quad (\boldsymbol{x}_o,\boldsymbol{x}_o+\boldsymbol{d})\in S \end{aligned} \tag{11}$$

$$t_i(\boldsymbol{y}_o)=0, \quad \boldsymbol{y}_o\in S_r \tag{12}$$

where $i=1,2,3$. t_i denotes the traction component from the Cauchy relation: $t_i=\sigma_{ij}n_j$.

It should be noted that the formulation of the axial shear problem shows that it is mathematically completely analogous to the transverse conductivity problem where the displacement u_1' corresponds to the temperature T', axial shear moduli C_{66}, C_{55} correspond to transverse conductivities k_{22}, k_{33} and axial shear stresses σ_{12}, σ_{13} correspond to the transverse heat flux Φ_2, Φ_3. This analogy is well-known for isotropic and transversely isotropic phases and is employed in this manuscript.

3. Adaptive deep homogenization neural network theory

Analytical solutions to the unit cell boundary value problems described by Eq. (4) and Eq. (10), with periodic constraints, are scarce. In this section, instead of solving these equations analytically or numerically, the deep neural networks are harnessed as universal function approximators to predict solutions to the governing differential equations of the unit cells with prescribed periodicity boundary constraints and interfacial conditions. Following the framework of DHN proposed by the present authors, the fluctuating temperature $T'(\boldsymbol{y})$ is approximated by fully connected residual network layers, which take the spatial coordinates $\boldsymbol{y}$ and output $T'(\boldsymbol{y})$ as follows:

$$[T'] = NN[y_2, y_3] \tag{13}$$

namely, a mapping from $\mathbb{R}^2 \to \mathbb{R}^1$, and in the case of an elastic heterogeneous solid, the network models can be represented by:

$$\begin{aligned} [u_1'] &= NN[y_2, y_3],\\ [u_2', u_3'] &= NN[y_2, y_3] \end{aligned} \tag{14}$$

namely, mappings from $\mathbb{R}^2 \to \mathbb{R}^1$ and $\mathbb{R}^2 \to \mathbb{R}^2$, respectively. Deep neural networks are comprised of multiple hidden layers in which the inputs of each layer and outputs are propagated through the network as:

$$q_j^l = \sigma\left(w_{i,j}^l q_j^{l-1} + b_j^l + q_j^{l-1}\right) \tag{15}$$

where q_j^{l-1} and q_j^l denote the inputs and output of the lth layer. $w_{i,j}^l$ and b_j^l are the weights and biases of the lth layer. σ represents the nonlinear activation function. The network parameters $\boldsymbol{\theta}$ (including the weights $\boldsymbol{w}$ and biases $\boldsymbol{b}$) are trained using gradient descent methods based on the backpropagation method by minimizing a loss function that penalizes the residuals for Eqs. (4)-(6) and Eqs. (10)-(12) as follows:

$$\mathscr{L} \ominus (\boldsymbol{\theta}) = \lambda_1 \mathscr{L}_{\text{PDE}}(\boldsymbol{\theta}) + \lambda_2 \mathscr{L}_{\text{per}}(\boldsymbol{\theta}) + \lambda_3 \mathscr{L}_{\text{int}}(\boldsymbol{\theta}) \tag{16}$$

where λ_i's are the weights associated with each loss term. $\mathscr{L}_{\text{PDE}}(\boldsymbol{\theta})$, $\mathscr{L}_{\text{per}}(\boldsymbol{\theta})$ and $\mathscr{L}_{\text{int}}(\boldsymbol{\theta})$ denote the PDE residuals, periodicity residuals, and interface heat flux or traction residuals, respectively, which should all vanish for a desired unit cell solution.

In the presence of multiple loss terms, the weight of each loss term in Eq. (16) must be carefully selected and fine-tuned to achieve a successful neural network solution. Otherwise, the network model can easily produce unreliable approximations if not fail to train altogether. We note that in heterogeneous porous materials, large deformation, and stress gradient/localization are typically observed in the vicinity of the pore boundary. The deep neural networks have difficulty reproducing such "stiff" PDE solutions and accurately capturing the localized effects in heterogeneous materials using the gradient-based optimization method. This presents an additional hurdle in the development of machine learning homogenization methods. To mitigate this issue, two techniques as described below are introduced in order to enforce exactly the periodicity boundary condition and adaptively determine the weights of each loss function term.

3.1. Implementation of periodic boundary condition

We employ the periodic function representation technique in our deep homogenization model to enforce the periodicity boundary conditions. This step is vital as the homogenization theory necessitates exact adherence to periodic boundary conditions. Therefore, the accuracy of the solution for the unit cell neural network critically depends on the

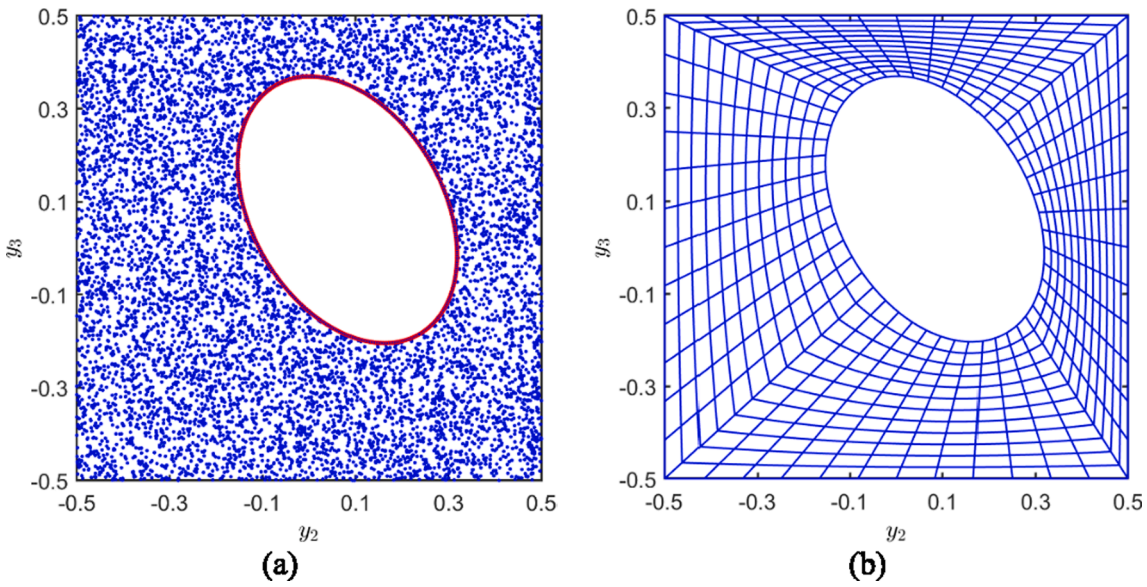

Fig. 3. (a) Training point distribution and (b) finite-element unit cell discretization.

incorporation of periodic boundary conditions.

As shown in Fig. 2, let us take the homogenization of thermo-conductive porous/fibrous media as an example. The fundamental concept involves expressing an arbitrary function represented by a network using a collection of independent sinusoidal functions. These sinusoidal functions have predetermined periods and adjustable training parameters, cf. Dong and Ni [24], so that the neural network predictions exactly satisfy the periodicity conditions of any order:

$$\begin{aligned} v_{2i}(y_2) &= \sigma[A_{2i}\cos(\omega_2 y_2 + \phi_2) + c_{2i}], \\ v_{3i}(y_3) &= \sigma[A_{3i}\cos(\omega_3 y_3 + \phi_3) + c_{3i}], \\ q_j(y_2, y_3) &= \sigma\left[\sum_{i=1}^{m} v_{2i}(y_2) W_{ij}^{(2)} + \sum_{i=1}^{m} v_{3i}(y_3) W_{ij}^{(3)} + B_j\right] \end{aligned} \tag{17}$$

with

$$\omega_2 = \frac{2\pi}{d_2}, \omega_3 = \frac{2\pi}{d_3} \tag{18}$$

In these equations, $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. $q_j(y_2, y_3)$ are the output of this layer. ω_2 and ω_3 are constants with fixed periods d_2 and d_3 that represent the characteristic lengths of the repeating unit cell, respectively. The inclusion of the activation function σ in the periodic layer guarantees that $v_{2i}(y_2)$ and $v_{3i}(y_3)$ encompass not only the frequencies ω_2 and ω_3, but also the higher frequencies with common periods in the pertinent direction. The training parameters of the periodic layer include:

$$A_{2i}, A_{3i}, \phi_{2i}, \phi_{3i}, c_{2i}, c_{3i}, W_{ij}^{(2)}, W_{ij}^{(3)}, B_j \tag{19}$$

The periodic layer defined in Eq. (17) enables the precise fulfillment of periodicity boundary conditions for the displacements/temperature field up to an infinite order (C^∞), hence the periodicity boundary conditions for tractions/heat flux are automatically enforced with precision at the machine level. As a result, the second term representing the periodicity boundary loss in Eq. (16) can be omitted directly from the total loss function by using this approach.

3.2. Adaptive weight determination

The DHN algorithm described in the previous subsection, with the unit cell periodicity boundary condition automatically satisfied, though provides a paradigm shift in approximating the unit cell solutions, can still produce inaccurate predictions in mimicking the response of heterogeneous media. This is because the competing effect between the PDE loss and the interfacial conditions, as well as the large deformation/stress gradient that exists in the vicinity of the phase interface, adversely affects the network performance. Indeed, the gradient descent back-propagation is a greedy procedure that may latch on some of the components at the expense of the others, therefore creating an imbalance in the rate of descent among the different loss components, if not on different collocation points, and preclude convergence to the accurate solution.

In this work, following the recent work by McClenny and Braga-Neto [25], instead of hard-coding weights in each loss term, we apply the fully trainable weights on each collocation point associated with the corresponding loss term, resembling the soft multiplicative attention masks commonly utilized in computer vision. These weights are simultaneously trained alongside the neural network weights. As a result, the PDE and interface collocation points in the vicinity of the pore boundary, which represents the difficult regions of the unit cell solution, are automatically assigned higher weights in the loss function. This compels the approximation to enhance its performance specifically in those critical areas. The loss function in Eq. (16) can be written as:

$$\mathscr{L} \ominus (\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{PDE}}, \boldsymbol{\lambda}_{\text{int}}) = \mathscr{L}_{\text{PDE}}(\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{PDE}}) + \mathscr{L}_{\text{int}}(\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{int}}) \tag{20}$$

where $\boldsymbol{\lambda}_{\text{PDE}} = \left(\lambda_p^1, \cdots, \lambda_p^{N_p}\right)$ and $\boldsymbol{\lambda}_{\text{int}} = \left(\lambda_i^1, \cdots, \lambda_i^{N_i}\right)$ are trainable, non-negative adaptive weights associated with the N_p PDE and N_i interface collocation points, respectively.

During the training of the new deep homogenization model, we attempt to minimize the total loss function $\mathscr{L} \ominus (\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{PDE}}, \boldsymbol{\lambda}_{\text{int}})$ with regard to the network parameters $\boldsymbol{\theta}$ based on the gradient descent, but concurrently maximize the total loss function with regard to the adaptive weights $\boldsymbol{\lambda}_{\text{PDE}}$ and $\boldsymbol{\lambda}_{\text{int}}$ based on the gradient ascent as follows:

$$\begin{aligned} \boldsymbol{\theta}^{k+1} &= \boldsymbol{\theta}^k - \eta_k \nabla_{\boldsymbol{\theta}} \mathscr{L} \ominus (\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{PDE}}, \boldsymbol{\lambda}_{\text{int}}), \\ \boldsymbol{\lambda}_{\text{PDE}}^{k+1} &= \boldsymbol{\lambda}_{\text{PDE}}^k + \rho_{\text{PDE}}^k \nabla_{\boldsymbol{\lambda}_{\text{PDE}}} \mathscr{L} \ominus (\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{PDE}}, \boldsymbol{\lambda}_{\text{int}}), \\ \boldsymbol{\lambda}_{\text{int}}^{k+1} &= \boldsymbol{\lambda}_{\text{int}}^k + \rho_{\text{int}}^k \nabla_{\boldsymbol{\lambda}_{\text{int}}} \mathscr{L} \ominus (\boldsymbol{\theta}, \boldsymbol{\lambda}_{\text{PDE}}, \boldsymbol{\lambda}_{\text{int}}) \end{aligned} \tag{21}$$

In the above equations, η_k denotes the non-negative learning rate for the network parameters at the kth step. $\rho_{(\bullet)}^k$ denotes a separate non-negative learning rate for the self-adaptation weights. It's important to highlight that the optimal hyperparameters for the neural network, including the number of hidden layers and neurons per layer, learning rate, and

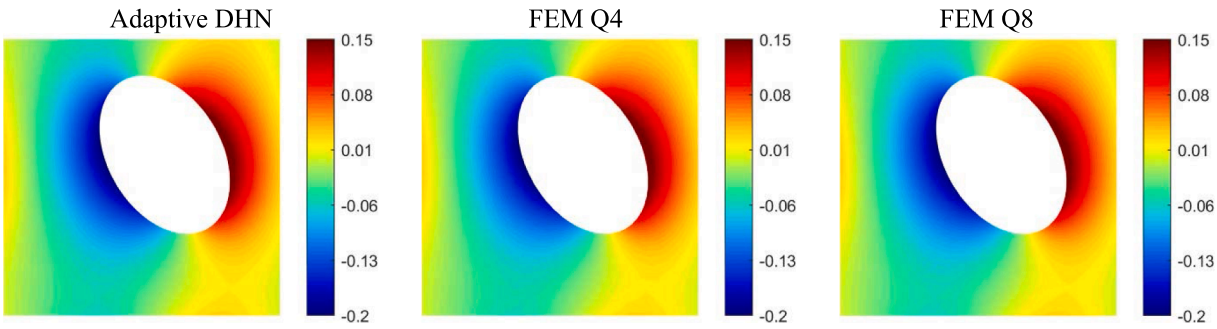

Fig. 4. Comparison of the fluctuating temperature $T^{'}(\mathrm{K})$ for periodic arrays with an elliptical cylindrical pore under macroscopic temperature gradient $\overline{H}_2 = 1\mathrm{K/m}$.

activation function, may vary depending on the specific case, as determined by our previous experience. In this manuscript, the hyperbolic tangent function $\sigma(x) = \tanh(x)$ for the periodic layer and all the hidden layers is utilized as a nonlinear activation function unless otherwise stated.

4. Numerical results

In order to demonstrate the modeling and predictive capabilities of the proposed methodology, several numerical experiments are conducted using the proposed theory. The neural network results are validated extensively against in-house finite-element-based homogenization theory and the elasticity-based homogenization predictions available in the literature.

4.1. Conductive porous media

The first investigated porous architecture is thermoconductive media containing 20 % pore volume fraction. The unit cell representative of such pore media is characterized by a square unit cell with an off-centered cylindrical porosity with an elliptical cross-sectional shape. The aspect ratio of the elliptical pore is 1.5 whose major axis is rotated by 120 degrees with respect to the horizontal axis. The lack of plane symmetry in the considered pore architectures significantly alters the local field distributions relative to the symmetric geometry and provides a very demanding test of the proposed network homogenization since the effect of the periodicity boundary condition is more important. The thermal conductivities of the bulk materials are prescribed as: $k_{22} = k_{33} = 10\ \mathrm{Wm^{-1}K^{-1}}$. A macroscopic transverse temperature gradient by $\overline{H}_2 = 1\mathrm{K/m}$ is applied with other components of temperature gradients set to zeros.

We employ a fully connected DHN model, as shown in Fig. 2, with 1 periodic layer, 3 hidden layers, and 30 neurons per layer. The DHN model is trained by Adam optimizer with a decreasing learning rate schedule to decrease the loss oscillation. The number of training points is 8k points in the bulk materials and 360 points on the interface, respectively, which were generated randomly using the Monte Carlo simulations, see Fig. 3(a). The unit cell in finite element simulation was discretized into 48 × 12 elements in circumferential and radial directions, respectively, see Fig. 3(b). The trained neural network model will be evaluated on a separate and unseen dataset containing 36,864 integration points in the finite-element unit cell model.

To demonstrate the performance of the proposed DHN technique on inferring local field distributions on unseen data, the results obtained

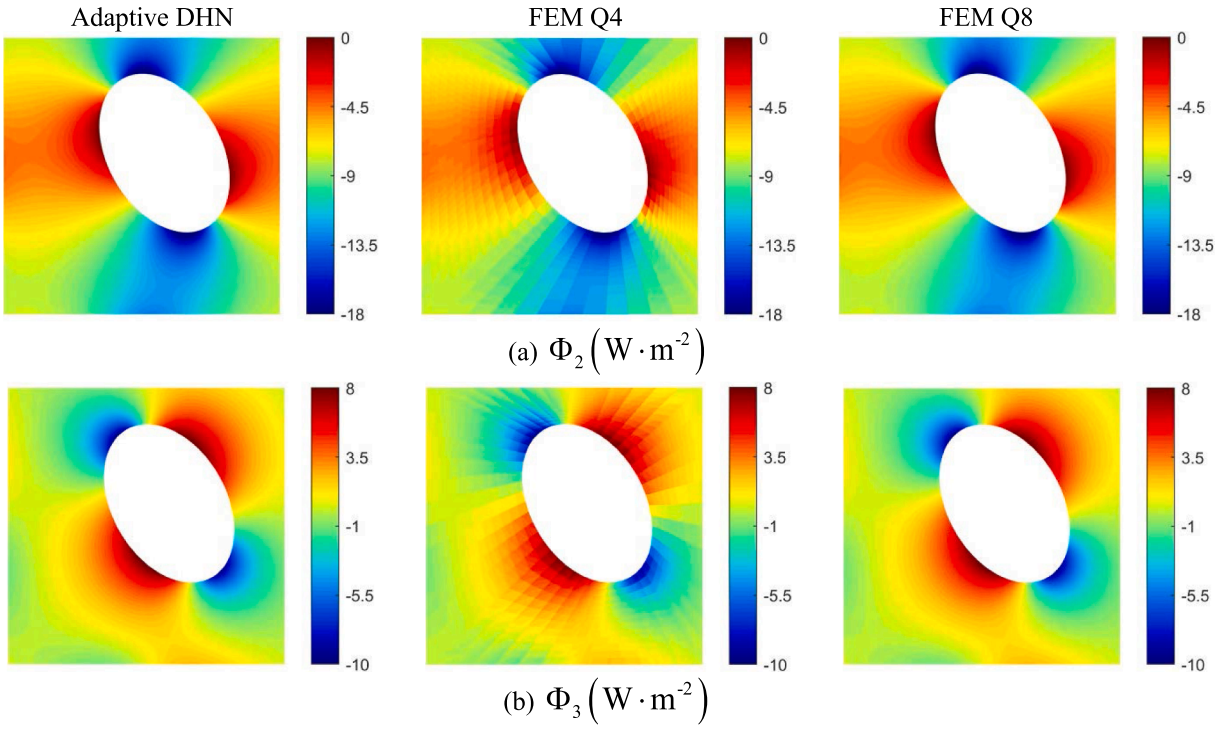

Fig. 5. Comparison of transverse heat flux for periodic arrays with an elliptical cylindrical pore under macroscopic temperature gradient $\overline{H}_2 = 1\mathrm{K/m}$.

Table 1
Comparison of the L_2 error norm generated by the DHN theory with regard to FEM Q4 and FEM Q8 predictions.

	T'	Φ_2	Φ_3
L_2 norm (w.r.t. FEMQ4)	0.0137	0.0383	0.1237
L_2 norm (w.r.t. FEMQ8)	0.0102	0.0062	0.0177

from the DHN predictions are quantitatively evaluated against the benchmark solutions, using the relative L_2 error as a metric:

$$L_2 \text{ error } = \frac{\|V - V^*\|_2}{\|V^*\|_2} \tag{22}$$

where V represents the DHN predictions (such as temperature, heat flux, displacements, and stresses) and V^* represents the corresponding benchmark solutions.

Figs. 4-5 depict the differences in the fluctuating temperature and transverse heat flux profiles predicted by the proposed adaptive DHN, 4- and 8-noded finite-element results. The neural network results are obtained after 10k iterations with learning rates of 1×10^{-2}, 5×10^{-3}, 2.5×10^{-3}, 1.25×10^{-3} and 6.25×10^{-4}(decaying by 50 % after each 2k iteration). The adaptive weights associated with all the collocation points are initialized to 1. They are updated in the first 5k iterations and are held constant thereafter. As observed in Figs. 4-5, the DHN can predict the fluctuating temperature and heat flux in the whole bulk domain well quantitatively, even in the vicinity of the pore boundary where pronounced local field concentrations are observed. We note that the temperature and heat flux fields predicted by the DHN theory are smoothly varying in the entire analysis domain. In contrast, the finite-element method ensures only the continuities of the predicted temperature, while the predicted heat flux is not continuous from one element to another, in particular with the lower-order finite-element technique.

Table 1 shows the L_2 error norms predicted by the proposed DHN technique with regard to the 4- and 8-noded finite-element reference

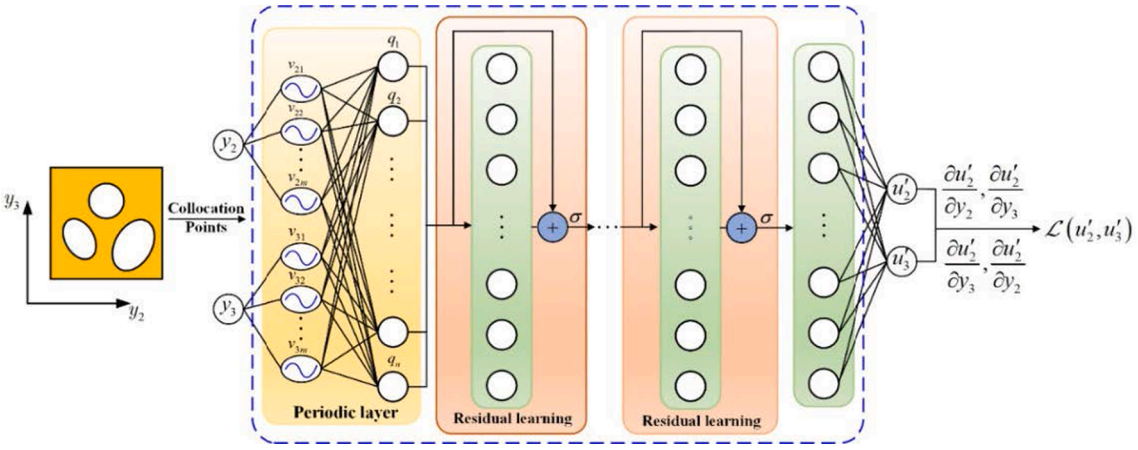

Fig. 6. Physically informed deep residual neural network acting as a surrogate model for homogenization of periodic elastic microstructured materials.

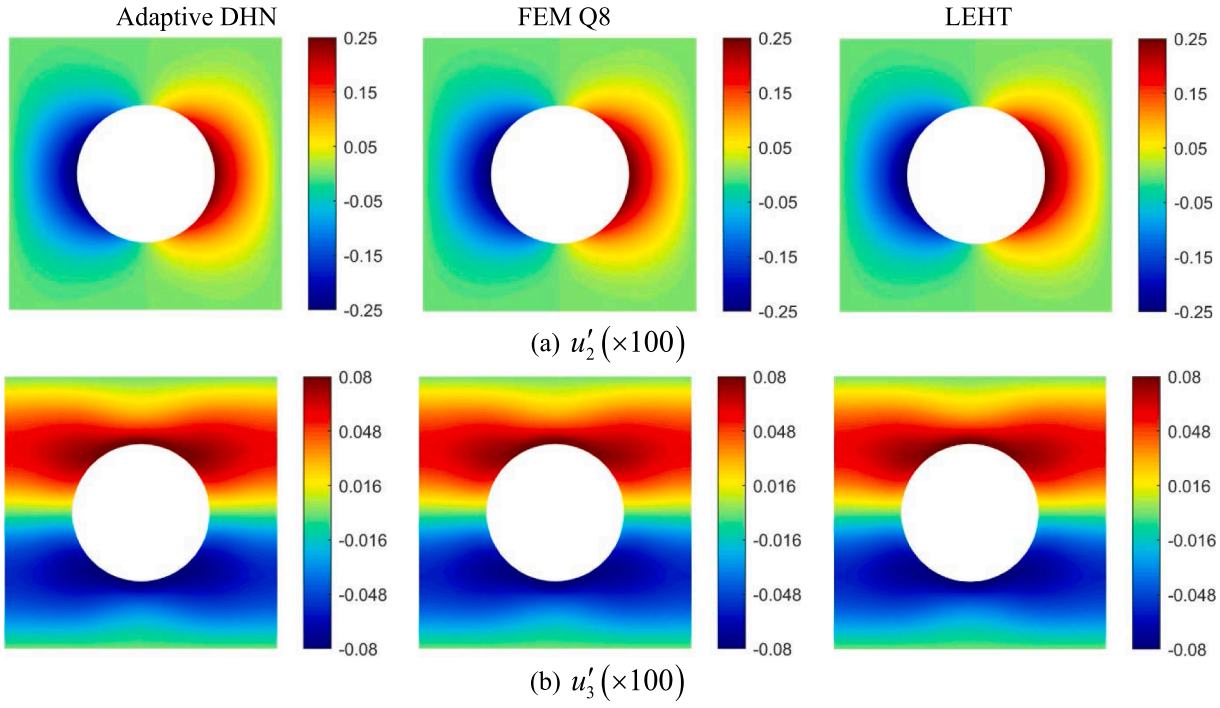

Fig. 7. Comparison of the fluctuating displacements u_2' and u_3' for periodic arrays with circular cylindrical pore under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ predicted by the DHN, 8-noded FEM and the LEHT predictions of Drago and Pindera [14].

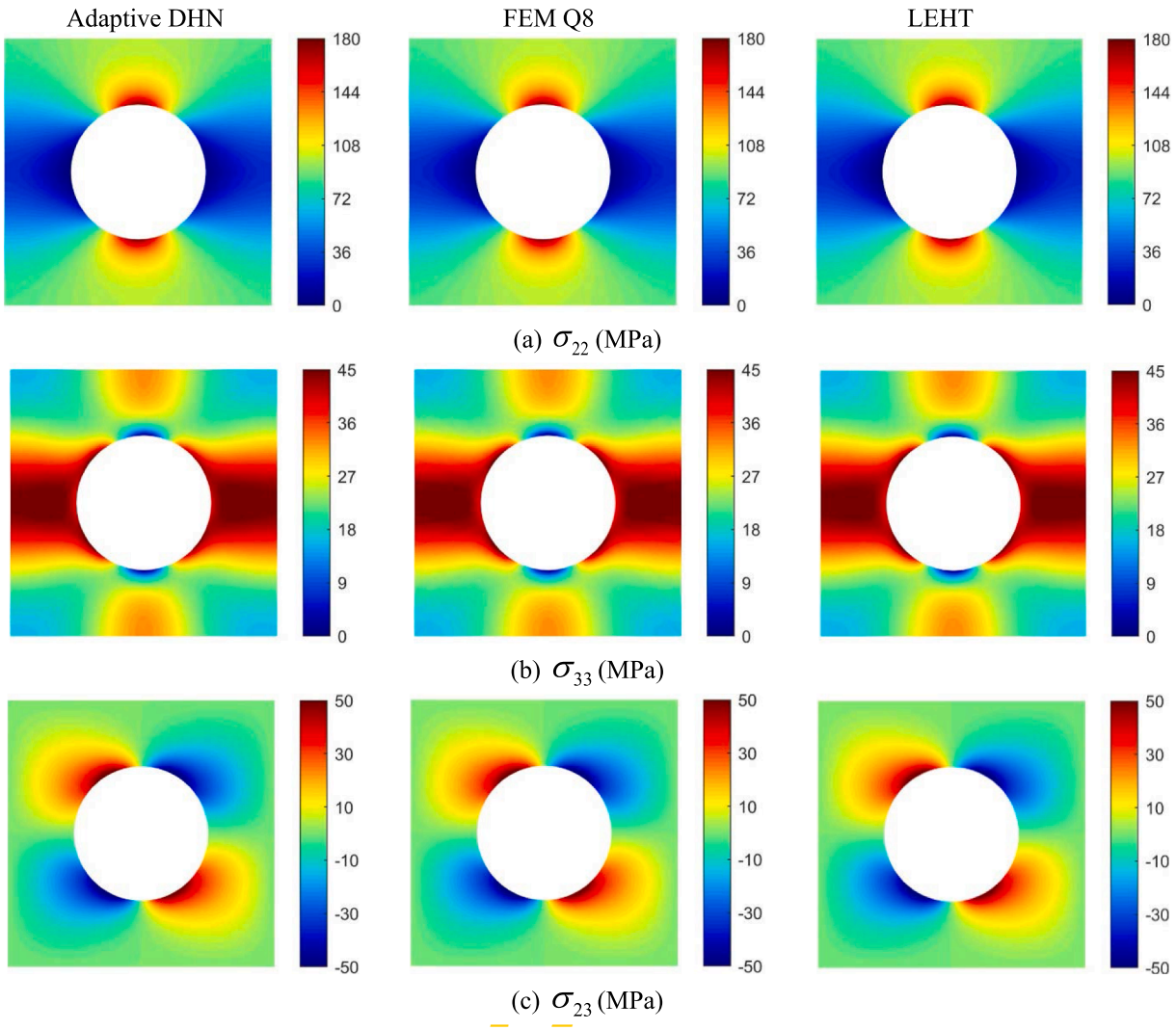

Fig. 8. Comparison of the local stress field σ_{22}, σ_{33} and σ_{23} for periodic arrays with circular cylindrical pore under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ predicted by the DHN, 8-noded FEM, and the LEHT predictions of Drago and Pindera [14].

solution. It is observed that the L_2 error for various fields predicted by the DHN technique is around 1 % with regard to the high-field 8-noded finite-element results. As expected, a little bit higher error is predicted with regard to 4-noded finite-element predictions.

4.2. *Elastic porous media*

In the second example, the proposed DHN theory is utilized to predict the fluctuating displacements (u_2', u_3'), and to infer the stress field distributions in elastic porous architectures under known macroscopic strain, as illustrated in Fig. 6. We demonstrate the accuracy of the proposed DHN technique vis-à-vis 8-noded finite-element results and the elasticity-based locally-exact homogenization theory predictions (LEHT), developed by Pindera and his coworkers [14]. The latter method utilizes precise solutions of governing differential equations by employing Fourier series representation of displacement fields and the balanced variational principle to enforce periodic boundary conditions. This stands in contrast to the finite-element method, which relies on the discretization of unit cells. Consequently, the LEHT serves as an excellent benchmark for evaluating the accuracy of the proposed DHN theory. However, it's crucial to note that the LEHT has a limitation: it can only consider single-inclusion unit cells arranged in square and hexagonal arrays.

We consider an aluminum periodic array with circular cylindrical pores. The Young's modulus and Poisson's ratio of the bulk aluminum are 72.4 GPa and 0.33, respectively. The pore volume fraction is prescribed as 20 %. The small diameter of the fibers in comparison to the overall dimensions of the unit cell results in elevated stress and deformation gradients at the pore interface, posing a rigorous test for the accuracy of the methods employed. A macroscopic transverse normal strain $\bar{\varepsilon}_{22} = 1\%$ is applied since it is the most demanding.

As before, the DHN model is constructed using fully connected neural network layers with 1 periodic layer, 3 hidden layers, and 30 neurons per layer. It is trained by Adam optimizer with the same decreasing learning rate schedule and the same number of epochs as in Section 4.1. The number of training points is 8k points (randomly sampled) in the bulk materials and 360 points on the interface, respectively (Not shown). The unit cell in finite element simulation was discretized into 48×12 elements in circumferential and radial directions, respectively, while 36 harmonic terms were utilized in the LEHT simulation such that reliable solutions were obtained. It's crucial to emphasize that the solution methodologies utilized in the DHN, FEM, and LEHT are fundamentally distinct, providing robust support for the thorough validation of the deep homogenization theory and the subsequent conclusions drawn.

Fig. 7 compares the difference between the proposed adaptive DHN, 8-noded finite-element, and the LEHT predictions of the fluctuating

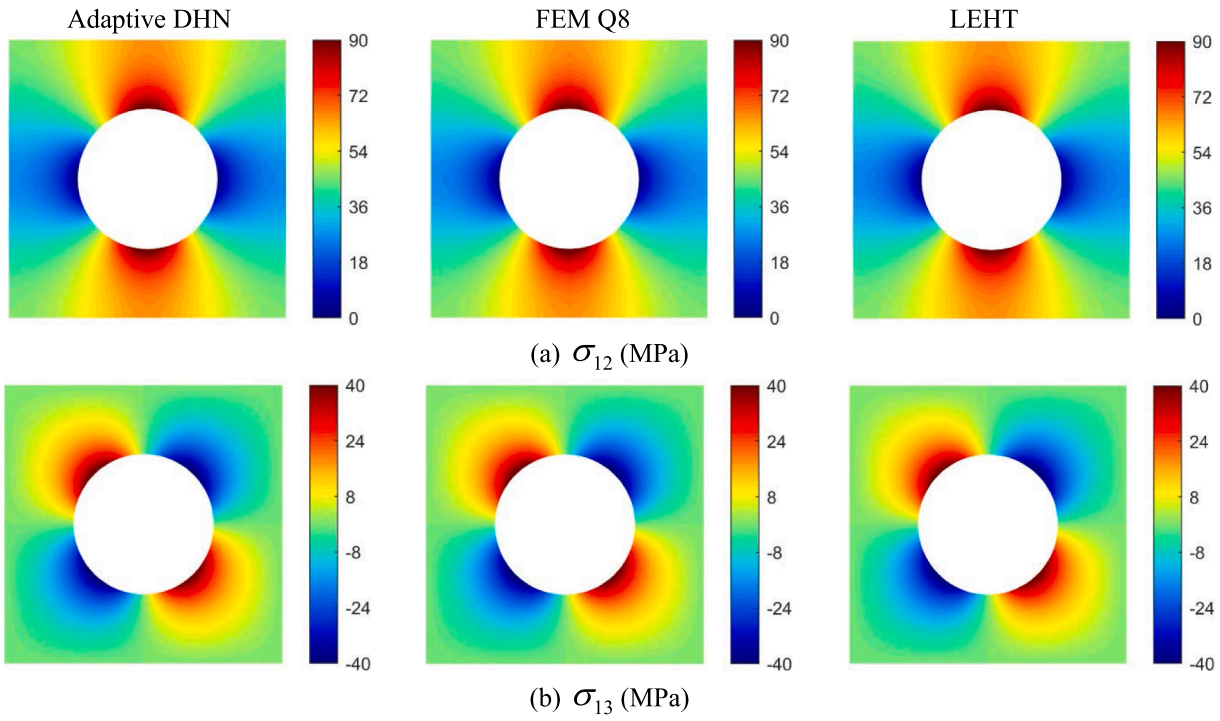

Fig. 9. Comparison of the local stress field σ_{12}, σ_{13} for periodic arrays with circular cylindrical pores under macroscopic strain $\bar{\varepsilon}_{12} = 1\%$ predicted by the DHN, 8-noded FEM and the LEHT predictions of Drago and Pindera [14].

Table 2
Comparison of the L_2 error norm predicted by the DHN with respect to the 8-noded FEM, and the LEHT predictions of Drago and Pindera [14].

	$u_2^{'}$	$u_2^{'}$	σ_{22}	σ_{33}	σ_{23}	σ_{12}	σ_{13}
L_2 norm(w.r.t. FEM Q8)	0.0078	0.0136	0.0051	0.0152	0.0139	0.0033	0.0013
L_2 norm(w.r.t. LEHT)	0.0079	0.0138	0.0050	0.0141	0.0141	4.07e-04	0.0015

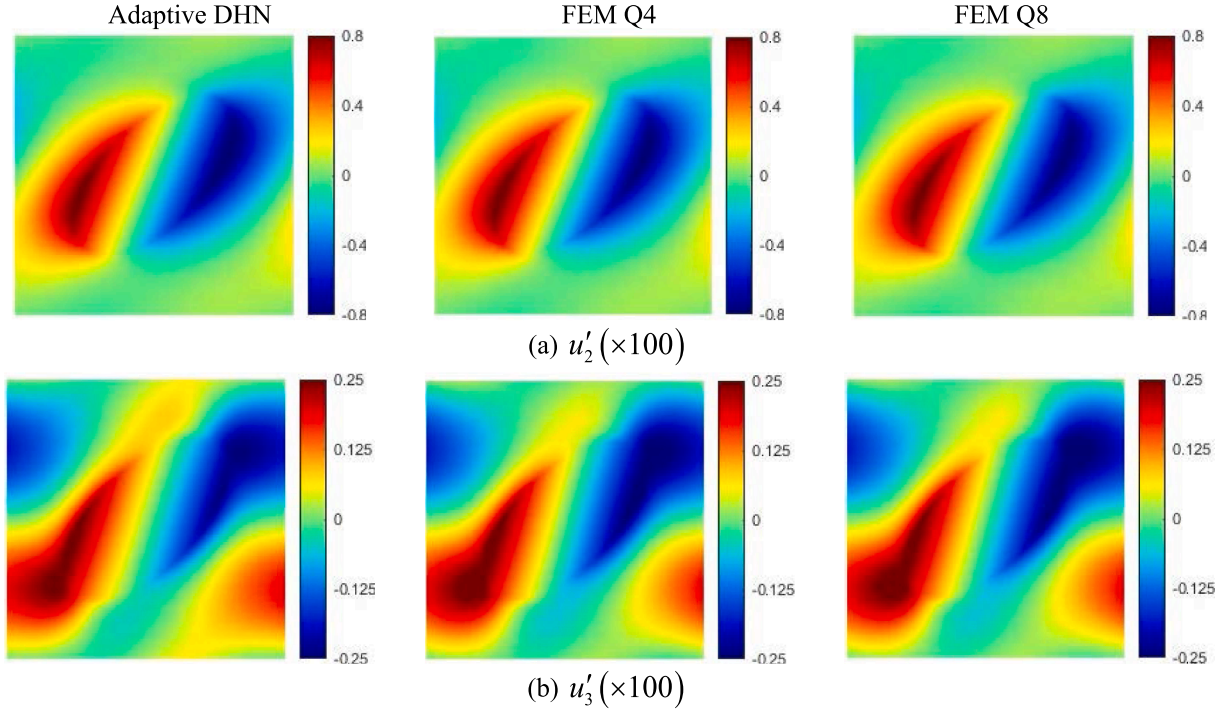

Fig. 10. Comparison of the fluctuating displacements $u_2^{'}$ and $u_3^{'}$ for periodic arrays of unidirectional fiber under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ predicted by the adaptive DHN, 4- and 8-noded FEM.

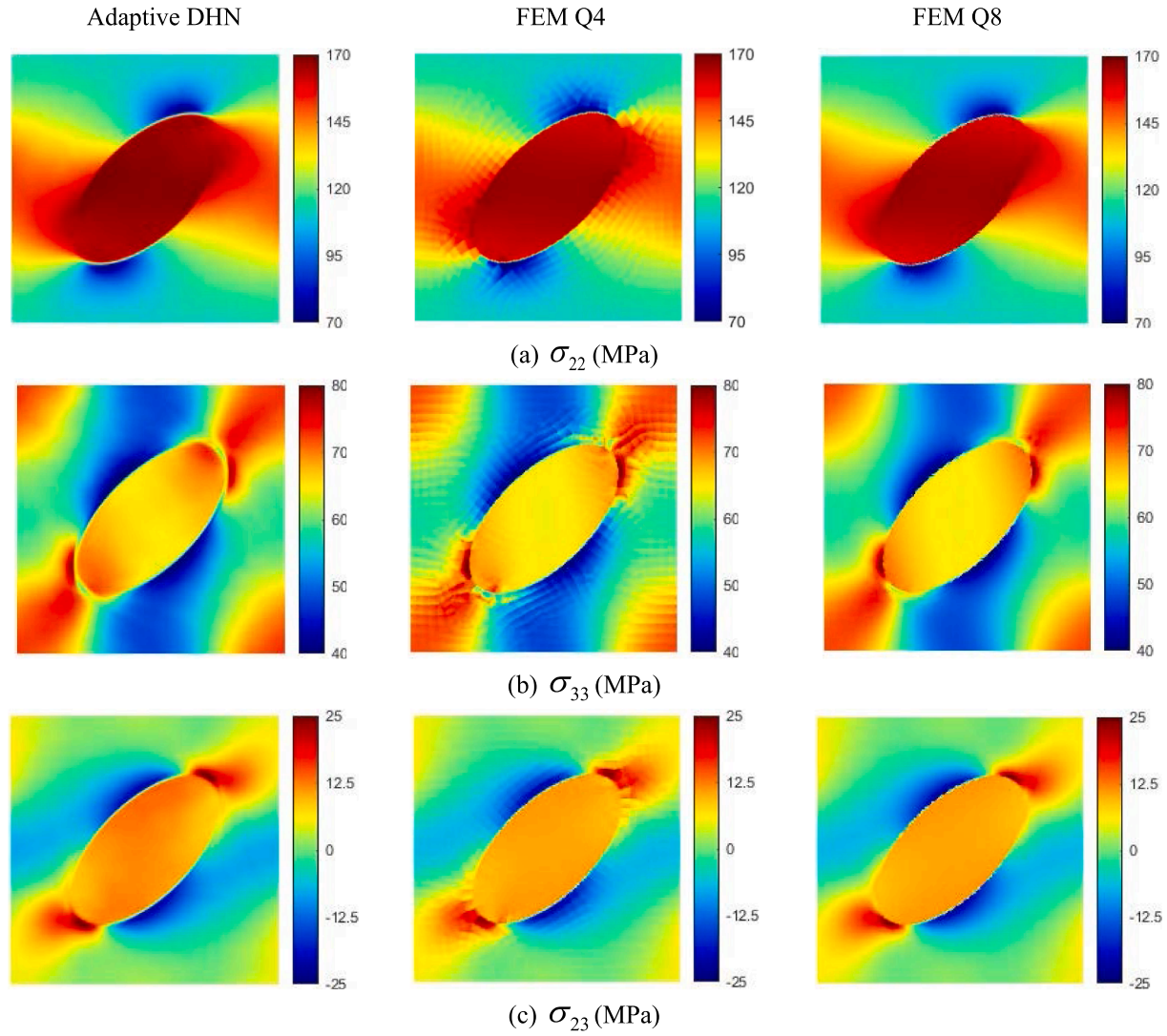

Fig. 11. Comparison of the local stress field σ_{22}, σ_{33} and σ_{23} for periodic arrays of unidirectional fiber under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ predicted by the adaptive DHN, 4- and 8-noded FEM.

displacement fields u_2' and u_3' . As expected, the DHN predicts indistinguishable results vis-à-vis the high-fidelity finite-element results and exact elasticity solution, providing good support for the developed adaptive DHN technique. Fig. 8 illustrates comparison of stress fields σ_{22} , σ_{33} and σ_{23} predicted by the three methods. Once again, the correlation between the three approaches is seen to be excellent for the three stress components. The stress concentrations at the pore boundary under this type of loading are captured by the DHN theory with sufficient accuracy.

As mentioned, the formulation of the axial shear problem by any combinations of $\bar{\varepsilon}_{12}$ and $\bar{\varepsilon}_{13}$ shows that it is mathematically entirely analogous to the transverse conductivity problem under macroscopic loading $\overline{H}_2$ and $\overline{H}_3$. Therefore, we use directly the network architecture developed in Section 4.1 to obtain the solution of stress fields σ_{12} and σ_{13} of elastic periodic porous arrays under axial shear loading by $\bar{\varepsilon}_{12} = 1\%$. For the FEM and LEHT computations, the strictly axial shear problem formulation is followed. The comparison of the axial shear stress distributions is displayed in Fig. 9. Similar comments can be applied to these results.

We end this section by comparing the L_2 error norm predicted by the DHN technique with respect to the finite-element and locally exact homogenization predictions, respectively (see Table 2). As anticipated, the differences between the proposed adaptive DHN and the benchmark solutions are negligible.

4.3. *Fibrous composites*

Next, we examine the fluctuating displacements and local stress fields in square array of unidirectional fiber with 20 % fiber volume fraction. The distribution of the Lamé constants in the unit cell microstructure is represented by the following expressions:

$$\lambda(\boldsymbol{y}) = 29.14\left(1+\tanh\frac{1-\frac{(y_2+y_3)^2}{8}-\frac{(y_2-y_3)^2}{2}}{0.05}\right)+52.84(\text{GPa}) \tag{23}$$

$$\mu(\boldsymbol{y}) = 69.72\left(1+\tanh\frac{1-\frac{(y_2+y_3)^2}{8}-\frac{(y_2-y_3)^2}{2}}{0.05}\right)+27.22(\text{GPa}) \tag{24}$$

As before, a macroscopic transverse normal strain $\bar{\varepsilon}_{22} = 1\%$ is applied. We employ a fully connected DHN model with 1 periodic layer, 3 hidden layers, and 30 neurons per layer. 10k collocation points generated randomly are utilized to train the network model. The neural network results are obtained after 15k iterations by Adam optimizer with learning rates of 1×10^{-2}, 5×10^{-3}, 2.5×10^{-3}, 1.25×10^{-3} and 6.25×10^{-4}(decaying by 50 % after every 2.5k iteration). The adaptive weights

Table 3
Comparison of the execution time between the proposed technique and the conventional finite-element method (averaged based on three runs).

Methods	DHN	FEM Q4	FEM Q8
Execution time (seconds)	2053.1	2.2	9.1

associated with all the collocation points are initialized to 1. They are updated in the first 7.5k iterations and are held constant thereafter.

Fig. 10 shows the comparison of the fluctuating displacements $u_2^{'}$ and $u_3^{'}$ distributions generated by the proposed technique and the 4- and 8-noded FEM approaches. It is observed a good agreement between the DHN and finite-element reference results. Comparison of the local stress field σ_{22}, σ_{33} and σ_{23} generated by the three approaches is given in Fig. 11. Similar comments apply in this case regarding the proposed theory's accuracy, as well as the smoothness of the stress field of the DHN and the conventional finite-element techniques.

Table 3 presents the comparison of the execution time between the proposed technique and the conventional finite-element method. These simulations were performed on a personal computer with 13th Gen Intel (R) Core (TM) I9-13900H @ 2.60 GHz, 32 GB memory, 64-bit operating system, and x64-based processor. We note that the proposed DHN approach consumes significantly higher execution time than that in both Q4 and Q8 FEM techniques. This is because the DHN approach transforms a linear elasticity unit cell problem into a nonlinear optimization problem with a very rough energy landscape. Gradient-based optimization within this landscape is particularly computationally expensive. Despite the neural network approach cannot compete with the conventional numerical method in terms of computational efficiency, advances in computational power have unlocked the potential of the neural network approach for solving boundary value problems in a reasonable time [22].

5. Multi-inclusion porous media

The applicability of the proposed DHN theory is further demonstrated by considering a unit cell with locally irregular pore distributions. This microstructure is an ideal candidate for assessing the performance of the deep homogenization theory since the pore–pore interaction produces significant deformation/stress gradients which are more crucial in this type of unit cell microstructure [29,30]. In particular, the loss function for the random microstructure is quite complex and non-convex and it is harder to optimize relative to the single inclusion case. It is also the objective of this section to justify the adaptive homogenization theory developed in this manuscript. Specifically, the PDE residuals in the bulk material, the traction residuals on pore interfaces, and the displacement and stress fields predicted by the adaptive DHN and the conventional nonadaptive counterpart are compared.

The unit cell architecture is characterized by a unit square containing locally irregular pores as shown in Fig. 12. The pores at the unit cell boundary are precisely chopped such that a complete pore with periodic geometry can be formed at the opposite sides of the unit cell. The total volume fraction of the pores is set to be 20 % as before. A macroscopic transverse normal strain $\bar{\varepsilon}_{22} = 1\%$ is applied to the unit cell. The finite-element unit cell for comparison with the DHN techniques is discretized into 1141 eight-noded elements with unconstructed meshes (not shown). To assess the performance of DHN models on an equal footing, both the adaptive and nonadaptive DHN models are constructed with 1 periodic layer with 30 neurons and 5 hidden layers with 50 neurons per layer. The number of training points is given as 7969 collocation points

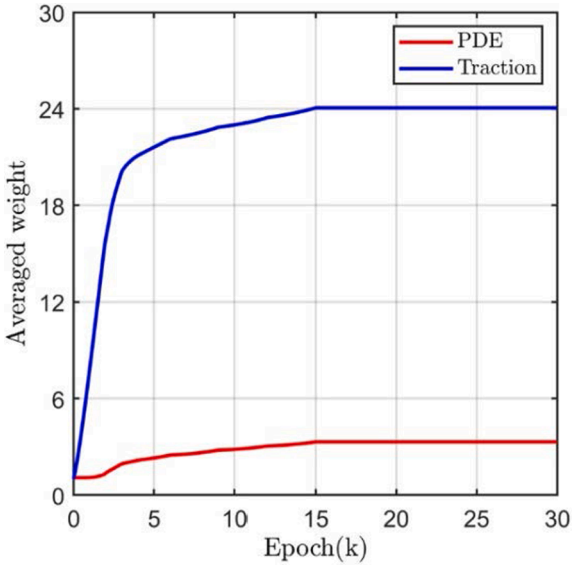

Fig. 13. Comparison of the average weights associated with collocation points within the bulk materials and on the interface for evaluating PDE and tractions loss as a function of training epoch.

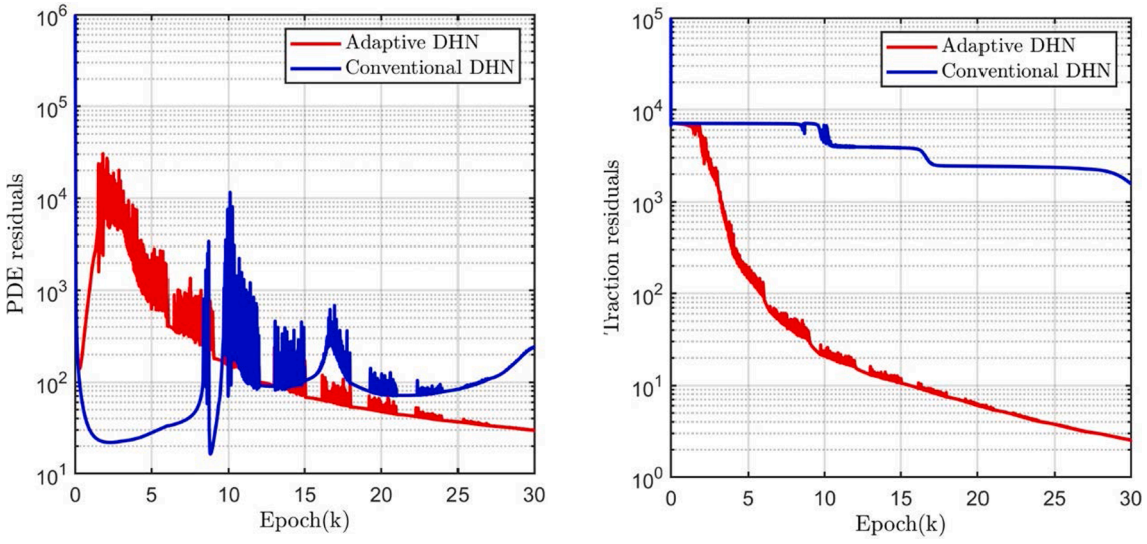

Fig. 12. Comparison of the PDE and traction residuals as a function of training epoch generated by the adaptive DHN and conventional DHN approaches.

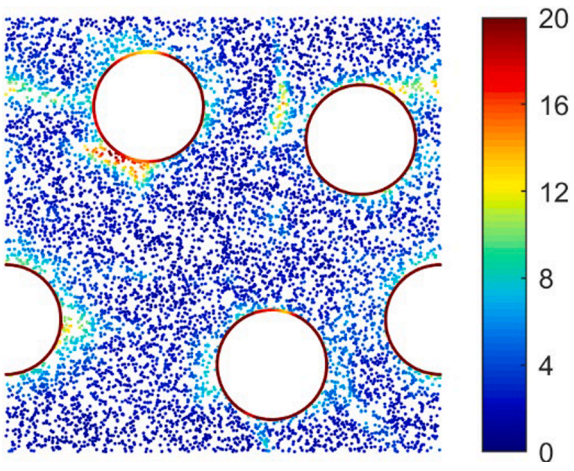

Fig. 14. Pointwise weight distributions associated with collocation points within the bulk materials and on the interface.

inside the bulk materials and 360 points at each complete pore boundary. The DHN models are trained by Adam optimizer with 30k iterations with an initial learning rate of 1×10^{-2} which is decreased by half for every 3k iterations. While the adaptive weights are updated in the first 15k iterations and hold constants thereafter, the loss weights for every collocation point are set to 1 for the non-adaptive method.

Fig. 12 presents the evolution of the PDE and traction residuals as a function of the training epoch by adaptive and non-adaptive DHN schemes. Note that for a direct comparison of the two approaches, the weights were removed from the loss residuals in the adaptive technique. It is demonstrated that in the adaptive DHN method, both PDE and traction residuals continuously decrease after 2k iterations and significantly low loss values have been achieved after 30k epochs relative to the non-adaptive scheme. Furthermore, in the non-adaptive DHN method, a substantial competing effect between the PDE and traction residuals exists which leads to significant variation in the PDE residuals and high ultimate loss values, indicating the conventional DHN might fail to train.

Fig. 13 shows the comparison of the average weights associated with collocation points within the bulk materials and on the interface for evaluating PDE and traction loss, respectively, as a function of the training epoch. As observed, while the averaged weights are continuously increasing, the rate of increase of the traction weights is much faster than that of the PDE weights since the traction loss tends to be neglected by the neural network. This indicates that the adaptive DHN theory has learned that traction loss is a more important part of the solution. The pointwise weight distributions after 30k training iterations for the PDE and traction training points are displayed in Fig. 14. It is observed that higher weight values are obtained for the collocation points in the vicinity of the pore boundary which is consistent with the observation in Fig. 13. Note that significant stress concentrations typically occur in the vicinity of the pore boundary, the proposed DHN method adaptively assigns higher weight values to the collocation points in the affected region, indicating the effectiveness of the proposed technique.

Figs. 15-16 illustrate the comparison of the predicted fluctuating displacements and stress fields by the conventional DHN, adaptive DHN, and high-fidelity finite-element baseline solutions. It is demonstrated that the adaptive DHN theory shows a good level of accordance with the finite-element predictions. The displacement and stress concentrations in the vicinity of the pore interface are captured with high accuracy. In contrast, the non-adaptive DHN captures only the displacement and stress components at certain pore boundaries, while most of the predictions are completely off from the baseline solution. Therefore, introducing adaptative weights into the DHN has been justified.

6. Conclusions

An adaptive deep homogenization neural network model has been proposed for the micromechanical analysis of unidirectionally periodic heterogeneous arrays with different microstructures. This method adopts a two-scale expansion of the temperature and displacement fields

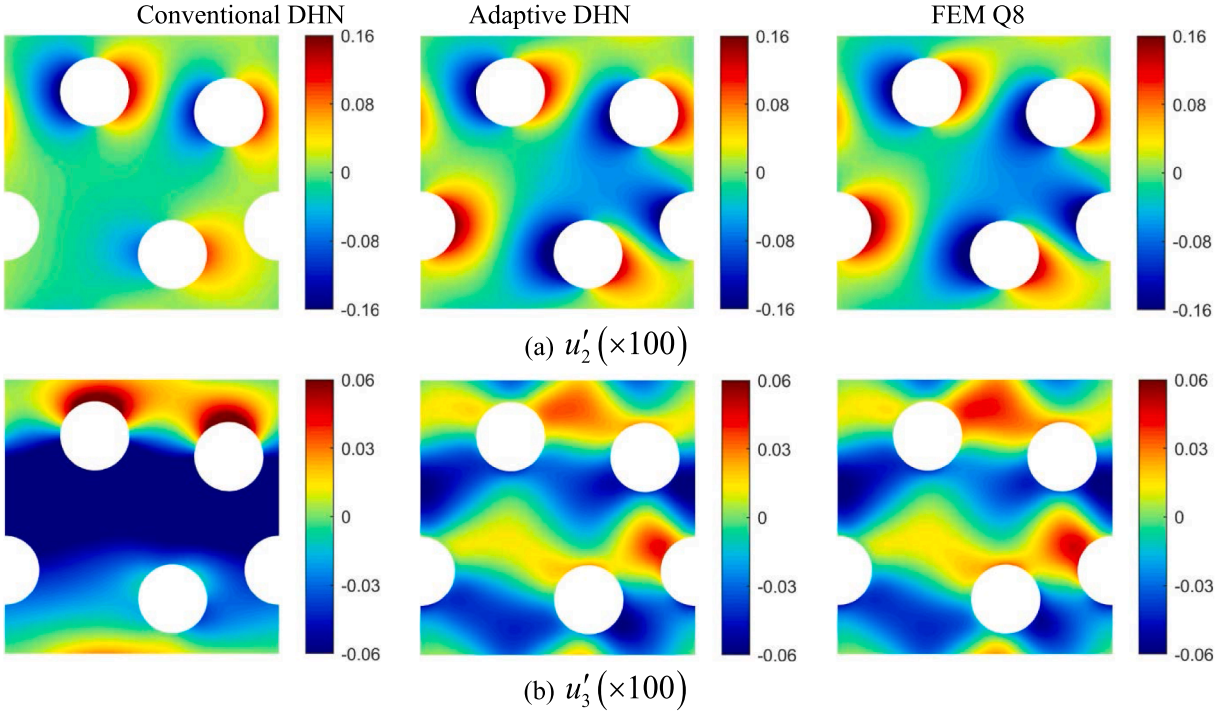

Fig. 15. Comparison of the fluctuating displacements u_2' and u_3' for periodic arrays with circular cylindrical pore under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ for the unit cell with locally irregular pore distribution

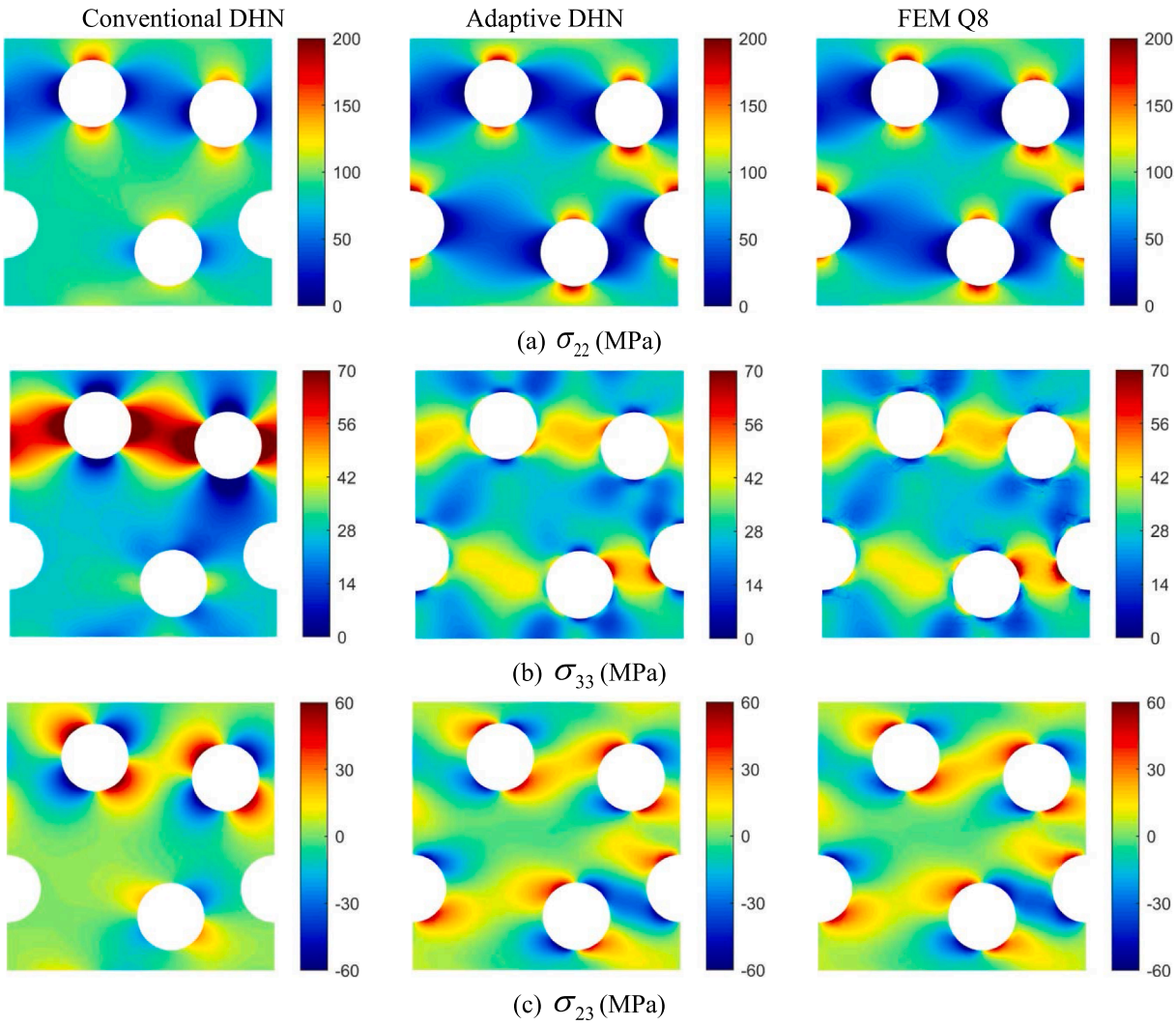

Fig. 16. Comparison of the local stress field σ_{22}, σ_{33} and σ_{23} for periodic arrays with circular cylindrical pore under macroscopic strain $\bar{\varepsilon}_{22} = 1\%$ for the unit cell with a locally irregular pore distribution

into macroscopic and microscopic contributions. The microscopic problem is solved by satisfying periodicity boundary conditions, stress equilibrium, or steady-state heat conduction equations, together with interfacial traction-free or adiabatic boundary conditions, respectively.

To facilitate training the neural network, a periodic layer is introduced in order to impose exactly the periodicity boundary conditions at the unit cell boundaries. Furthermore, fully-trainable weights are applied on each collocation point associated with the PDE and interfacial loss terms. These weights are trained concurrently with the network weights. The numerical results presented in this work demonstrate that the DHN is capable of accurately capturing the displacement and stress field of the periodic microstructural arrays with comparable accuracy to the high-fidelity finite element and exact solutions. It has also been shown that the trainable weights in the new DHN technique facilitate more accurate neural network solutions by automatically weighting more collocation points in difficult regions of the unit cell solution in the loss function.

The present DHN framework can be further extended in order to describe the nonlinear material behavior and continuum damage mechanisms involving path dependence and irreversibility. Such extension necessitates the implementation of a radial return algorithm at each loading step that ensures the proper storage and updating of state variables to obey the Kuhn–Tucker consistency conditions. This work will be conducted in a forthcoming publication.

CRediT authorship contribution statement

Jiajun Wu: Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Qiang Chen:** Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Jindong Jiang:** Writing – review & editing, Software. **George Chatzigeorgiou:** Writing – review & editing, Validation, Conceptualization. **Fodil Meraghni:** Writing – review & editing, Validation, Conceptualization.

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] Chen Q, Wang G, Pindera M-J. Finite-volume homogenization and localization of nanoporous materials with cylindrical voids. Part 1: Theory and validation. Eur J Mech-A/Solids 2018;70:141–55.

[2] Khatam H, Chen L, Pindera M-J. Elastic and plastic response of perforated metal sheets with different porosity architectures. J Eng Mater Technol 2009;131.
[3] Zieliński TG, Opiela KC, Pawłowski P, Dauchez N, Boutin T, Kennedy J, et al. Reproducibility of sound-absorbing periodic porous materials using additive manufacturing technologies: Round robin study. Addit Manuf 2020;36:101564.
[4] Huang M, You W, Zhang L, Li Y, Niu M, Lv Y. Structure and circuit modeling of frequency domain polarization characteristics for porous composite material. Compos Sci Technol 2022;224:109457.
[5] Wang Y, Chen Z, Mei D, Zhu L, Wang S, Fu X. Highly sensitive and flexible tactile sensor with truncated pyramid-shaped porous graphene/silicone rubber composites for human motion detection. Compos Sci Technol 2022;217:109078.
[6] Fritzen F, Forest S, Böhlke T, Kondo D, Kanit T. Computational homogenization of elasto-plastic porous metals. Int J Plast 2012;29:102–19.
[7] Koiter W, Langer R. Stress distribution in an infinite elastic sheet with a doubly-periodic set of equal holes. Madison: The University of Wisconsin Press; 1960. p. 191–213.
[8] Grigolyuk E, Kursin L, Fil'shtinskii L. A method for the solution of doubly-periodic problems in the theory of elasticity. Prikl Mekh. 1965;1:22-31.
[9] Meijers P. Doubly-Periodic Stress Distributions in Perforated Plates. Delft: Technological University; 1967.
[10] Slot T, O'Donnell WJ. Effective elastic constants for thick perforated plates with square and triangular penetration patterns. Journal of Engineering for Industry 1971;93:935–42.
[11] Khatam H, Pindera M-J. Plastic deformation modes in perforated sheets and their relation to yield and limit surfaces. Int J Plast 2011;27:1537–59.
[12] Chen Q, Wang G, Pindera M-J. Homogenization and localization of nanoporous composites - a critical review and new developments. Compos B Eng 2018;155:329–68.
[13] Chen Q, Pindera M-J. Homogenization and localization of elastic-plastic nanoporous materials with Gurtin-Murdoch interfaces: an assessment of computational approaches. Int J Plast 2020;124:42–70.
[14] Drago AS, Pindera M-J. A locally exact homogenization theory for periodic microstructures with isotropic phases. J Appl Mech 2008;75.
[15] Wang G, Chen Q, He Z, Pindera M-J. Homogenized moduli and local stress fields of unidirectional nano-composites. Compos B Eng 2018;138:265–77.
[16] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707.
[17] Nguyen-Thanh VM, Anitescu C, Alajlan N, Rabczuk T, Zhuang X. Parametric deep energy approach for elasticity accounting for strain gradient effects. Comput Methods Appl Mech Eng 2021;386:114096.
[18] Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh VM, Guo H, Hamdia K, et al. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: concepts, implementation and applications. Comput Methods Appl Mech Eng 2020;362:112790.
[19] Haghighat E, Raissi M, Moure A, Gomez H, Juanes R. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Comput Methods Appl Mech Eng 2021;379:113741.
[20] Abueidda DW, Lu Q, Koric S. Meshless physics-informed deep learning method for three-dimensional solid mechanics. Int J Numer Meth Eng 2021;122:7182–201.
[21] Jiang J, Zhao J, Pang S, Meraghni F, Siadat A, Chen Q. Physics-informed deep neural network enabled discovery of size-dependent deformation mechanisms in nanostructures. Int J Solids Struct 2022;236–237:111320.
[22] Jiang J, Wu J, Chen Q, Chatzigeorgiou G, Meraghni F. Physically informed deep homogenization neural network for unidirectional multiphase/multi-inclusion thermoconductive composites. Comput Methods Appl Mech Eng 2023;409:115972.
[23] Wu J, Jiang J, Chen Q, Chatzigeorgiou G, Meraghni F. Deep homogenization networks for elastic heterogeneous materials with two- and three-dimensional periodicity. Int J Solids Struct 2023;284:112521.
[24] Dong S, Ni N. A method for representing periodic functions and enforcing exactly periodic boundary conditions with deep neural networks. J Comput Phys 2021;435:110242.
[25] McClenny LD, Braga-Neto UM. Self-adaptive physics-informed neural networks. J Comput Phys 2023;474:111722.
[26] Bensoussan A, Lions J-L, Papanicolaou G. Asymptotic analysis for periodic structures: North Holland, Amsterdam,Netherlands, 1978.
[27] He Z, Pindera M-J. Locally exact asymptotic homogenization of periodic materials under anti-plane shear loading. Eur J Mech A Solids 2020;81:103972.
[28] Suquet PM. Homogenization Techniques for Composite Media. Lecture Notes in Physics. Berlin, Heidelberg: Springer Berlin Heidelberg; 1987. p. 193-8.
[29] Yin S, He Z, Pindera M-J. A new hybrid homogenization theory for periodic composites with random fiber distributions. Compos Struct 2021;269:113997.
[30] Yin S, Pindera M-J. Hybrid homogenization theory with surface effects: Application to columnar nanoporous materials. Eur J Mech A Solids 2023;101:105050.
[31] Wang L, Yang Q, Yang Y, Luo K, Bai R. Porous polyurethane hydrogels incorporated with CMC for eliminating methylene blue from water. Int. J. Smart Nano Mater 2023;14(1):57–76.