Interaction-based material network: a general framework for (porous) microstructured materials

Van Dung Nguyen[a,1], Ludovic Noels[a,*]

[a] *Computational & Multiscale Mechanical of Materials (CM3), Department of Aerospace and Mechanical Engineering, University of Liège, Quartier Polytech 1, Allée de la Découverte 9, B-4000 Liège, Belgium*

Abstract

A material network consisting of discrete material nodes and their interactions can represent complex microstructure responses. Under this interaction viewpoint, the material network can be viewed as a trainable system involving fitting parameters including not only the weights of the material nodes but also the parameters characterizing their interactions. As opposed to the other existing works, this interaction-based material network does not rely on the micromechanics of multiple-phase laminates but on constraining all requirements of a truly microscopic boundary value problem including the stress and strain averaging principles and the Hill-Mandel energetically consistent condition. Consequently, the proposed framework can be applied to microstructures with the presence of voids, which is not achievable with the laminate theory. To make a material network become a surrogate of a full-field microscopic model, this work proposes two different training procedures to calibrate its fitting parameters. On the one hand, a nonlinear training procedure is proposed considering sequential data collected from finite element simulations on the full-field model subjected to proportional loading paths. On the other hand, a linear elastic training procedure considers only the elastic response of the heterogeneous material. The accuracy and efficiency of the proposed framework for microstructures with the presence of voids are demonstrated by comparing the predictions of the trained material networks with the ones of the direct numerical simulations in both contexts of virtual testing and multiscale simulations. It is also shown that the linear elastic training procedure requires a lower computational cost but could lead to less accurate predictions in comparison with the nonlinear training procedure.

Keywords: Material network, Porous material, Machine learning, Inelastic training

1. Introduction

With the rapidly increasing application of heterogeneous materials in engineering fields, the accurate evaluation of the relationship between the structural behavior and the intrinsic properties of the material microstructure is mandatory. To enable the numerical simulation of structures made of heterogeneous materials, an effective framework known as computational homogenization (FE^2) has been developed to directly account for complex material microstructure and constitutive behaviors at lower scales [1, 2, 3, 4, 5, 6, 7]. In the FE^2 framework, the constitutive relationship at each macroscopic material point is obtained from the resolution of a microscopic Boundary Value

*Corresponding author, Phone: +32 4 366 48 26, Fax: +32 4 366 95 05

Email addresses: vandung.nguyen@ulg.ac.be (Van Dung Nguyen), L.Noels@ulg.ac.be (Ludovic Noels)

[1] Postdoctoral Researcher at the Belgian National Fund for Scientific Research (FNRS)

Problem (BVP) defined at that point. However, the requirement of solving microscopic BVPs at all the integration points by the finite element method involves high numerical resources in terms of time and memory, which limits the applicability of this technique to reduced size problems. We will address this issue within the context of material network, which is a connected network of discrete material nodes.

Many studies have been focusing on building surrogate models of the microscopic BVP in the FE^2 framework for improving computational efficiency. The surrogate model is then considered at each macroscopic integration point in the place of the original full-field microscopic BVP, in which the computational costs and accuracy are balanced. One possibility is to use Reduced Order Models (ROM), which have been extensively developed to reduce the computation cost in the finite element analysis, see *e.g.* the works [8, 9, 10, 11, 12, 13, 14, 15], generally for the microscopic BVPs. In these ROM-based surrogate models, governing equations of the full-field model can be solved with a reduced number of degrees of freedom in a reduced order space based on offline full-field solution snapshots, leading to considerably reduced computational cost. Alternatively, the surrogate model of a microscopic BVP can be constructed based on machine learning (ML) techniques, *e.g.* artificial neural networks and deep learning, in which the complex input-output relationships representing the constitutive behavior can be approximated through a training step using data obtained either from offline simulations or from experimental observations [16, 17, 18, 19, 20, 21, 22, 20, 23, 24, 25, 26, 27]. However, since both the ROM-based and ML-based surrogate models rely on the offline database which is not easily built large enough when dealing with irreversible material responses, their prediction could be inaccurate when extrapolating beyond the offline sampling space.

An alternative way to build a computationally efficient surrogate of the microscopic BVP relies on a network structure based on mechanistic building blocks. Following this idea, the deep Material Network (MN) method was proposed in [28] for two-phase composite materials with a binary hierarchical topological structure based on two-phase laminates as mechanistic building blocks. The deep MN was applied to three-dimensional microstructures [29] and augmented by cohesive zone models [30] and material failure [31]. The thermodynamic consistency of the deep MN was investigated in [32, 33] in which the rotation-free deep MNs based on laminates with variable direction of lamination were proposed. More complex micro-mechanical models such as mean-field homogenization in the mechanistic building blocks was also recently considered for woven composite materials [34]. A deep MN consists of fitting parameters, which are identified for a given microstructure through a training process using only linear elastic simulations. Once trained, this deep MN is able to predict nonlinear material behaviors both accurately and efficiently although only linear elastic data were used during the offline training, thanks to the ability of the deep MN to learn the topology representation of the material microstructure. Recently, a MN has been framed from the interaction viewpoint [35] by satisfying all requirements of a truly microscopic BVP, *i.e.* stress and strain averaging principles and Hill-Mandel energetically consistent condition. The interaction viewpoint allows building a material network, a so-called interaction-based MN, with the material nodes and their interactions consisting of multiple interaction mechanisms. The interaction viewpoint allows not only the easy evaluation of the response of material networks with an arbitrary constitutive law considered at each material node, through the resolution of well-defined governing equations, but also the consideration of an arbitrary number of phases independently of the network architecture. A multiple-phase laminate mechanistic building block can be considered as a special case consisting of multiple interaction mechanisms. In these mentioned works with the exception of [34], the material network is built as a tree-hierarchical architecture of multiple-phase laminate blocks in analogy to deep artificial neural networks. However, the use of multiple-phase laminates as mechanistic building blocks is not always trivial with microstructures embedding voids. When a void is considered as a phase in a multiple-phase laminate, the balance of the laminate following the

laminate direction leads in a zero-stress component in this direction. A laminate of voided laminate also leads to zero-in-plane stress and, as a result, the possible stress state in each material node is too restrictive, which limits the learning ability of the material network in general as it will be discussed. We will address this gap within the framework of the interaction-based MN.

It is noted that the interaction-based MN does not require the definition of an underlying micro-mechanical model, such as multi-phase laminates [32, 33], although multi-phase laminates provide an efficient way to construct interaction mechanisms [35]. In this paper, the interaction-based MN is employed in a general form in which each interaction mechanism is no longer extracted from multiple-phase laminates but characterized by fitting parameters which are tuned with an offline training process. The voids are considered as a usual phase giving a zero-stress under all possible deformations. Since the network interactions do not involve the laminate theory but are unknowns characterized by fitting parameters, the zero-homogenized stress problem is circumvented. The main contributions of the work are:

- The interaction MN is built as a network of material nodes with interactions consisting of multiple interaction mechanisms. Each interaction consists of not only a direction but also of other fitting parameters characterizing the contribution of each material node involved in this mechanism. As a result, the fitting parameters of the interaction MN consist of not only the weights of the material nodes but also the fitting parameters characterizing the network interactions.

- The inference of the fitting parameters is constrained in the sense of the scale transition, including the stress and strain averaging principles and the enforcement of the Hill-Mandel (energetic consistency) condition. The latter ensures that the first law of thermodynamics is verified. The second-law of thermodynamics is also satisfied by using a thermodynamically consistent material law governing the mechanical behavior in each material node.

- Because there exist different constraints on fitting parameters in order to satisfy all requirements of a truly microscopic BVP, they are eliminated by direct constraint elimination in order to obtain an unconstrained minimization problem in the offline training process.

- A complete offline training framework of the interaction-based MN is proposed with two different strategies:

  - Linear elastic training procedure with the linear elastic data in terms of the homogenized elastic tensors collected from the direct finite element simulations of the microscopic BVP by considering the underlying phases being elastic and their elastic tensors being artificially and randomly generated.
  - Nonlinear training procedure with the sequential data collected from direct finite element simulations of the microscopic BVP following strain histories provided by a sequence of discrete values, including micro-structural evolution.

  For a given microstructure, the linear elastic training procedure does not require any information about the constitutive laws governing its physical phases unlike the nonlinear training procedure.

- The proposed framework is applied for microstructures with the presence of voids embedded in an elastoplastic solid. The accuracy of the trained material networks is verified by comparing the MN predictions with the ones of the direct numerical simulations of the full-field model in both contexts of virtual testing and multiscale simulations. Moreover, the evolution of the microstructure, *e.g.* porosity, is tracked in both the MN and direct numerical simulations.

The paper is organized as follows. In Section 2, the theory of the FE^2 technique is briefly recalled. A complete framework to make an interaction-based MN being a surrogate of a microscopic BVP is detailed in Section 3, in which the theory of the interaction MN is recalled and the fitting parameters are defined. In Section 4, the offline training strategy is proposed to infer the fitting parameters. In Section 5, the proposed framework is applied on microstructures with voids embedded in an elastoplastic matrix in which the accuracy, the efficiency and the thermodynamics consistency of the trained material networks, in both contexts of virtual testing and multiscale simulations, are demonstrated. Finally, conclusions are given in Section 6.

2. Computational homogenization

2.1. Macroscopic problem

Let us consider a body B subjected to a volumetric force $\bar{\mathbf{b}}$ and whose boundary is ∂B. The stress equilibrium equation and the strain-displacement relationship are respectively given by

$$\bar{\boldsymbol{\nabla}} \cdot \bar{\boldsymbol{\sigma}} + \bar{\mathbf{b}} = \mathbf{0} \text{ and } \bar{\boldsymbol{\varepsilon}} = \bar{\boldsymbol{\nabla}} \otimes^s \bar{\mathbf{u}} \text{ on } B\,, \tag{1}$$

where $\bar{\boldsymbol{\sigma}}$ is the Cauchy stress tensor, $\bar{\boldsymbol{\nabla}}$ is the gradient operator, $\bar{\boldsymbol{\varepsilon}}$ is the strain tensor, $\bar{\mathbf{u}}$ is the displacement vector, and $\otimes^s$ is the symmetric dyadic operator, *i.e.* $\mathbf{a} \otimes^s \mathbf{b} = \frac{1}{2}(\mathbf{a} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{a})$ for all arbitrary tensors $\mathbf{a}$ and $\mathbf{b}$. In order to find the displacement field $\bar{\mathbf{u}}$ for prescribed boundary conditions, a constitutive model must be introduced and can be rewritten under a general form:

$$\bar{\boldsymbol{\sigma}}(t) = \bar{\boldsymbol{\Sigma}}\left(\bar{\boldsymbol{\varepsilon}}(t), \bar{\mathbf{q}}(\tau \leq t)\right)\,, \tag{2}$$

in which $\bar{\mathbf{q}}$ is a vector of internal variables introduced to account for the history dependency. In a multi-scale problem, this vector of internal variables encompasses the thermodynamics internal variables distributions of the lower scale BVPs. The macroscopic problem stated by Eqs. (1 - 2) can be solved by means of the finite element method. The iterative finite element resolution requires the macroscopic tangent operator, $\bar{\mathbf{C}} = \frac{\partial \bar{\boldsymbol{\sigma}}}{\partial \bar{\boldsymbol{\varepsilon}}}$, which is directly deduced from the material law (2). In the context of concurrent multiscale analyzes, the constitutive law (2) at each macroscopic material point is obtained from the resolution of a microscopic problem attached to that point through the so-called downscaling and upscaling processes.

2.2. Microscopic problem

The microscopic BVP is defined on a representative volume element (RVE) V whose boundary is ∂V. By maintaining the small strain assumptions and neglecting the effect of body forces, the stress equilibrium equation and the strain-displacement relationship are respectively given as

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = \mathbf{0} \text{ and } \boldsymbol{\varepsilon} = \boldsymbol{\nabla} \otimes^s \mathbf{u} \text{ on } V\,. \tag{3}$$

In order to find the microscopic displacement field $\mathbf{u}$, constitutive models also have to be introduced. Assuming that the RVE consists of P constituents

$$V = \bigcup_{0}^{P-1} V^p \text{ with } V^p \text{ being the volume occupied by the constituent } p\,, \tag{4}$$

their volume fractions are given by

$$\omega^0 = \frac{V^0}{V}, \dots, \text{ and } \omega^{P-1} = \frac{V^{P-1}}{V}, \text{ satisfying } \sum_{p=0}^{P-1} \omega^p = 1\,. \tag{5}$$

The mechanical behavior of each constituent is explicitly provided by

$$\boldsymbol{\sigma}(t) = \boldsymbol{\Sigma}^p\left(\boldsymbol{\varepsilon}(t), \mathbf{q}(\tau \leq t)\right) \text{ with } p = 0, \dots, P-1\,, \tag{6}$$

in which $\mathbf{q}$ is a vector of thermodynamics internal variables defined to account for the history dependency, *e.g.* the plastic strain tensor and equivalent plastic strain in the context of J2-plasticity. It is noted that one of the physical phases can be voids, in which case the stress tensor is equal to $\mathbf{0}$ for all strain states.

2.3. Scale transition

A microscopic boundary condition must be introduced to conduct the deformation state over the RVE through a downscaling process of $\bar{\boldsymbol{\varepsilon}} = \langle \boldsymbol{\varepsilon} \rangle_V$, where $\langle \bullet \rangle_V$ denotes the volume average of the quantity $\bullet$ over a volume V

$$\langle \bullet \rangle_V = \frac{1}{V} \int_V \bullet \, dV\,. \tag{7}$$

The macroscopic stress $\bar{\boldsymbol{\sigma}} = \langle \boldsymbol{\sigma} \rangle_V$ and the corresponding tangent operator $\bar{\mathbf{C}}$ are obtained through an upscaling process, see Appendix A for details. From the computational point of view, the microscopic model acts as a usual constitutive law. The stress-strain relationship is always available through the resolutions of the microscopic BVP although its explicit form cannot generally be achieved. However, a FE^2 simulation requires performing a large number of the iterative resolutions of the microscopic BVP (at all macro-scale Gauss points and for all macroscopic Newton-Raphson iterations), leading to an intractable computation time and memory. In order to speed up the multiscale analysis, this work considers the material networks as surrogates to replace the costly microscopic models.

3. Interaction-based material network as surrogate of the microscopic problem

On the one hand, the microscopic problem described in Section 2.2 results in an implicit stress-strain response, which can be written under a functional form as

$$\begin{bmatrix} \bar{\boldsymbol{\sigma}}(t) \\ \bar{\boldsymbol{z}}(t) \end{bmatrix} = \boldsymbol{\mathcal{L}}^{\text{ch}}\left(\bar{\boldsymbol{\varepsilon}}(t)\,; \bar{\mathbf{q}}(\tau) \text{ with } 0 \leq \tau \leq t, \mathcal{G}, \mathcal{P}\right)\,, \tag{8}$$

where $\boldsymbol{\mathcal{L}}^{\text{ch}}$ denotes a multiple-input and multiple-output function, $\mathcal{G}$ is the RVE geometrical descriptor, and $\mathcal{P}$ is the material descriptor of the P constituents of the microstructure. The outputs $\bar{\boldsymbol{z}}(t)$ corresponds to properties of interest of the microscopic problem such as the microstructure evolution, which are thus not inputs of the micromechanical model. In Eq. (8), the vector of internal variables $\bar{\mathbf{q}}$ corresponds to the set of internal variables defining the microscopic problem, *i.e.* the thermodynamics internal variables of the local material law at the different integration points of its finite element discretization. Alternatively, Eq. (8) can be rewritten as

$$\begin{bmatrix} \bar{\boldsymbol{\sigma}}(t) \\ \bar{\boldsymbol{z}}(t) \end{bmatrix} = \boldsymbol{\mathcal{L}}^{\text{ch}}\left(\bar{\boldsymbol{\varepsilon}}(\tau) \text{ with } 0 \leq \tau \leq t, \mathcal{G}, \mathcal{P}\right)\,, \tag{9}$$

since the internal variables evolution $\bar{\mathbf{q}}(t)$ depend directly from the strain history $\bar{\boldsymbol{\varepsilon}}(\tau)$ with $0 \leq \tau \leq t$.

On the other hand, its surrogate $\boldsymbol{\mathcal{L}}^{\mathrm{sg}}$ can be expressed as

$$\begin{bmatrix} \bar{\boldsymbol{\sigma}}(t) \\ \bar{\boldsymbol{z}}(t) \end{bmatrix} = \boldsymbol{\mathcal{L}}^{\mathrm{sg}}\left(\bar{\boldsymbol{\varepsilon}}(t); \bar{\mathbf{q}}(\tau) \text{ with } 0 \leq \tau \leq t, \mathcal{P}, \mathbf{X}(\mathcal{G})\right), \tag{10}$$

where $\mathbf{X}$ is a vector consisting of the fitting parameters, which represents the RVE geometrical descriptor. In the surrogate model (10), the nature of the vector of the thermodynamics internal variables $\bar{\mathbf{q}}$ will be discussed in the next section. These fitting parameters do not depend on the material descriptor $\mathcal{P}$, as it will be shown in Section 5.6 which demonstrates the ability of the trained material network to extrapolate beyond the original material law used for the training. Alternatively, Eq. (10) can be rewritten as

$$\begin{bmatrix} \bar{\boldsymbol{\sigma}}(t) \\ \bar{\boldsymbol{z}}(t) \end{bmatrix} = \boldsymbol{\mathcal{L}}^{\mathrm{sg}}\left(\bar{\boldsymbol{\varepsilon}}(\tau) \text{ with } 0 \leq \tau \leq t, \mathcal{P}, \mathbf{X}(\mathcal{G})\right). \tag{11}$$

In the following sections, the theory of the interaction-based MN is first recalled. It is noted that the general form of the interaction-based MN does not rely on any micro-mechanical model. The fitting parameters $\mathbf{X}$ are then summarized. Since there exist constraints between the fitting parameters, a constrained optimization problem needs to be considered for the training process. A direct constraint elimination strategy is thus proposed to eliminate these constraints since an unconstrained optimization problem facilitates the training process when using a gradient descent-based optimizer. Finally, the training process is detailed to tune the fitting parameters such that the discrepancy between the predictions of the surrogate $\boldsymbol{\mathcal{L}}^{\mathrm{sg}}$ and the ones of the microscopic model $\boldsymbol{\mathcal{L}}^{\mathrm{ch}}$ is small.

3.1. Interaction-based material network

The interaction-based material network derived in [35] is briefly recalled. Let us consider a material network, denoted by $\mathcal{M}$, consisting of N discrete material nodes indexed from 0 to $N-1$ as

$$\mathcal{K} = [0, 1, \ldots, N-1], \tag{12}$$

where $\mathcal{K}$ is its list of indices. An arbitrary material node $i \in \mathcal{K}$ possesses

- a positive weight W^i in order to quantify its contribution to the network; and
- a physical state consisting of the strain $\boldsymbol{\varepsilon}^i$, the stress $\boldsymbol{\sigma}^i$, and internal state $\mathbf{q}^i$ governed by a constitutive law.

The vector of internal variables $\bar{\mathbf{q}}$ of the surrogate model (10) thus corresponds to the set of vectors of thermodynamics internal variables used in the material nodes constitutive law, *i.e.* $\bar{\mathbf{q}} = [\mathbf{q}^{i^T} \quad i = 0, \ldots, N-1]^T$, *i.e.* the plastic strain tensor and equivalent plastic strain in the context of J2-plasticity.. We consider a microscopic model defined on the RVE V described in Section 2.2, in which the material network $\mathcal{K}$ corresponds to a decomposition of the volume V into N non-overlapping parts:

$$V = \bigcup_{i=0}^{N-1} V^i \text{ satisfying } V^i \cap V^j = \emptyset \, \forall i \neq j, \text{ and } \frac{V^i}{V} = \frac{W^i}{\sum_{c \in \mathcal{K}} W^c}. \tag{13}$$

Each part $V^i \subset V$ corresponds to the material node i in a homogenization sense in which the strain average, Hill-Mandel and stress average relations arising from the scale transition, see Appendix A, are postulated as

$$\boldsymbol{\varepsilon}^i = \langle \boldsymbol{\varepsilon} \rangle_{V^i} \,, \tag{14}$$

$$\boldsymbol{\sigma}^i : \delta\boldsymbol{\varepsilon}^i = \langle \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon} \rangle_{V^i} \,, \text{ and} \tag{15}$$

$$\boldsymbol{\sigma}^i = \langle \boldsymbol{\sigma} \rangle_{V^i} \text{ with } i = 0, \dots, N-1 \,, \tag{16}$$

where $\langle \bullet \rangle_{V^i} = \frac{1}{V^i} \int_{V^i} \bullet \, dV$ denotes the averaging operator over V^i. As a result of Eqs. (13 - 16), the strain average, Hill-Mandel and stress average relations arising from the scale transition, see Appendix A, can be rewritten under the discrete forms as

$$\frac{1}{\sum_{c \in \mathcal{K}} W^c} \sum_{i \in \mathcal{K}} W^i \boldsymbol{\varepsilon}^i = \bar{\boldsymbol{\varepsilon}} \,, \tag{17}$$

$$\frac{1}{\sum_{c \in \mathcal{K}} W^c} \sum_{i \in \mathcal{K}} W^i \boldsymbol{\sigma}^i : \delta\boldsymbol{\varepsilon}^i = \bar{\boldsymbol{\sigma}} : \delta\bar{\boldsymbol{\varepsilon}} \,, \text{ and} \tag{18}$$

$$\frac{1}{\sum_{c \in \mathcal{K}} W^c} \sum_{i \in \mathcal{K}} W^i \boldsymbol{\sigma}^i = \bar{\boldsymbol{\sigma}} \,. \tag{19}$$

The behavior of each material node is governed by a material law specified in Eq. (6) as

$$\boldsymbol{\sigma}^i(t) = \boldsymbol{\Sigma}^{p_i} \left(\boldsymbol{\varepsilon}^i(t), \mathbf{q}^i(\tau \leq t) \right) \,, \tag{20}$$

where p_i denotes the constituent index of the material node i among the P physical constituents. Moreover, one can define the phase decomposition of the list of indices $\mathcal{K}$

$$\mathcal{K} = \bigcup_{p=0}^{P-1} \mathcal{K}^p \text{ satisfying } \mathcal{K}^p \cap \mathcal{K}^q = \emptyset \; \forall p \neq q \,, \tag{21}$$

in which $\mathcal{K}^p$ includes all material nodes governed by the material law indexed by p.

Equations (17 - 19) govern the response of the material network with $\boldsymbol{\varepsilon}^0, \dots, \boldsymbol{\varepsilon}^{N-1}$ being unknowns. To tie the material nodes together, network interactions are defined through a so-called interaction mapping $\mathcal{I}$, which is defined by

$$\begin{aligned} \mathcal{I} : \left(\mathbf{a}^0, \dots, \mathbf{a}^{M-1} \right) &\rightarrow \left(\boldsymbol{\varepsilon}^0, \dots, \boldsymbol{\varepsilon}^{N-1} \right) \\ &: \quad \boldsymbol{\varepsilon}^i = \bar{\boldsymbol{\varepsilon}} + \sum_{j=0}^{M-1} \alpha^{i,j} \mathbf{a}^j \otimes^s \mathbf{n}^j \text{ with } i = 0, \dots, N-1 \,. \end{aligned} \tag{22}$$

In the last equation, the following ingredients are considered:

- M is the number of interaction mechanisms;
- Each mechanism indexed by j with $j = 0, \dots, M-1$ is characterized by:
  - the interaction coefficients $\boldsymbol{\alpha}^j$, which are defined by the vector

$$\boldsymbol{\alpha}^j = \left[\alpha^{0,j} \; \dots \; \alpha^{N-1,j} \right]^T \,; \tag{23}$$

  - the interaction direction $\mathbf{n}^j$, which is a unit vector; and

– the incompatible vector $\mathbf{a}^j$, which is a new unknown vector .

- The M pairs $\left(\boldsymbol{\alpha}^j, \mathbf{n}^j\right)$ are the fitting parameters assumed to be known, *e.g.* following an offline training stage, while the M vectors $\mathbf{a}^j$ with $j = 0, \ldots, M-1$ are the new unknowns substituting for the local strains $\boldsymbol{\varepsilon}^i$ with $i \in \mathcal{K}$.

Equation (22) is inspired from the strain averaging over a polygonal (polyhedral) domain in 2D (3D). Indeed, let us assume that a material node i corresponds to a polygonal (polyhedral) subdomain V_0^i of the microstructure. When downscaling, see Appendix A for details, the local strain inside the subdomain V_0^i can be decomposed into a homogenized part and a fluctuation part as

$$\boldsymbol{\varepsilon} = \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} \otimes^s \mathbf{w} \,, \tag{24}$$

where $\mathbf{w}$ is the local fluctuation field, which is present due to the local inhomogeneities. The strain in the material node i corresponds to the homogenized strain over the subdomain V_0^i, leading to

$$\boldsymbol{\varepsilon}^i = \bar{\boldsymbol{\varepsilon}} + \frac{1}{V_0^i} \int_{V_0^i} \boldsymbol{\nabla} \otimes^s \mathbf{w} \, dV \,. \tag{25}$$

The boundary ∂V_0^i of the polygonal (polyhedral) domain V_0^i can be decomposed into M_i straight (planar) parts, *i.e.* $\partial V_0^i = \cup_{j=0}^{M_i-1} \Gamma_0^{i,j}$ so that each part $\Gamma_0^{i,j}$ is characterized by its outward unit normal vector $\mathbf{n}^{i,j}$. Using the Gauss theorem, Eq. (25) can be rewritten as

$$\boldsymbol{\varepsilon}^i = \bar{\boldsymbol{\varepsilon}} + \sum_{j=0}^{M_i-1} \frac{\Gamma_0^{i,j}}{V_0^i} \mathbf{w}^{i,j} \otimes^s \mathbf{n}^{i,j} \,, \tag{26}$$

$$\text{with } \mathbf{w}^{i,j} = \frac{1}{\Gamma_0^{i,j}} \int_{\Gamma_0^{i,j}} \mathbf{w} \, dA \,, \tag{27}$$

It is found that Eq. (26) takes the same form as Eq. (22), in which the meanings of $\alpha^{i,j}$, $\mathbf{a}^j$, and $\mathbf{n}^j$ are elucidated. To obtain such a polygonal (polyhedral) subdomain for each material node, we assume that there exists a list of M pairs $(\mathbf{a}^j, \mathbf{n}^j)$ with $j = 0, \ldots, M-1$. In Eq. (26), one can define $\mathbf{n}^k$ such that $\mathbf{n}^{i,j} = \mathbf{n}^k$ or $\mathbf{n}^{i,j} = -\mathbf{n}^k$, implying that the value of $\alpha^{i,j}$ with $i \in \mathcal{K}$ in Eq. (22) does not need to be positive but can be negative or equal to 0. Consequently, the form (22) is proposed, in which $\alpha^{i,j}$ with $i \in \mathcal{K}$ and $\mathbf{n}^j$ are the fitting parameters that have to be inferred using a training process.

An interaction mechanism does not have to involve all material nodes in the network. If a material node i does not contribute to the interaction mechanism j, then the coefficient $\alpha^{i,j} = 0$. As a result, one can define a compacted form of $\boldsymbol{\alpha}^j$ specified by Eq. (23) for the mechanism j as

$$\tilde{\boldsymbol{\alpha}}^j = \left[\alpha^{i,j} \text{ for } i \in \mathcal{V}^j\right]^T \,, \tag{28}$$

where $\mathcal{V}^j$ is the ordered list of indices of all material nodes present in the interaction mechanism j. One has clearly

$$\alpha^{i,j} = 0 \; \forall i \notin \mathcal{V}^j \text{ with } j = 0, \ldots, M-1 \,. \tag{29}$$

Similarly, one can define the list of active interactions for a material node i, denoted by $\mathcal{T}^i$, consisting of all interaction mechanisms in which the material node i contributes. One has the following

properties

$$i \notin \mathcal{V}^j \equiv j \notin \mathcal{T}^i \text{ and } i \in \mathcal{V}^j \equiv j \in \mathcal{T}^i \,. \tag{30}$$

One can define the architecture $\boldsymbol{\mathcal{A}}$ of the interaction-based MN as a list of all the ordered lists of indices of all material nodes present in each interaction mechanism as

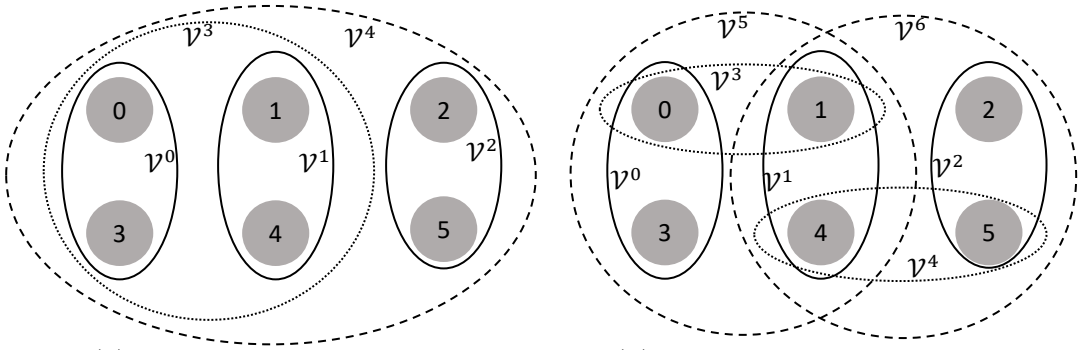
$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \mathcal{V}^0 & \mathcal{V}^1 & \dots & \mathcal{V}^{M-1} \end{bmatrix} . \tag{31}$$

It is noted that the architecture must be chosen before performing the offline training to identify the fitting parameters. There does not exist any assumption on the network architecture and the resolution strategy considered in this section can be performed as long as the interaction mapping (22) is known.

With the material nodes given in a list of indices $\mathcal{K}$ in Eq. (12), one can define a list $\mathbb{P}$ consisting of all sub-lists[2] of $\mathcal{K}$ except empty list and sub-lists with only 1 element, as the following

$$\mathbb{P} = [\mathcal{C} | \mathcal{C} \subseteq \mathcal{K} \text{ and size}\,(\mathcal{C}) \geq 2] \,, \tag{32}$$

where size $(\mathcal{C})$ denotes the number of elements in $\mathcal{C}$. Generally, an architecture $\boldsymbol{\mathcal{A}}$ of an interaction-based MN specified by Eq. (31) is a sub-list of $\mathbb{P}$. An interaction-based material network can form either a hierarchical or a non-hierarchical architecture, see Fig. 1 for examples of a material network of 6 material nodes within a hierarchical and a non-hierarchical architecture. Finding a network architecture for a given microstructure should rely on an optimization problem stated as finding $\boldsymbol{\mathcal{A}} \subset \mathbb{P}$ to obtain the best outcome following a suitable criterion. Such an optimization problem is beyond the scope of this paper and will be considered in a future work. In the presented applications, we consider hierarchical architectures for porous materials as detailed in Section 5.2.



(a) Hierarchical architecture (b) Non-hierarchical architecture

Figure 1: Interaction-based material network constructed from 6 material nodes: (a) hierarchical architecture $\boldsymbol{\mathcal{A}} = \begin{bmatrix} \mathcal{V}^0 & \mathcal{V}^1 & \dots & \mathcal{V}^4 \end{bmatrix}$ and (b) non-hierarchical architecture $\boldsymbol{\mathcal{A}} = \begin{bmatrix} \mathcal{V}^0 & \mathcal{V}^1 & \dots & \mathcal{V}^6 \end{bmatrix}$. Each interaction is formed with material nodes inside each contour.

Equation (22) can be rewritten in a more convenient form

$$\boldsymbol{\varepsilon}^i = \bar{\boldsymbol{\varepsilon}} + \sum_{j \in \mathcal{T}^i} \alpha^{i,j} \mathbf{a}^j \otimes^s \mathbf{n}^j \text{ with } i \in \mathcal{K} \,. \tag{33}$$

[2]A list of indices $\mathcal{C}$ is a sub-list of $\mathcal{K}$, denoted by $\mathcal{C} \subseteq \mathcal{K}$, if $\mathcal{C}$ is an empty-list or $\mathcal{K}$ contains all components of $\mathcal{C}$. Two sub-lists $\mathcal{C}^0$ and $\mathcal{C}^1$ are equal if and only if $\mathcal{C}^0 \subseteq \mathcal{C}^1$ and $\mathcal{C}^1 \subseteq \mathcal{C}^0$.

Using the interaction mapping described in Eq. (22) and Eq. (16), Eqs. (17, 18) can be respectively rewritten as

$$\sum_{j=0}^{M-1} \left(\sum_{i \in \mathcal{V}^j} W^i \alpha^{i,j} \right) \mathbf{a}^j \otimes^s \mathbf{n}^j = \mathbf{0}\,, \text{ and} \tag{34}$$

$$\sum_{j=0}^{M-1} \left[\left(\sum_{i \in \mathcal{V}^j} W^i \boldsymbol{\sigma}^i \alpha^{i,j} \right) \cdot \mathbf{n}^j \right] \cdot \delta \mathbf{a}^j = \mathbf{0}\,. \tag{35}$$

Eqs. (34, 35) govern the response of the material network and respectively represent the kinematic constraints and the weak form, from which the solution of this material network can be derived. It is noted that Eq. (34) is equivalent to the constraint on the fluctuation field arising from the scale transition theory, see Appendix A (in particular Eq. (A.4)), which has to be satisfied using a suitable boundary condition in the context of the computational homogenization, *e.g.* the periodic boundary condition. Within the context of the material networks, such a boundary condition does not exist. Instead, to fulfill *a priori* Eq. (34), an obvious choice follows

$$\sum_{i \in \mathcal{V}^j} W^i \alpha^{i,j} = 0 \text{ for } j = 0\,, \dots\,, M-1\,. \tag{36}$$

In [35], for each interaction mechanism j, $\alpha^{i,j}$ for $i \in \mathcal{V}^j$ were chosen as a function of the weights of the material nodes inside this mechanism in order to satisfy Eq. (36). However, this way of constructing the interaction mechanisms is equivalent to putting the material nodes within multiple-phase laminates as proposed in [32]. In order to be more general, $\alpha^{i,j}$ and W^i for $i \in \mathcal{V}^j$ for the mechanism j are considered as being independent and Eq. (36) as a constraint. Consequently, $\alpha^{i,j}$ and W^i for $i \in \mathcal{V}^j$ for the mechanism j need to be inferred during the training stage during which Eq. (36) must be enforced. The presented formalism is a general framework that can handle arbitrary network architectures as long as satisfying the condition (36). Therefore a material network based on laminate building blocks can be still considered, but with a lower number of trainable parameters to be inferred during the training stage.

Since Eq. (34) is satisfied independently to the values of $\mathbf{a}^j$, one can choose $\mathbf{a}^0, \mathbf{a}^1, \dots, \mathbf{a}^{M-1}$ being independent. As a result, Eq. (35) results in the following system of M equations

$$\left(\sum_{i \in \mathcal{V}^j} W^i \boldsymbol{\sigma}^i \alpha^{i,j} \right) \cdot \mathbf{n}^j = \mathbf{0} \text{ for } j = 0\,, \dots\,, M-1\,. \tag{37}$$

It is noted that the resolution of Eq. (37) automatically fulfills the Hill-Mandel condition stated by Eq. (18). The homogenized stress $\bar{\boldsymbol{\sigma}}$ is computed using Eq. (19) as an observable quantity.

To summarize, the homogenized stress $\bar{\boldsymbol{\sigma}}$ is estimated from $\bar{\boldsymbol{\varepsilon}}$ in the context of the material network by the following procedure:

- A material network of N material nodes is built with their weights W^i and M interaction mechanisms characterized by $\alpha^{i,j}$ and $\mathbf{n}^j$ with $i \in \mathcal{V}^j$ and $j = 0\,, \dots\,, M-1$. These fitting parameters are identified through an offline training stage.

- With the unknowns $\mathbf{a}^j$, $j = 0\,, \dots\,, M-1$, the evaluation of an interaction-based material network follows an iterative resolution as detailed in [35]. This resolution procedure is detailed in Appendix C and summarized by the following steps:

- Downscaling: the homogenized strain tensor $\bar{\boldsymbol{\varepsilon}}$ is downscaled to all the material nodes through the network interactions specified by Eq. (22), leading to the local strain at each material node, which depends not only on $\bar{\boldsymbol{\varepsilon}}$ but also on the unknowns. The local constitutive law of each material node (20) is used to estimate the local stresses and tangent operators, in which the nonlinearities could be present.
- Nonlinear system resolution: the system of equations (37) is iteratively solved to find the unknowns following the methodology detailed in Appendix C.
- Upscaling: the homogenized stress tensor $\bar{\boldsymbol{\sigma}}$ is computed by Eq. (19). When this material network is used within a macro-scale finite element simulation as a constitutive law, the homogenized tangent operator $\bar{\mathbf{C}} = \dfrac{\partial \bar{\boldsymbol{\sigma}}}{\partial \bar{\boldsymbol{\varepsilon}}}$ needs to be estimated in order to perform the Newton-Raphson iterations.

- Several quantities can be extracted from the material network through a volume averaging operation. For an arbitrary quantity $\bullet$ defined at the local physical state of the material nodes, a homogenized operator over $\mathcal{C}$, denoted by $\mathcal{H}_{\mathcal{C}}(\bullet)$ with $\mathcal{C} \subseteq \mathcal{K}$, is defined by

$$\mathcal{H}_{\mathcal{C}}(\bullet) \;=\; \frac{1}{\sum_{i\in\mathcal{C}} W^i} \sum_{i\in\mathcal{C}} W^i \bullet^i \,. \tag{38}$$

  It is noted that $\mathcal{H}_{\mathcal{K}}(\bullet)$ and $\mathcal{H}_{\mathcal{K}^p}(\bullet)$ are respectively equivalent to the operators $\langle\bullet\rangle_V$ and $\langle\bullet\rangle_{V^p}$ considered in the microscopic model, where V^p is the volume occupied by the constituent p. One has clearly $\bar{\boldsymbol{\varepsilon}} = \mathcal{H}_{\mathcal{K}}(\boldsymbol{\varepsilon})$ and $\bar{\boldsymbol{\sigma}} = \mathcal{H}_{\mathcal{K}}(\boldsymbol{\sigma})$, where $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ denote respectively the strain and stress measures.

3.2. Thermodynamic consistency of the interaction-based material networks

In this section, we show that interaction-based material networks respect the first and second laws of thermodynamics. As a result, thermodynamically consistent predictions are obtained. Since a pure mechanical problem is considered, these two laws are written in the context of an isothermal process.

The first law of thermodynamics states the energy conservation. For an interaction-based material network constructed with a list of indices $\mathcal{K}$ under an isothermal condition, the specific macroscopic mechanical power (per unit volume), which is given by $\bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}}$, is applied on the material network. The energy conservation through the scales implies the following condition

$$\bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}} = \frac{1}{\sum_{i\in\mathcal{K}} W^i} \sum_{i\in\mathcal{K}} W^i \boldsymbol{\sigma}^i : \dot{\boldsymbol{\varepsilon}}^i \,, \tag{39}$$

where the right-hand-side term corresponds to the specific mechanical power over the material network. The interaction-based material networks always respect the first law of thermodynamics since Eq. (39) corresponds to Eq. (18) in its incremental form, which is a constraint used when inferring the material network fitting parameters.

The second law of thermodynamics leads to the well-known Clausius-Duhem inequality establishing the positiveness of the dissipation rate. Under isothermal condition, the Clausius-Duhem inequality reads

$$\dot{\bar{\mathcal{D}}} = \bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}} - \dot{\bar{\Psi}} \geq 0 \,, \tag{40}$$

where $\dot{\bar{\mathcal{D}}}$ denotes the specific dissipation rate, $\bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}}$ is the specific mechanical power, $\bar{\Psi}$ is the specific reversible energy. Considering the material network, the specific reversible energy is defined[3] by the following form

$$\dot{\bar{\Psi}} = \frac{1}{\sum_{i\in\mathcal{K}} W^i} \sum_{i\in\mathcal{K}} W^i \dot{\Psi}^i \,, \tag{41}$$

where $\dot{\Psi}^i$ is the local specific reversible energy at the material node i. Using Eqs. (39, 41), Eq. (40) becomes

$$\dot{\bar{\mathcal{D}}} = \frac{1}{\sum_{i\in\mathcal{K}} W^i} \sum_{i\in\mathcal{K}} W^i \left(\boldsymbol{\sigma}^i : \dot{\boldsymbol{\varepsilon}}^i - \dot{\Psi}^i \right) \geq 0 \,. \tag{42}$$

In the last equation, the term $\boldsymbol{\sigma}^i : \dot{\boldsymbol{\varepsilon}}^i - \dot{\Psi}^i$ corresponds to the local dissipation rate at the material node i. In general, the local material law governing the mechanical behavior of this material node is thermodynamically consistent, *i.e.* $\mathcal{D}^i = \boldsymbol{\sigma}^i : \dot{\boldsymbol{\varepsilon}}^i - \dot{\Psi}^i \geq 0$. As a result the second law of thermodynamics is always respected in the interaction-based material networks since the condition $\dot{\bar{\mathcal{D}}} \geq 0$ is always satisfied.

3.3. Fitting parameters

Let us consider $\boldsymbol{\mathcal{L}}^{\text{sg}}$ expressed by Eq. (11) being the functional form of the response of a material network described in the previous section. To make this model as a surrogate of the full-field model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ expressed by Eq. (9), the fitting parameters $\mathbf{X}$ must be identified. We assume that the network architecture $\boldsymbol{\mathcal{A}}$ following Eq. (31) has already been designed. A specific form of $\boldsymbol{\mathcal{A}}$ is employed in the numerical applications in Section 5.

3.3.1. Fitting parameters with their constraints

Within the context of the material network described in the previous section, there exist three following groups of the fitting parameters, each one with its constraints:

- The first group, denoted by $\mathbf{W}$ as a column vector, consists of all the weights of the material nodes

$$\mathbf{W} = \left[W^i \ \forall i \in \mathcal{K} \right]^T \,. \tag{43}$$

  These weights must satisfy the following constraints:

  - Each weight must be positive, *i.e.*

$$W^i > 0 \ \forall i \in \mathcal{K} \,. \tag{44}$$

  - The initial fraction of each constituent in the full-field model must be recovered at the end of the training process, *i.e.*

$$\omega^p = \frac{\sum_{i\in\mathcal{K}^p} W^i}{\sum_{i\in\mathcal{K}} W^i} \text{ for } p = 0, \ldots, P-1 \,, \tag{45}$$

[3] We note that the scale transition does not enforce the equivalence of elastic energy rate but of the total energy rate through the Hill-Mandel condition, here recasted in the form (39).

where ω^p is the fraction of the p^{th} constituent in the microscopic model as mentioned in Eq. (5).

- The second group, denoted by $\mathbf{A}$ as a column vector, consists of all the interaction coefficients

$$\mathbf{A} = \left[\left(\tilde{\boldsymbol{\alpha}}^j \right)^T \text{ for } j = 0 \, , \ldots \, , M-1 \right]^T \, , \tag{46}$$

where $\tilde{\boldsymbol{\alpha}}^j$ is given by Eq. (28) as a column vector. Each interaction mechanism $\tilde{\boldsymbol{\alpha}}^j$ must satisfy Eq. (36), which can be rewritten by

$$\left(\tilde{\boldsymbol{\alpha}}^j \right)^T \left[W^i \text{ for } i \in \mathcal{V}^j \right]^T = 0 \, . \tag{47}$$

- The third group, denoted by $\mathbf{N}$ as a column vector, consists of all the interaction directions

$$\mathbf{N} = \left[\left(\mathbf{n}^j \right)^T \text{ for } j = 0 \, , \ldots \, , M-1 \right]^T \, , \tag{48}$$

where a unit vector $\mathbf{n}^j$ is considered in Eq. (22), implying the following constraints

$$\mathbf{n}^j \cdot \mathbf{n}^j = 1 \text{ for } j = 0 \, , \ldots \, , M-1 \, . \tag{49}$$

The constraints given by Eqs. (44, 45, 47, 49) need to be enforced during the training process, implying that a constrained optimization problem needs to be considered. These constraints should be eliminated since an unconstrained optimization problem facilitates the training process when using a gradient descent-based optimizer. For this purpose, *ad hoc* changes of variables are considered to satisfy *a priori* these constraints.

3.3.2. Unconstrained fitting parameters

In order to satisfy both Eqs. (44, 45), the following parameterization of the weights is considered

$$W^i = \omega^p \frac{f\left(Z^i\right)}{\sum_{l \in \mathcal{K}^p} f\left(Z^l\right)} \text{ for } i \in \mathcal{K}^p \text{ and } p = 0, \ldots, P-1 \, , \tag{50}$$

where Z^l with $l = 0 \, , \ldots \, N-1$ are the new un-constrained fitting parameters, $f\left(\bullet\right)$ is the activation function whose output is always positive, and ω^p is the fraction of the constituent p. This work considers the smoothed version of relu, abbreviated by "relu$^{\text{s}}\left(\bullet\right)$", as the activation function

$$f\left(\bullet\right) = \text{relu}^{\text{s}}\left(x\right) = \frac{1}{s} \ln\left(1 + \text{e}^{s\bullet}\right) \, , \tag{51}$$

where s is the sharpness. In the remaining of the paper, the value $s = 10$ is chosen as considered in [35]. As a result of the parameterization (50), Eqs. (44, 45) are always satisfied. The new fitting parameters for the weights are defined by the following column vector

$$\widetilde{\mathbf{W}} = \left[Z^i \text{ for } i \in \mathcal{K} \right]^T \, . \tag{52}$$

Equation (47) is enforced by a direct constraint elimination. Let us consider the interaction mechanism j whose list of indices of the material nodes is $\mathcal{V}^j$. Moreover, one can define the first element of this list by $\mathcal{V}_0^j$ and a location operator $\text{loc}\left(l, \mathcal{V}^j\right)$, which returns the location of the node

l in the list $\mathcal{V}^j$, *i.e.*

$$\mathcal{V}^j_{\mathrm{loc}(l,\mathcal{V}^j)} = l\,. \tag{53}$$

Eq. (47) is enforced by the following constraint elimination

$$\begin{cases} \alpha^{\mathcal{V}^j_0,j} \text{ is fixed to an arbitrarily non-zero value}\,, \text{ and} \\ \alpha^{i,j} = -\dfrac{W^{\mathcal{V}^j_0}\alpha^{\mathcal{V}^j_0,j}}{W^i}\dfrac{\beta^{\mathrm{loc}\left(i,\mathcal{V}^j\right)-1,j}}{\sum_{c=0}^{L_j-2}\beta^{c,j}} \text{ for } i \in \mathcal{V}^j \text{ and } \mathrm{loc}\left(i,\mathcal{V}^j\right) > 0\,, \end{cases} \tag{54}$$

where L_j is the number of material nodes in $\mathcal{V}^j$ and $\beta^{r,j}$ for $r = 0,\dots,L_j-2$ are new fitting parameters. These new fitting parameters for this interaction mechanism can be collected in a column vector

$$\boldsymbol{\beta}^j = \left[\beta^{r,j} \text{ for } r = 0,\dots,L_j-2\right]^T\,. \tag{55}$$

Finally, the new fitting parameters for all the interaction coefficients are given by the following column vector

$$\widetilde{\mathbf{A}} = \left[\left(\boldsymbol{\beta}^j\right)^T \text{ for } j = 0\,,\dots,M-1\right]^T\,. \tag{56}$$

The constraints (49) are enforced using an angle parameterization [35]. Since any unit vector can be expressed in terms of either the polar angle in a bidimensional state or the spherical angles in a general three-dimensional state, the vector $\mathbf{n}^j$ is rewritten by

$$\mathbf{n}^j = \begin{cases} \left[\cos\left(2\pi\varphi^j\right) \quad \sin\left(2\pi\varphi^j\right) \quad 0\right]^T & \text{in plane strain}\,, \\ \left[\cos\left(2\pi\varphi^j\right)\sin\left(\pi\theta^j\right) \quad \sin\left(2\pi\varphi^j\right)\sin\left(\pi\theta^j\right) \quad \cos\left(\pi\theta^j\right)\right]^T & \text{otherwise}\,, \end{cases} \tag{57}$$

with $\left[\varphi^j\ \theta^j\right] \in [0\ 1) \times [0\ 1]$ being normalized angles. The relationship above allows defining the operator angle $(\bullet)$ by

$$\mathrm{angle}\left(\mathbf{n}\right) = \begin{cases} [\varphi] & \text{in plane strain}\,, \\ [\varphi\ \theta]^T & \text{otherwise}\,, \end{cases} \tag{58}$$

for any arbitrary unit vector $\mathbf{n}$. As a result, the new fitting parameters for all the interaction interactions are given by the following column vector

$$\widetilde{\mathbf{N}} = \left[\mathbf{e}^{jT} \text{ for } j = 0\,,\dots,M-1\right]^T\,, \tag{59}$$

where the notation $\mathbf{e}^j = \mathrm{angle}\left(\mathbf{n}^j\right)$ is used.

3.3.3. Summary

From the known network architecture $\boldsymbol{\mathcal{A}}$, the fitting parameters were detailed. The use of Eqs. (50, 54, 57) allows Eqs. (44, 45, 47, 49) to be satisfied. The fitting parameters to be inferred are summarized in a column vector $\mathbf{X}$ as

$$\mathbf{X} = \left[\widetilde{\mathbf{W}}^T \quad \widetilde{\mathbf{A}}^T \quad \widetilde{\mathbf{N}}^T\right]^T\,. \tag{60}$$

A training stage needs to be performed to identify these fitting parameters by minimizing a loss function as detailed in Section 4.

4. Machine learning algorithm for tuning fitting parameters

The surrogate model $\boldsymbol{\mathcal{L}}^{\text{sg}}$ is an interaction-based MN, in which a number of fitting parameters $\mathbf{X}$ needs to be inferred such that the discrepancy between the predictions of the surrogate $\boldsymbol{\mathcal{L}}^{\text{sg}}$ and the ones of the microscopic model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ is small. This training process is detailed in this section. Assuming that the network architecture $\boldsymbol{\mathcal{A}}$ defined by Eq. (31) of $\boldsymbol{\mathcal{L}}^{\text{sg}}$ is known, we propose two different strategies to train $\boldsymbol{\mathcal{L}}^{\text{sg}}$:

- *Nonlinear training* (denoted by NT): We assume that the constitutive laws governing the physical phases of the microstructure are known. Since the response of $\boldsymbol{\mathcal{L}}^{\text{sg}}$ is path-dependent, *e.g.* due to plasticity, the training database is provided by nonlinear paths. For each loading path defined by a strain sequence, the response of $\boldsymbol{\mathcal{L}}^{\text{sg}}$ in terms of the stress sequence and of the microstructure evolution, *e.g.* porosity, are tracked. The fitting parameters of $\boldsymbol{\mathcal{L}}^{\text{sg}}$ can be found by minimizing the discrepancy between the predictions of $\boldsymbol{\mathcal{L}}^{\text{sg}}$ and the corresponding predictions of the full-field model $\boldsymbol{\mathcal{L}}^{\text{ch}}$. We note that the microstructure evolution $\bar{\boldsymbol{z}}$ and the stress sequence are outputs of the microscopic model (8) and of the surrogate model (10) but are not part of their thermodynamics internal variables $\bar{\mathbf{q}}$.

- *Linear elastic training* (denoted by LET): This method was employed in other works [28, 29, 32, 34, 35], in which linear elastic analyses are used. Each material node belongs to a physical phase whose behavior is assume to be linearly elastic and characterized by an elastic tangent operator. The corresponding homogenized response of the interaction-based MN can be expressed by the homogenized elastic tangent operator, which is a function of the elastic tangent operators of the underlying phases and of the fitting parameters. As a result, these fitting parameters result from an optimization procedure using an offline database consisting of the elastic tangent operators of the underlying phases as inputs and the corresponding homogenized elastic tangent operators as outputs. This training database can be obtained with the full-field model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ in which the elastic behaviors of the phases can be randomly and numerically generated. Since the interaction-based MN does not involve hierarchical multiple-phase laminates as employed in [28, 29, 32, 34, 35], a unified approach needs to be used to estimate its homogenized tangent operator in this work.

In the following sections, the nonlinear training procedure is first presented. We will show that this procedure can be easily adapted for the linear elastic training.

4.1. Nonlinear training

4.1.1. Sequential data structure

When employing history-dependent material behaviors, the response of the microscopic model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ and of the material network $\boldsymbol{\mathcal{L}}^{\text{sg}}$ must be provided in terms of sequential data. A sequence $\boldsymbol{\mathcal{B}}$ is defined by

$$\boldsymbol{\mathcal{B}} = \begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \dots & \mathbf{b}_{T-1} \end{bmatrix}, \tag{61}$$

where T is the length of the sequence $\boldsymbol{\mathcal{B}}$, *i.e.* $T = \text{len}(\boldsymbol{\mathcal{B}})$, and each component $\mathbf{b}_k$ can be scalars, vectors, second-order tensors, *etc.*

4.1.2. Sequential offline data

For the sake of readability, an output of the microscopic full-field model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ is represented with the notation $\widehat{\bullet}$ while the notation $\bar{\bullet}$ is used to refer to its counterpart obtained with the material network $\boldsymbol{\mathcal{L}}^{\text{sg}}$. During the training process, $\boldsymbol{\mathcal{L}}^{\text{ch}}$ and $\boldsymbol{\mathcal{L}}^{\text{sg}}$ are loaded with the same strain history given by the strain sequences

$$\widehat{\boldsymbol{\mathcal{E}}} = \bar{\boldsymbol{\mathcal{E}}} = \begin{bmatrix} \bar{\boldsymbol{\varepsilon}}_0 & \bar{\boldsymbol{\varepsilon}}_1 & \dots & \bar{\boldsymbol{\varepsilon}}_{T-1} \end{bmatrix} . \tag{62}$$

As a result, Eqs. (9, 11) can be rewritten as

$$\begin{bmatrix} \widehat{\boldsymbol{\mathcal{S}}} \\ \widehat{\mathcal{Z}}^0 \\ \vdots \\ \widehat{\mathcal{Z}}^{L_{\bar{\mathcal{Z}}}-1} \end{bmatrix} = \boldsymbol{\mathcal{L}}^{\text{ch}} \left(\widehat{\boldsymbol{\mathcal{E}}}, \mathcal{G}, \mathcal{P} \right) , \text{ and} \tag{63}$$

$$\begin{bmatrix} \bar{\boldsymbol{\mathcal{S}}} \\ \bar{\mathcal{Z}}^0 \\ \vdots \\ \bar{\mathcal{Z}}^{L_{\bar{\mathcal{Z}}}-1} \end{bmatrix} = \boldsymbol{\mathcal{L}}^{\text{sg}} \left(\bar{\boldsymbol{\mathcal{E}}}, \mathcal{G}, \mathcal{P}; \mathbf{X} \right) , \tag{64}$$

where $\widehat{\boldsymbol{\mathcal{E}}}$, $\widehat{\boldsymbol{\mathcal{S}}}$, and $\widehat{\mathcal{Z}}^q$ with $q = 0, \dots, L_{\bar{\mathcal{Z}}} - 1$ are respectively the sequences of the homogenized strain, the sequences of the homogenized stress, and the sequences $L_{\bar{\mathcal{Z}}}$ of properties of interest such as microstructure evolution, which are used or predicted by the microscopic model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ while $\bar{\boldsymbol{\mathcal{E}}}$, $\bar{\boldsymbol{\mathcal{S}}}$, and $\bar{\mathcal{Z}}^q$ with $q = 0, \dots, L_{\bar{\mathcal{Z}}} - 1$ are respectively their counterparts in the material network $\boldsymbol{\mathcal{L}}^{\text{sg}}$. Any quantity characterizing the microstructure evolution can be considered as a homogenized state, *e.g.* using the homogenization operator (38) for $\boldsymbol{\mathcal{L}}^{\text{sg}}$ and the operator (7) as its counterpart for $\boldsymbol{\mathcal{L}}^{\text{ch}}$, to define the properties of interest z^q. Since $\boldsymbol{\mathcal{L}}^{\text{sg}}$ does not possesses any microstructure, $\mathcal{K}^p$ consisting of all material nodes of the phase indexed by p and $\mathcal{K}$ are considered in the homogenization operator (38) for $\boldsymbol{\mathcal{L}}^{\text{sg}}$ while V^p and V are respectively considered in the homogenization operator (7) for $\boldsymbol{\mathcal{L}}^{\text{ch}}$. In the numerical examples with porous materials in Section 5, the homogenized porosity is investigated. When using $L_{\bar{\mathcal{Z}}} = 0$, the training process is considered without any homogenized properties of interest.

Training the surrogate model $\boldsymbol{\mathcal{L}}^{\text{sg}}$ for a given RVE geometrical descriptor $\mathcal{G}$ and a given material descriptor $\mathcal{P}$ of P constituents requires collecting a set of sequential strain paths and their corresponding sequential responses from the microscopic model $\boldsymbol{\mathcal{L}}^{\text{ch}}$. For this purpose, a training dataset $\mathbb{D}_{\text{train}}$ of N_{train} samples of strain paths is first generated

$$\widehat{\boldsymbol{\mathcal{E}}}^{[0]}, \dots, \widehat{\boldsymbol{\mathcal{E}}}^{[N_{\text{train}}-1]} . \tag{65}$$

Then the corresponding outputs of $\boldsymbol{\mathcal{L}}^{\text{ch}}$ are collected as

$$\widehat{\boldsymbol{\mathcal{S}}}^{[0]}, \dots, \widehat{\boldsymbol{\mathcal{S}}}^{[N_{\text{train}}-1]} \tag{66}$$

$$\widehat{\mathcal{Z}}^{q,[0]}, \dots, \widehat{\mathcal{Z}}^{q,[N_{\text{train}}-1]} \text{ with } q = 0 , \dots , L_{\bar{\mathcal{Z}}} - 1 . \tag{67}$$

To investigate overfitting, a validation dataset $\mathbb{D}_{\text{valid}}$ consisting of N_{valid} samples is also created.

4.1.3. Loss function and its gradient with respect to fitting parameters

The training stage requests the definition of the error between the predictions obtained by the microscopic model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ and the ones obtained by the material network $\boldsymbol{\mathcal{L}}^{\text{sg}}$. For an offline sample l, the corresponding error reads

$$C^{[l]}\left(\mathbf{X}\right) = \epsilon^{\text{rel}}\left(\widehat{\boldsymbol{\mathcal{S}}}^{[l]}, \bar{\boldsymbol{\mathcal{S}}}^{[l]}\right) + \sum_{q=0}^{L_{\bar{\mathcal{Z}}}-1} \lambda^q \epsilon^{\text{rel}}\left(\widehat{\bar{\mathcal{Z}}}^{q,[l]}, \bar{\bar{\mathcal{Z}}}^{q,[l]}\right), \tag{68}$$

where $\bar{\boldsymbol{\mathcal{S}}}^{[l]}$ and $\bar{\bar{\mathcal{Z}}}^{q,[l]}$ predicted by the material network are function of $\mathbf{X}$, λ^q are the penalty parameters introduced in order to efficiently combine different losses, and $\epsilon^{\text{rel}}\left(\boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{H}}\right)$ is the relative difference operator between two arbitrary sequences $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{H}}$ of the same length T as

$$\epsilon^{\text{rel}}\left(\boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{H}}\right) = \frac{1}{T}\sum_{k=0}^{T-1} \frac{||\mathbf{b}_k - \mathbf{h}_k||_2}{||\mathbf{b}_k||_2} \quad \text{with} \tag{69}$$

$$||\bullet||_2 = \begin{cases} |\bullet| & \text{if } \bullet \text{ is a scalar}, \\ \left[\sum_i \left(|\bullet_i|\right)^2\right]^{\frac{1}{2}} & \text{if } \bullet \text{ is a vector}, \text{ and} \\ \left[\sum_{i,j} \left(|\bullet_{ij}|\right)^2\right]^{\frac{1}{2}} & \text{if } \bullet \text{ is a second-order tensor}. \end{cases} \tag{70}$$

In practice $\lambda^q = 1\ \forall q$ are chosen in Eq. (68) since all errors are dimensionless. The gradient of the error (68) with respect to the tuning parameters $\mathbf{X}$ must be estimated in the context of the gradient descent algorithm. One has

$$\mathbf{g}^{[l]} = \frac{\partial C^{[l]}}{\partial \mathbf{X}}, \tag{71}$$

where the computation details can be found in Appendix D. Over N_{train} samples, the total loss function is computed as

$$C_{\text{train}} = \frac{1}{N_{\text{train}}} \sum_{l=0}^{N_{\text{train}}-1} C^{[l]}. \tag{72}$$

The validation error C_{valid} of the validation set is also estimated using Eq. (72).

For minimizing the difference between the predictions obtained by the microscopic model $\boldsymbol{\mathcal{L}}^{\text{ch}}$ and the ones obtained by the material network $\boldsymbol{\mathcal{L}}^{\text{sg}}$, optimization iterations for adjusting the tuning parameters $\mathbf{X}$ are performed using a stochastic gradient descent algorithm with adaptive moment estimation (Adam) [36], see Appendix E for details.

4.2. Linear elastic training

The training procedure detailed in the previous section is adapted for the linear elastic training in this section. First, a unified strategy for estimating the homogenized elastic tensor of an interaction-based MN is presented. Then the training procedure using only elastic data is detailed by defining a new cost function in the optimization iterations.

4.2.1. Linear elastic response of an interaction-based material network and its gradient with respect to fitting parameters

For the ease of implementation, the Voigt's notations of stress and strain tensors are considered with the order $\begin{bmatrix}00 & 11 & 22 & 01 & 02 & 12\end{bmatrix}$. As a result, the material tensor $\mathbf{C}$ can also be written in a matrix form $\text{mat}\,(\mathbf{C})$, see Appendix F. When a plane strain state is employed, one has its reduced 9-component form $\text{mat}^{2D}\,(\mathbf{C})$.

From P elastic tensors $\mathbf{C}^{(0)}, \ldots, \mathbf{C}^{(P-1)}$ governing the P phases, the corresponding elastic tensor $\bar{\mathbf{C}}$ of the homogenized material and its gradient with respect to the fitting parameters can be estimated in a unified way, which does not rely on any assumption about the architecture $\boldsymbol{\mathcal{A}}$ of the material network under consideration, see Appendix F.

In the following, the homogenized elastic matrix is defined as

$$\bar{\mathbf{L}} = \begin{cases} \text{mat}\left(\bar{\mathbf{C}}\right) & \text{for a general 3-dimensional strain state}\,, \\ \text{mat}^{2D}\left(\bar{\mathbf{C}}\right) & \text{for a plane strain state}\,. \end{cases} \tag{73}$$

As a result, the P elastic tensors $\mathbf{C}^{(0)}, \ldots, \mathbf{C}^{(P-1)}$ governing the P phases can be rewritten as $\mathbf{L}^{(0)}$, ..., $\mathbf{L}^{(p-1)}$ in their matrix forms and the estimation of the corresponding homogenization elastic matrix $\bar{\mathbf{L}}$ can be rewritten under a functional form

$$\bar{\mathbf{L}} = \bar{\mathbf{L}}\left(\mathbb{X}; \mathbf{X}\right)\,, \tag{74}$$

where

$$\mathbb{X} = \left[\mathbf{L}^{(0)}, \ldots, \mathbf{L}^{(P-1)}\right]\,. \tag{75}$$

The gradient of $\bar{\mathbf{L}}$ with respect to the fitting parameters, $\dfrac{\partial \bar{\mathbf{L}}}{\partial X_j}$ for each component X_j of $\mathbf{X}$, is also known following Appendix F.

4.2.2. Training procedure

In order to identify $\mathbf{X}$, a training dataset $\mathbb{D}_{\text{train}}$ of N_{train} samples of material tensors of the physical phases is first generated

$$\mathbb{X}^{[0]}, \ldots, \mathbb{X}^{[N_{\text{train}}-1]}\,. \tag{76}$$

Then the corresponding outputs of $\boldsymbol{\mathcal{L}}^{\text{ch}}$ are collected as

$$\hat{\mathbf{L}}^{[0]}, \ldots, \hat{\mathbf{L}}^{[N_{\text{train}}-1]}\,. \tag{77}$$

A validation dataset $\mathbb{D}_{\text{valid}}$ consisting of N_{valid} samples is also created to investigate overfitting.

The same optimization procedure as in the non-linear training case of Section 4.1, which is described in Appendix E, is used but with the following cost function

$$C^{[l]}\left(\mathbf{X}\right) = \frac{||\bar{\mathbf{L}}\left(\mathbb{X}^{[l]}; \mathbf{X}\right) - \hat{\mathbf{L}}^{[l]}||_{\text{F}}}{||\hat{\mathbf{L}}^{[l]}||_{\text{F}}}\,, \tag{78}$$

where $||\bullet||_{\text{F}}$ represents the Frobenius norm. The gradient of the error (68) with respect to the tuning parameters $\mathbf{X}$ must be estimated in the context of the gradient descent algorithm. The

corresponding gradient of the loss function (78) with respect to the fitting parameters reads

$$\mathbf{g}^{[l]} = \left(\frac{\bar{\mathbf{L}}\left(\mathbb{X}^{[l]};\mathbf{X}\right) - \hat{\mathbf{L}}^{[l]}}{||\bar{\mathbf{L}}\left(\mathbb{X}^{[l]};\mathbf{X}\right) - \hat{\mathbf{L}}^{[l]}||_{\mathrm{F}} \times ||\hat{\mathbf{L}}^{[l]}||_{\mathrm{F}}} \right) \circ \frac{\bar{\mathbf{L}}\left(\mathbb{X}^{[l]};\mathbf{X}\right)}{\partial \mathbf{X}} , \tag{79}$$

where $\frac{\bar{\mathbf{L}}\left(\mathbb{X}^{[l]};\mathbf{X}\right)}{\partial \mathbf{X}}$ is a block vector whose components are matrices with the dimension of $\bar{\mathbf{L}}$ and $\circ$ represents the element-wise product of a matrix with each component of the block vector.

5. Numerical examples: porous materials

In this section, a surrogate of the microscopic problem defined on a voided microstructure is constructed based on the interaction-based MN using the theory in Section 3 and the training procedure in Section 4. The contribution of voids in the interaction-based MN is accounted for by considering a network of material nodes of the solid phase and of the material nodes representing the voids. A void material node has the stress being equal to zero for any strain state and does not contain any internal state, while the strain at this node represents its volumetric change.

In the following, we consider the microscopic problem $\boldsymbol{\mathcal{L}}^{\mathrm{ch}}$ defined on a microstructure V consisting of a solid phase (V^0) and voids (V^1). The initial solid volume fraction is equal to ω^0 and the void volume fraction is equal to $\omega^1 = 1 - \omega^0$, in which the values of ω^0 and ω^1 are known. The solid phase is governed by a J_2 elastoplastic model, see Appendix G for details. The material parameters of the solid phase are reported in Tab. 1. Both the NT and LET procedures detailed in Section 4 are used to train the material network $\boldsymbol{\mathcal{L}}^{\mathrm{sg}}$ as surrogate of the microscopic model $\boldsymbol{\mathcal{L}}^{\mathrm{ch}}$. In the NT procedure, the training of $\boldsymbol{\mathcal{L}}^{\mathrm{ch}}$ is achieved by minimizing the error between not only the homogenized stresses $\bar{\boldsymbol{\sigma}} = \langle \boldsymbol{\sigma} \rangle_V$ following Eq. (19), but also between the microstructure evolution, which is characterized by the global porosity evolution. In the LET procedure, only the homogenized elastic behavior is employed. The online prediction of the trained material networks successively trained with these two strategies is investigated in both contexts of a virtual test and of FE^2 multiscale simulations.

Table 1: Elastoplastic parameters of the solid phase.

Young modulus (E [GPa])	3
Poisson ratio (ν [-])	0.3
Isotropic flow stress $\sigma_y(\gamma)$ [MPa] (γ - equivalent plastic strain)	$100 + 60\gamma$

5.1. Evaluation of the porosity evolution

Let us consider an infinitesimal volume $dV(x)$ at a material point $\mathbf{x}$ undergoing a strain $\boldsymbol{\varepsilon}$. The volumetric change of $dV(x)$ reads

$$\delta dV(x) = \mathrm{tr}\left(\boldsymbol{\varepsilon}(\mathbf{x})\right) dV(x) , \tag{80}$$

where $\mathrm{tr}(\bullet)$ is the trace operator of an arbitrary second-order tensor $\bullet$. Consequently, the volumetric change of the whole solid part can be estimated by

$$\Delta V^0 = \int_{V^0} \delta dV(x) = \int_{V^0} \mathrm{tr}(\boldsymbol{\varepsilon})\, dV = \langle \mathrm{tr}(\boldsymbol{\varepsilon}) \rangle_{V^0} V^0 , \tag{81}$$

where V^0 denotes the volume occupied by the solid part. Moreover, the volume V of the RVE undergoes the homogenized strain $\bar{\boldsymbol{\varepsilon}}$, leading to the change of the RVE volume

$$\Delta V = \operatorname{tr}(\bar{\boldsymbol{\varepsilon}})\, V\,. \tag{82}$$

As a result, the porosity evolution $\hat{\eta}$ can be extracted from the solution of the microscopic problem by

$$\hat{\eta} = 1 - \frac{V^0 + \Delta V^0}{V + \Delta V} = 1 - \frac{\langle 1 + \operatorname{tr}(\boldsymbol{\varepsilon})\rangle_{V^0}}{1 + \operatorname{tr}(\bar{\boldsymbol{\varepsilon}})}\omega^0\,. \tag{83}$$

Equivalently, the porosity evolution $\bar{\eta}$ can be obtained with the material network $\boldsymbol{\mathcal{L}}^{\mathrm{sg}}$ as

$$\bar{\eta} = 1 - \frac{\mathcal{H}_{\mathcal{K}^0}(1 + \operatorname{tr}(\boldsymbol{\varepsilon}))}{1 + \operatorname{tr}(\bar{\boldsymbol{\varepsilon}})}\omega^0\,, \tag{84}$$

where $\mathcal{K}^0$ is the list of indices of all material nodes of the solid phase and $\mathcal{H}_{\mathcal{K}^0}(1 + \operatorname{tr}(\boldsymbol{\varepsilon}))$ is defined in Eq. (38). As a result, one can consider $L_{\bar{\mathcal{Z}}} = 1$ in Eqs. (63, 64) and use

$$\begin{cases} \hat{z}^0 = \langle z^0 \rangle_{V^0} \\ \bar{z}^0 = \mathcal{H}_{\mathcal{K}^0}\left(z^0\right) \end{cases} \qquad \text{with } z^0 = 1 + \operatorname{tr}(\boldsymbol{\varepsilon})\,. \tag{85}$$

It is noted that the porosity evolution is not used in the linear elastic training, in which only the homogenized elastic tangent operator is considered. Once trained, the predictions with the interaction-based MN of both the homogenized stress and the porosity evolution are compared to the corresponding evaluations of the microscopic full-field model.

5.2. Interaction-based material network with a hierarchical architecture

The framework in Sections 3 and 4 is not limited to any particular architecture. However, the network architecture $\boldsymbol{\mathcal{A}}$ of the material network following Eq. (31) must be defined before performing the training and is not modified during the training. MNs based on laminate building blocks should not be used due to their limitations. Indeed, in a laminate consisting of void as a phase (so-called voided laminate), the stress state is constrained as the result of the stress balance in the lamination direction. Any combination of co-planar voided laminates will also yield zero-stress in the plane, which limits the training ability of the network as shown in Appendix B. In this work, we investigate only the following hierarchical architecture of N_{levels} levels, which are recursively constructed as follows:

- At level 0: $2^{N_{\mathrm{levels}}-1}$ separate interaction mechanisms of 4 material nodes are created. It is noted that an arbitrary number of material nodes larger than 2 in each interaction mechanism at this lowest level can be employed[4]. In this work, we consider 4 material nodes per interaction mechanism since it is the smallest number of material nodes that can be used under the assumption that the solid phase and voids play an equivalent role in the material network, *i.e.* the number of material nodes of the solid phase and the ones of the voids in each interaction mechanism are equal.

[4]The learning ability of the material network is limited with a number of 2 material nodes because of voids. Indeed, considering an interaction j of two material nodes: a solid node i and a void node k, Eq. (37) becomes $\boldsymbol{\sigma}^i \cdot \mathbf{n}^j = \mathbf{0}$ because $\boldsymbol{\sigma}^k = \mathbf{0}$. As a result, the stress state in the material node i is too restrictive, which limits the learning ability of the material network.

- At level $k+1$ with $0 \leq k \leq N_{\text{levels}} - 2$: we divide the $2^{N_{\text{levels}}-1-k}$ interaction mechanisms at level k into $2^{N_{\text{levels}}-2-k}$ pairs, from which $2^{N_{\text{levels}}-2-k}$ interaction mechanisms are created with all the material nodes in each pair. Clearly, this recursive procedure stops at level $N_{\text{levels}} - 1$ when all material nodes are present in a unique interaction mechanism.

Figure 2 shows an example of an interaction-based MN with $N_{\text{levels}} = 3$, $N = 16$- number of material nodes numbered from 0 to 15, and $M = 7$ - number of interaction mechanisms. Its architecture $\boldsymbol{\mathcal{A}}$ reads

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \mathcal{V}^0 & \mathcal{V}^1 & \mathcal{V}^2 & \mathcal{V}^3 & \mathcal{V}^4 & \mathcal{V}^5 & \mathcal{V}^6 \end{bmatrix}, \tag{86}$$

in which

$$\begin{cases} \mathcal{V}^0 &= [0, 1, 8, 9] \ , \\ \mathcal{V}^1 &= [2, 3, 10, 11] \ , \\ \mathcal{V}^2 &= [4, 5, 12, 13] \ , \\ \mathcal{V}^3 &= [6, 7, 14, 15] \ , \\ \mathcal{V}^4 &= [0, 1, 2, 3, 8, 9, 10, 11] \ , \\ \mathcal{V}^5 &= [4, 5, 6, 7, 12, 13, 14, 15] \ , \text{ and} \\ \mathcal{V}^6 &= [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] \ . \end{cases} \tag{87}$$

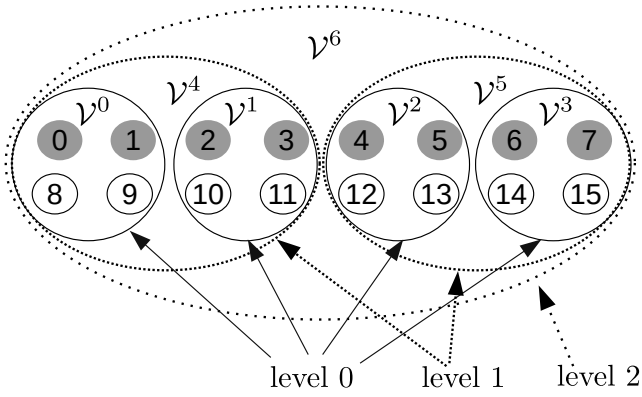


Figure 2: Example of an architecture of 16 material nodes with $N_{\text{levels}} = 3$. The filled nodes and blank nodes correspond to the solid nodes and the the void nodes respectively. The number inside each node corresponds to its index.

In general, an interaction-based MN constructed by the procedure above consists of $2^{N_{\text{levels}}+1}$ material nodes and $2^{N_{\text{levels}}} - 1$ interaction mechanisms. The number of fitting parameters in an architecture of N_{levels} levels, denoted by $N_{\mathbf{X}}$, relates to N_{levels} by the following relation

$$N_{\mathbf{X}} = \underbrace{2^{N_{\text{levels}}+1}}_{\text{size of } \widetilde{\mathbf{W}}} + \underbrace{N_{\text{levels}} 2^{N_{\text{levels}}+1} - 2^{N_{\text{levels}}} + 1}_{\text{size of } \widetilde{\mathbf{A}}} + \underbrace{(2^{N_{\text{levels}}} - 1)d}_{\text{size of } \widetilde{\mathbf{N}}}, \tag{88}$$

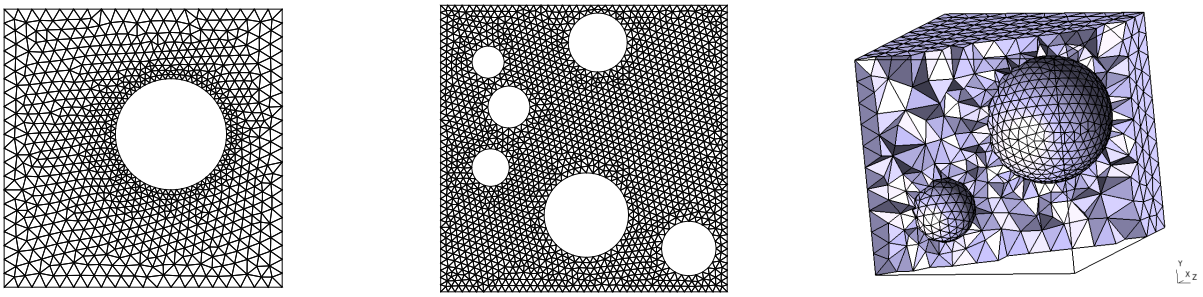
where $d = 1$ for a plane strain problem and $d = 2$ for a general 3-dimensional problem. In the following sections, we consider there different architectures with N_{levels} equal to 3, 4, and 5.

5.3. Offline training for various microstructures

Interaction material networks are trained for three different microstructures as shown in Fig. 3, which illustrates their finite element meshes:

- A two-dimensional microstructure with a single void, see Fig. 3(a), is denoted by 2D-RVE1. The volume fraction of the void is equal to 0.1257 and the finite element mesh consists of 1160 six-node triangular elements.
- A two-dimensional microstructure with multiple voids, see Fig. 3(b), is denoted by 2D-RVE2. The volume fraction of the voids is equal to 0.17 and the finite element mesh consists of 3544 six-node triangular elements.
- A three-dimensional microstructure with spherical voids, see Fig. 3(c), is denoted by 3D-RVE. The volume fraction of the voids is equal to 0.1272 and the finite element mesh consists of 10315 10-node tetrahedral elements.

Both the NT and LET procedures are performed for these three microstructures.



(a) 2D-RVE1 (b) 2D-RVE2 (c) 3D-RVE

Figure 3: Microstructures with voids embedded in an elastoplastic matrix: (a) two-dimensional single void (denoted by 2D-RVE1), (b) two-dimensional multiple voids (denoted by 2D-RVE2), and (c) three-dimensional with spherical voids (denoted by 3D-RVE) microstructures.

5.3.1. Nonlinear training

To collect the sequential data as described in Section 4.1.2, multi-axial and proportional loading paths in terms of strain sequences are generated. Once the bound $\bar{\boldsymbol{\varepsilon}}^{\max}$, *i.e.* so-called maximal homogenized strain, is known, the whole path is given by

$$\bar{\boldsymbol{\mathcal{E}}} = \left[\frac{i+1}{T} \bar{\boldsymbol{\varepsilon}}^{\max} \text{ for } i = 0, \dots, T-1 \right] , \tag{89}$$

where T is the number of data points. The maximal homogenized strain $\bar{\boldsymbol{\varepsilon}}^{\max}$ is generated such that a wide range of the multi-axial states is covered and the following condition is satisfied

$$R_{\max} = \sqrt{(\bar{\varepsilon}_{00}^{\max})^2 + (\bar{\varepsilon}_{11}^{\max})^2 + (\bar{\varepsilon}_{22}^{\max})^2 + (\bar{\varepsilon}_{01}^{\max})^2 + (\bar{\varepsilon}_{02}^{\max})^2 + (\bar{\varepsilon}_{12}^{\max})^2} , \tag{90}$$

where $R_{\max}$ is the maximal radius of the sampling space. $R_{\max}$ is chosen to obtain a nonlinearity effect in the local state inside the microscopic model while remembering that the model is written in a small strain setting. In this section, we set $R_{\max} = 0.2$ since a considerable plastic strain

is accumulated at this strain level. To satisfy the condition (90), $\bar{\boldsymbol{\varepsilon}}^{\max}$ can be parameterized by generalized spherical coordinates as

$$\begin{cases} \bar{\varepsilon}_{00}^{\max} = R_{\max}\cos\left(\phi_0\right) , \\ \bar{\varepsilon}_{11}^{\max} = R_{\max}\sin\left(\phi_0\right)\cos\left(\phi_1\right) , \\ \bar{\varepsilon}_{22}^{\max} = R_{\max}\sin\left(\phi_0\right)\sin\left(\phi_1\right)\cos\left(\phi_2\right) , \\ \bar{\varepsilon}_{01}^{\max} = R_{\max}\sin\left(\phi_0\right)\sin\left(\phi_1\right)\sin\left(\phi_2\right)\cos\left(\phi_3\right) , \\ \bar{\varepsilon}_{02}^{\max} = R_{\max}\sin\left(\phi_0\right)\sin\left(\phi_1\right)\sin\left(\phi_2\right)\sin\left(\phi_3\right)\cos\left(\phi_4\right) , \text{ and} \\ \bar{\varepsilon}_{12}^{\max} = R_{\max}\sin\left(\phi_0\right)\sin\left(\phi_1\right)\sin\left(\phi_2\right)\sin\left(\phi_3\right)\sin\left(\phi_4\right) . \end{cases} \tag{91}$$

$$\text{with} \begin{cases} [\phi_0\,\phi_1\,\phi_2\,\phi_3\,\phi_4] \in [0\,\pi]^4 \times [0\,2\pi) & \text{for a 3-dimensional path}, \\ [\phi_0\,\phi_1] \in [0\,\pi] \times [0\,2\pi)\,, \phi_2 = \frac{\pi}{2}\,, \text{ and } \phi_3 = \phi_4 = 0 & \text{for a plane strain path}. \end{cases}$$

To randomly generate $\bar{\boldsymbol{\varepsilon}}^{\max}$, all angles are randomly generated using uniform distributions over their admissible ranges.

An offline dataset $\mathbb{D}_{\text{offline}}$ of N_{sample} samples is generated. Since $L_{\bar{\mathcal{Z}}} = 1$ is considered in Eqs. (63, 64) and the property of interest (85), which characterizes the microstructure evolution, also needs to be extracted. The data generation process is summarized as follows:

- **for** l **from** 0 **to** $N_{\text{sample}} - 1$:
  - (i) generate $\bar{\boldsymbol{\mathcal{E}}}^{[l]}$ as follows:
    - For a 3-dimensional state:
    $$\phi_i \in U\left(0, \pi\right) \text{ for } i = 0, 1, 2, 3\,, \text{ and } \phi_4 \in U\left(0, 2\pi\right); \tag{92}$$
    - For a plane strain problem:
    $$\phi_0 \in U\left(0, \pi\right)\,, \phi_1 \in U\left(0, 2\pi\right)\,, \phi_2 = \frac{\pi}{2}\,, \phi_3 = 0\,, \text{ and } \phi_4 = 0\,; \tag{93}$$
    - estimate $\bar{\boldsymbol{\varepsilon}}^{\max}$ following Eq. (91);
    - create a strain sequence following Eq. (89);
  - (ii) run the microscopic simulation as described in Section 2.2:
    - initialize the sequences of $\widehat{\boldsymbol{\mathcal{S}}}^{[l]}$ and $\widehat{\mathcal{Z}}^{0,[l]}$ with the same length as $\bar{\boldsymbol{\mathcal{E}}}^{[l]}$;
    - **for** k **from** 0 **to** len $\left(\bar{\boldsymbol{\mathcal{E}}}^{[l]}\right) - 1$:
      - solve (incrementally) the microscopic problem for the homogenized strain $\bar{\boldsymbol{\varepsilon}}_k^{[l]}$;
      - extract the homogenized stress following Appendix A and store in $\hat{\boldsymbol{\sigma}}_k^{[l]}$;
      - extract the homogenized value of z^0 following Eq. (85) and store in $\hat{z}_k^{0,[l]}$;
      - store thermodynamics internal variables of the microscopic problem for next step $k + 1$;
  - (iii) append $\bar{\boldsymbol{\mathcal{E}}}^{[l]}$, $\widehat{\boldsymbol{\mathcal{S}}}^{[l]}$, and $\widehat{\mathcal{Z}}^{0,[l]}$ to the offline dataset $\mathbb{D}_{\text{offline}}$.
  - (iv) exit.

From $\mathbb{D}_{\text{offline}}$, two non-overlapping subsets $\mathbb{D}_{\text{train}}$ and $\mathbb{D}_{\text{valid}}$ respectively of N_{train} samples and of N_{valid} samples are randomly extracted for the training and validation stages. For the cases of 2D-RVE1 and 2D-RVE2, 60 samples are generated for each case and then randomly divided into 50

training samples and 10 validation samples. In the case of 3D-RVE, 100 samples are generated and divided into 80 training samples and 20 validation samples.

The histories of the training error C_{train} and validation error C_{valid} evolutions using the NT procedure are reported in Fig. 4 for the different material network architectures with N_{levels} equal to 3, 4, and 5 and for the microstructures reported in Fig. 3. The oscillation of the training and validation errors are observed at the beginning as a result of the stochastic gradient descent stepping. The saturation of the training and validation errors are found and further reduction of these errors could not be achieved by continuing more epochs. We can see from the figures that a material network with a higher N_{levels} involving more interaction mechanisms allows reducing the saturation values of these errors. For the case of 3D-RVE with $N_{\text{levels}} = 5$, the training process is stopped after 400 epochs when the saturation of the training error is attained while in other cases, the maximal number of epochs is set to 500.

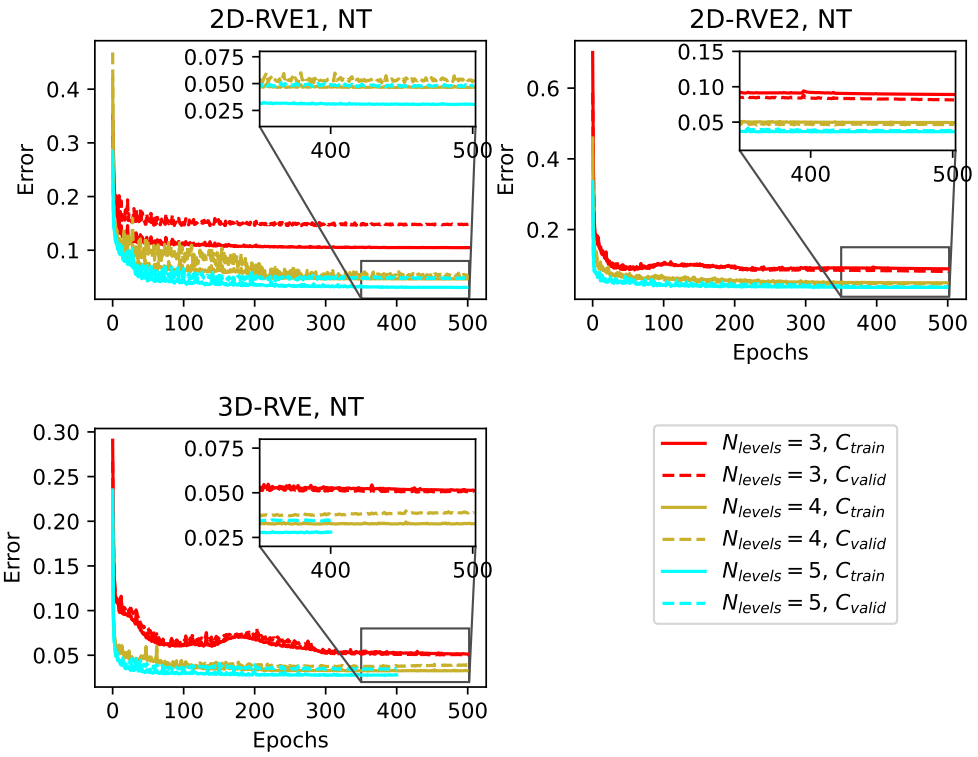


Figure 4: Nonlinear training (NT) - histories of the training error C_{train} and validation error C_{valid} for the RVEs reported in Fig. 3.

The computational times for generating the offline dataset and for the NT procedure are summarized in Tab. 2. For generating the offline dataset $\mathbb{D}_{\text{offline}}$, the cases of 2D-RVE1, 2D-RVE2, and 3D-RVE respectively took 3.5h, 6.5h, and 800h with a single processor, in which the three-dimensional analysis is seen to be time-consuming. Since the generation of each sample is independent, the wall-clock time for the generation process linearly decreases with the number of processors. The training time increases with the number of fitting parameters involved in the material networks and the number of samples considered during the training.

5.3.2. Linear elastic training

An offline dataset $\mathbb{D}_{\text{offline}}$ of N_{sample} samples is generated for the linear elastic training mentioned in Section 4.2.2. For a given microstructure, the data generation process is summarized as follows:

Table 2: Computational time for generating the offline training dataset and for training the network with the nonlinear training (NT) procedure using a single processor. The number of fitting parameters $N_{\mathbf{X}}$ following Eq. (88) is also reported.

Case	$\mathbb{D}_{\text{offline}}$ (N_{sample})	$N_{\text{levels}} = 3$ ($N_{\mathbf{X}}$)	$N_{\text{levels}} = 4$ ($N_{\mathbf{X}}$)	$N_{\text{levels}} = 5$ ($N_{\mathbf{X}}$)
2D-RVE1, NT	3.5h (60)	2.3h (64)	6.3h (160)	19.1h (384)
2D-RVE2, NT	6.5h (60)	2.2h (64)	6h (160)	20.3h (384)
3D-RVE, NT	800h (100)	4.2h (71)	11.5h (175)	29.7h (415)

- **for** l **from** 0 **to** $N_{\text{sample}} - 1$:

  (i) generate the input $\mathbb{X}^{[l]}$ of P elastic matrices. The generation procedure for a general 3-dimensional problem is found in [29], in which P material matrices are generated through randomly generating the Young's moduli and Poison ratios of an arbitrary orthotropic elastic matrix. The particular plane strain problem follows the work [35] and is reported in Appendix H;

  (ii) assign these P random elastic matrices in $\mathbb{X}^{[l]}$ governing the elastic behaviors of the P physical phases in this microstructure;

  (iii) compute the corresponding homogenization material tensor $\hat{\mathbf{L}}^{[l]}$ of the microscopic model defined on this microstructure with the periodic boundary condition following the work [37];

  (iv) append the pair $\left(\mathbb{X}^{[l]}, \hat{\mathbf{L}}^{[l]}\right)$ to the offline dataset $\mathbb{D}_{\text{offline}}$.

Two non-overlapping subsets $\mathbb{D}_{\text{train}}$ and $\mathbb{D}_{\text{valid}}$ respectively of N_{train} samples and of N_{valid} samples are randomly extracted from $\mathbb{D}_{\text{offline}}$ for training and validation. For each microstructure reported in Fig. 3, 300 samples are generated and then randomly separated into 250 training samples and 50 validation samples.

The histories of the training error C_{train} and validation error C_{valid} evolutions using the LET procedure are reported in Fig. 5 for the different material network architectures with N_{levels} equal to 3, 4, and 5 and for the three microstructures reported in Fig. 3. The training iterations are stopped after 100 epochs when the saturation of the training error is observed. Similarly to the nonlinear training procedure reported in Fig. 4, a material network with a higher N_{levels} allows reducing the saturation value of the training error.

The computational times for generating the offline dataset and for the LET procedure are summarized in Tab. 3. The training time increases with the number of fitting parameters involved in the material networks. Compared to the NT procedure reported in Tab. 2, a LET procedure is less time consuming in both the data generation and the training. However, a material network trained with the NT procedure allows improving the accuracy of its online prediction as shown in the next sections.

Table 3: Computational time for generating the offline training dataset and for training the network with the linear elastic training (LET) procedure using a single processor. The number of fitting parameters $N_{\mathbf{X}}$ following Eq. (88) is also reported.

Case	$\mathbb{D}_{\text{offline}}$ (N_{sample})	$N_{\text{levels}} = 3$ ($N_{\mathbf{X}}$)	$N_{\text{levels}} = 4$ ($N_{\mathbf{X}}$)	$N_{\text{levels}} = 5$ ($N_{\mathbf{X}}$)
2D-RVE1, LET	133s (300)	0.16h (64)	0.4h (160)	1.1h (384)
2D-RVE2, LET	161s (300)	0.14h (64)	0.43h (160)	1.12h (384)
3D-RVE, LET	4.5h (300)	0.3h (71)	0.8h (175)	2.7h (415)

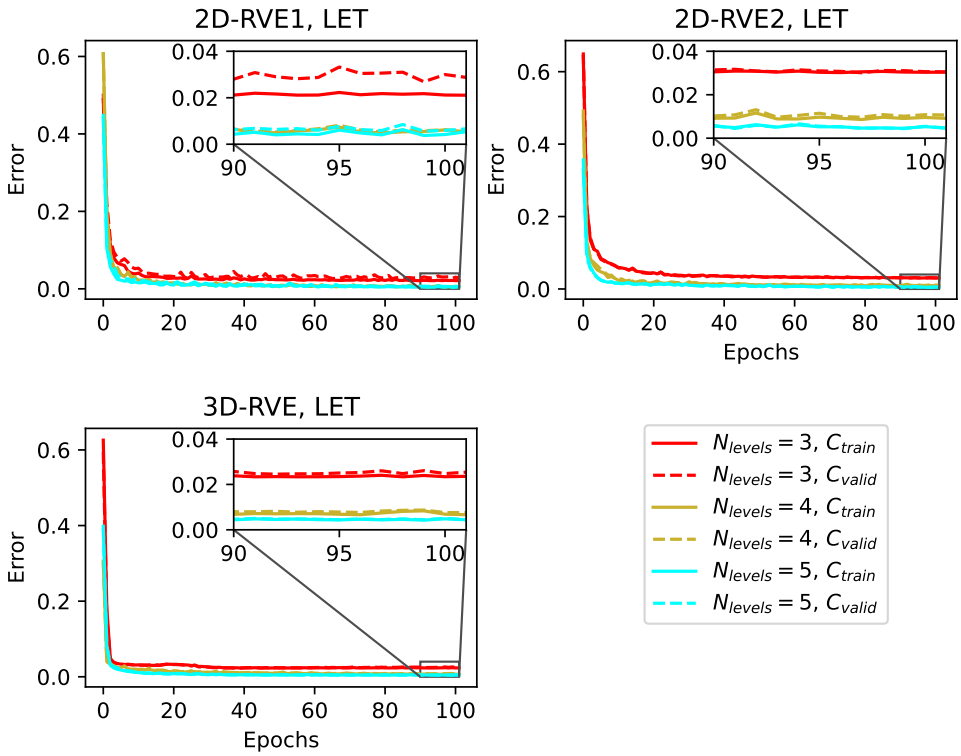


Figure 5: Linear elastic training (LET) - histories of the training error C_{train} and validation error C_{valid} for the RVEs reported in Fig. 3.

5.4. Online prediction of the trained material networks

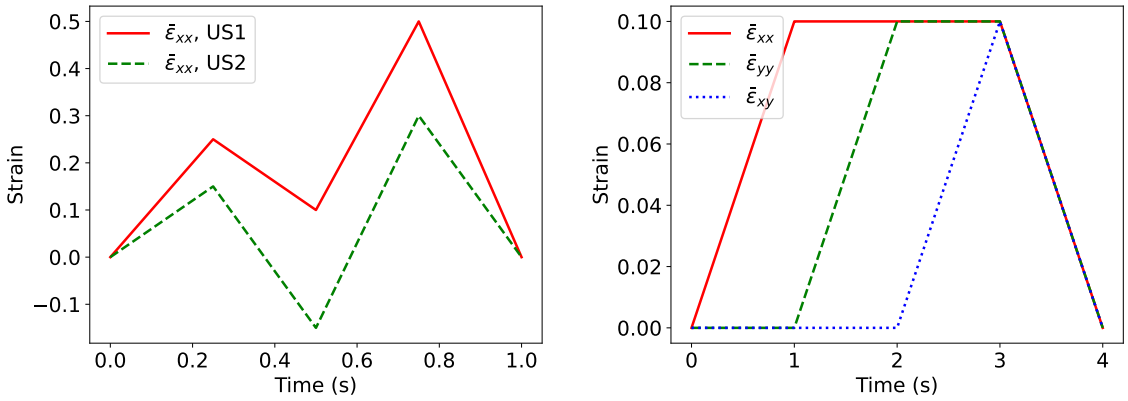
The online predictions of the trained material network using the NT and LET procedures described in Section 5.3 are assessed with loading cases that are not used for training. Three different scenarios involving loading-unloading, including non-proportional cases, and at different strain levels are considered:

- US1 and US2: uniaxial tensions following the x-direction are conducted with two prescribed $\bar{\varepsilon}_{xx}$ loading paths as illustrated in Fig. 6(a). All shear components of $\bar{\boldsymbol{\varepsilon}}$ are set to 0 and the remaining components are found by constraining the uniaxial stress condition. It is noted that a truly uniaxial stress state is obtained in a three-dimensional problem while, in a plane strain problem, the uniaxial stress state is obtained only in-plane.

- MA: a multiaxial strain loading path is prescribed by a non-proportional loading-unloading paths of $\bar{\varepsilon}_{xx}$, $\bar{\varepsilon}_{yy}$ and $\bar{\varepsilon}_{xy}$ while the other strain components are set to 0, as illustrated in Fig. 6(b).

5.4.1. Material networks trained with the nonlinear training (NT) procedure

In the case of the 2D-RVE1 microstructure, the evolutions of the homogenized stress and of the porosity are reported in Fig. 7 for the three trained material networks with the US1, US2, and MA loading conditions. The corresponding results using the microscopic problem (denoted by direct numerical simulation -DNS) are also reported for comparison purpose. In the cases of N_{levels} equal to 4 and 5, the full-field results are well reproduced by the material networks for all loading conditions while the case of N_{levels} equal to 3 yields a more erroneous prediction.



(a) US1 and US2 paths (b) MA path

Figure 6: Loading paths considered in the online predictions: (a) uniaxial stress conditions (US1 and US2) with prescribed loading-unloading path of $\bar{\varepsilon}_{xx}$ and other strain components obtained to reach the uniaxial stress condition and (b) multiaxial strain condition (MA) with prescribed loading-unloading paths of $\bar{\varepsilon}_{xx}$, $\bar{\varepsilon}_{yy}$, and $\bar{\varepsilon}_{xy}$ and other strain components set to 0.

In the case of the 2D-RVE2 microstructure, good predictions in terms of the evolutions of the homogenized stress and of the porosity are also obtained in comparison with the corresponding DNS results with the trained material networks when N_{levels} is equal to 4 and 5 as shown in Fig. 8. The case of N_{levels} equal to 3 leads to an erroneous prediction.

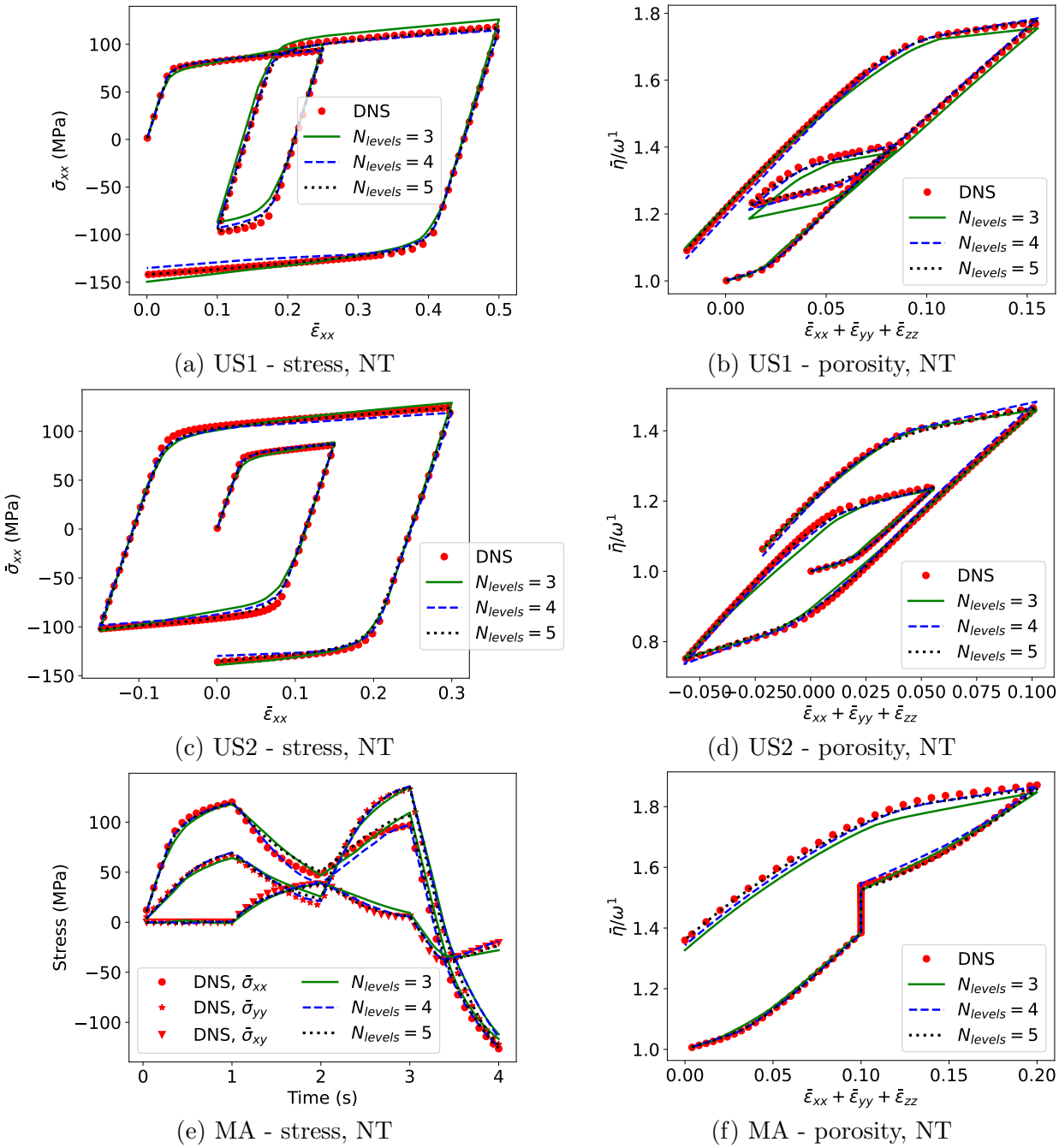
In the case of the 3D-RVE microstructure, good predictions in terms of the evolutions of the homogenized stress and of the porosity are also obtained with the three trained material networks compared to the corresponding DNS results using the microscopic problem as shown in Fig. 9.

It is thus shown that a sufficiently high number of material nodes and interaction mechanisms is necessary for a good prediction. Moreover, although the strain value goes beyond the range of 0.2 used during the training, the prediction remains accurate when extrapolating beyond this offline sampling space.

5.4.2. Material networks trained with the linear elastic training (LET) procedure

The evolutions of the homogenized stress and of the porosity are reported in Figs. 10, 11, and 12 for the three trained material networks with the US1, US2, and MA loading conditions and respectively for the 2D-RVE1, 2D-RVE2, and 3D-RVE microstructures. The corresponding DNS results are also reported for comparison purpose. For the cases of N_{levels} equal to 4 and 5, good predictions are obtained for the US1 and US2 loading conditions but they are less accurate for the MA loading condition. The simple network of N_{levels} equal to 3 yields an erroneous prediction in all loading conditions.

We found that a sufficiently high number of material nodes and interaction mechanisms is necessary for a good prediction. Although the LET procedure requires only elastic behavior for the training stage, satisfying predictions in the online prediction can still be obtained with a suitable network architecture. In comparison with the material networks trained by the NT procedure for a given network architecture, the ones trained by the LET procedure provides less accuracy while the ones trained by the NT procedure are time-consuming. The choice between the NT and LET procedures needs to consider the compromise between the offline computational cost and the accuracy.



(a) US1 - stress, NT

(b) US1 - porosity, NT

(c) US2 - stress, NT

(d) US2 - porosity, NT

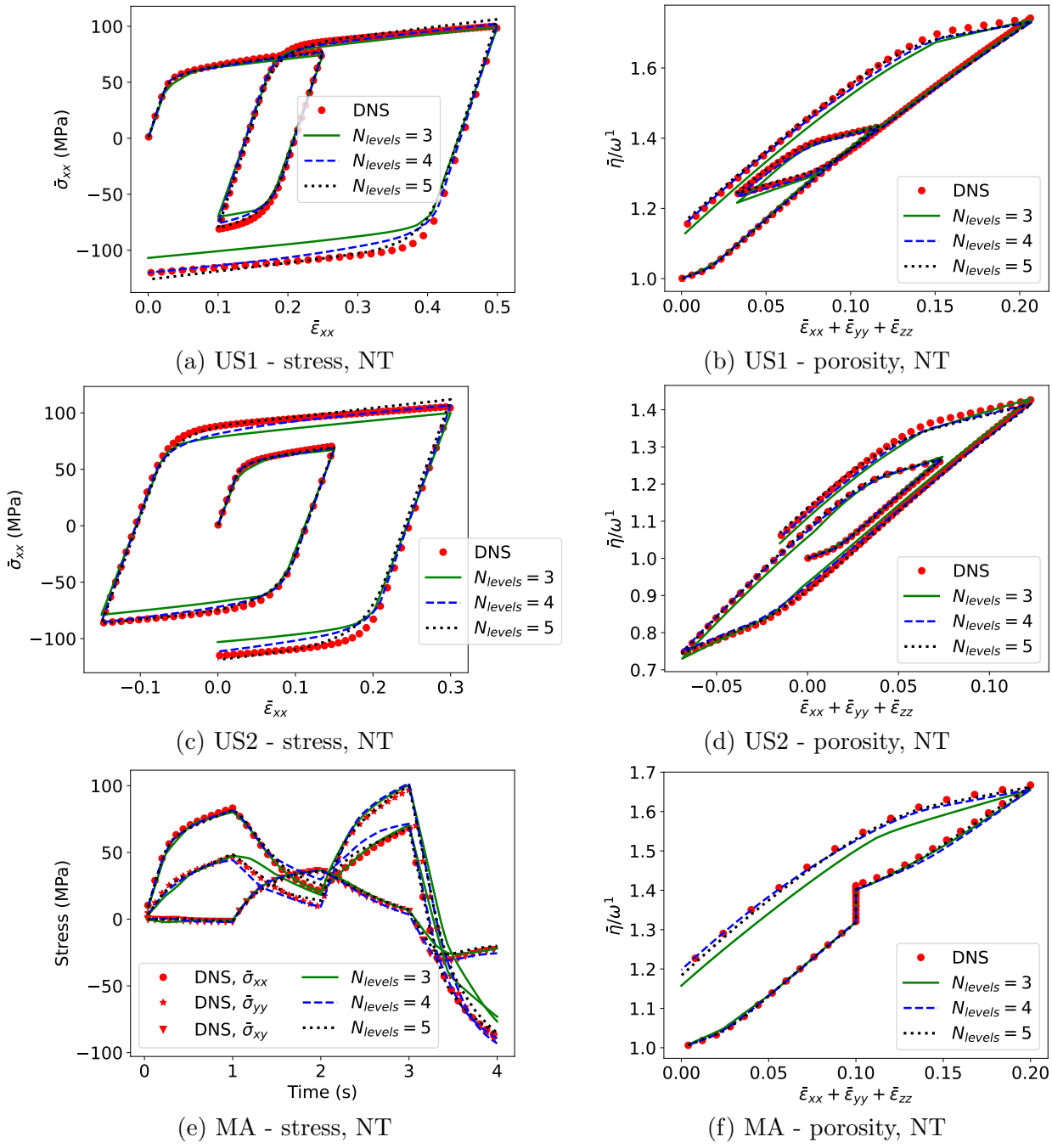
(e) MA - stress, NT

(f) MA - porosity, NT

Figure 7: 2D-RVE1, NT - comparison of the results predicted by the material network with different architectures and by the direct numerical simulation (DNS) under different loading conditions: US1 (a, b), US2 (c, d), and MA (e,f).

5.4.3. Computational speedup

The use of the material networks allows accelerating the prediction. The computational efficiency in comparison with the full-field finite element simulation is characterized by a speedup factor, which



(a) US1 - stress, NT

(b) US1 - porosity, NT

(c) US2 - stress, NT

(d) US2 - porosity, NT

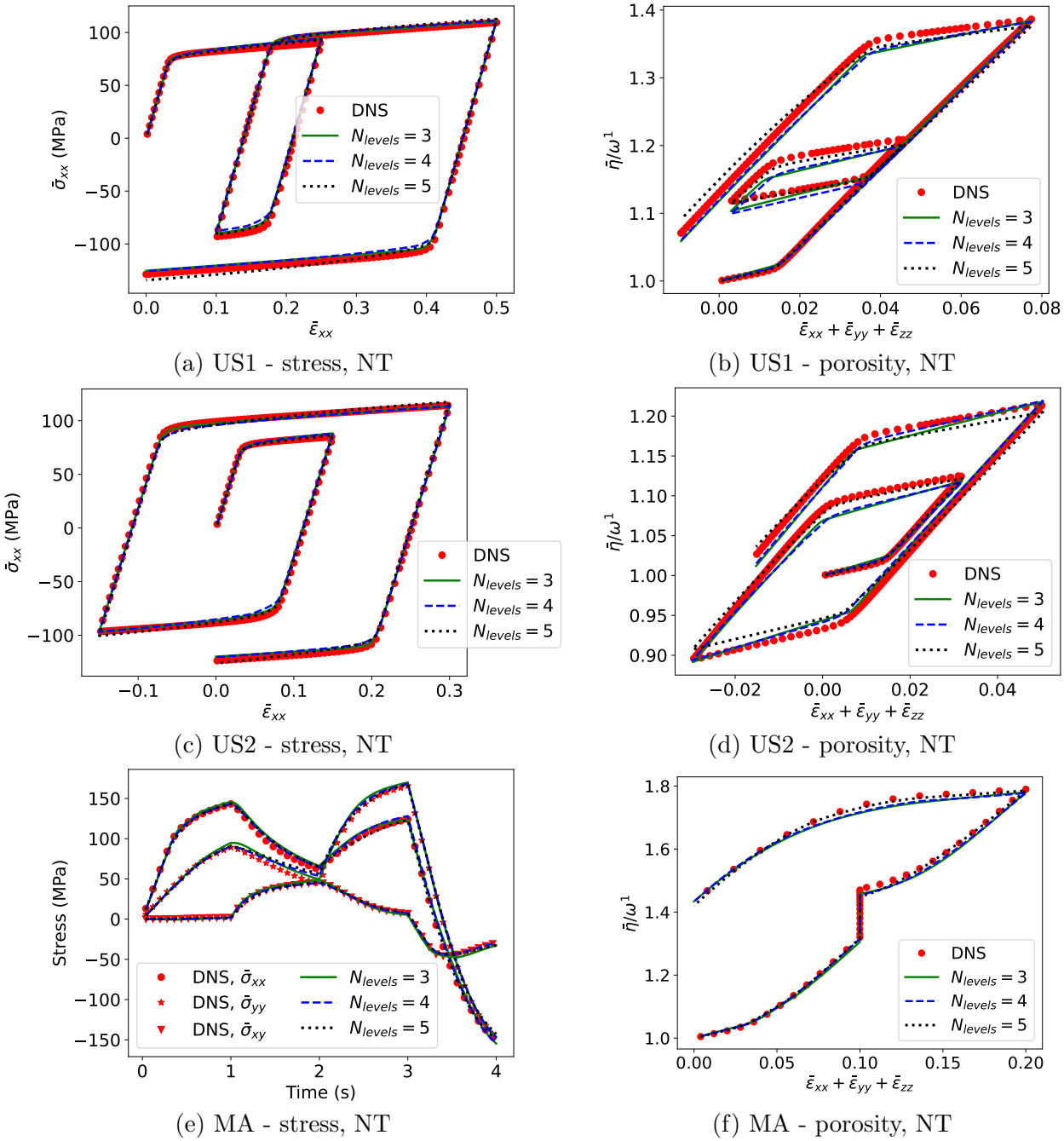
(e) MA - stress, NT

(f) MA - porosity, NT

Figure 8: 2D-RVE2, NT - comparison of the results predicted by the material network with different architectures and by the direct numerical simulation (DNS) under different loading conditions: US1 (a, b), US2 (c, d), and MA (e,f).

is defined as

$$\text{Speedup} = \frac{T_{\text{DNS}}}{T_{\text{MN}}}\,, \tag{94}$$



(a) US1 - stress, NT

(b) US1 - porosity, NT

(c) US2 - stress, NT

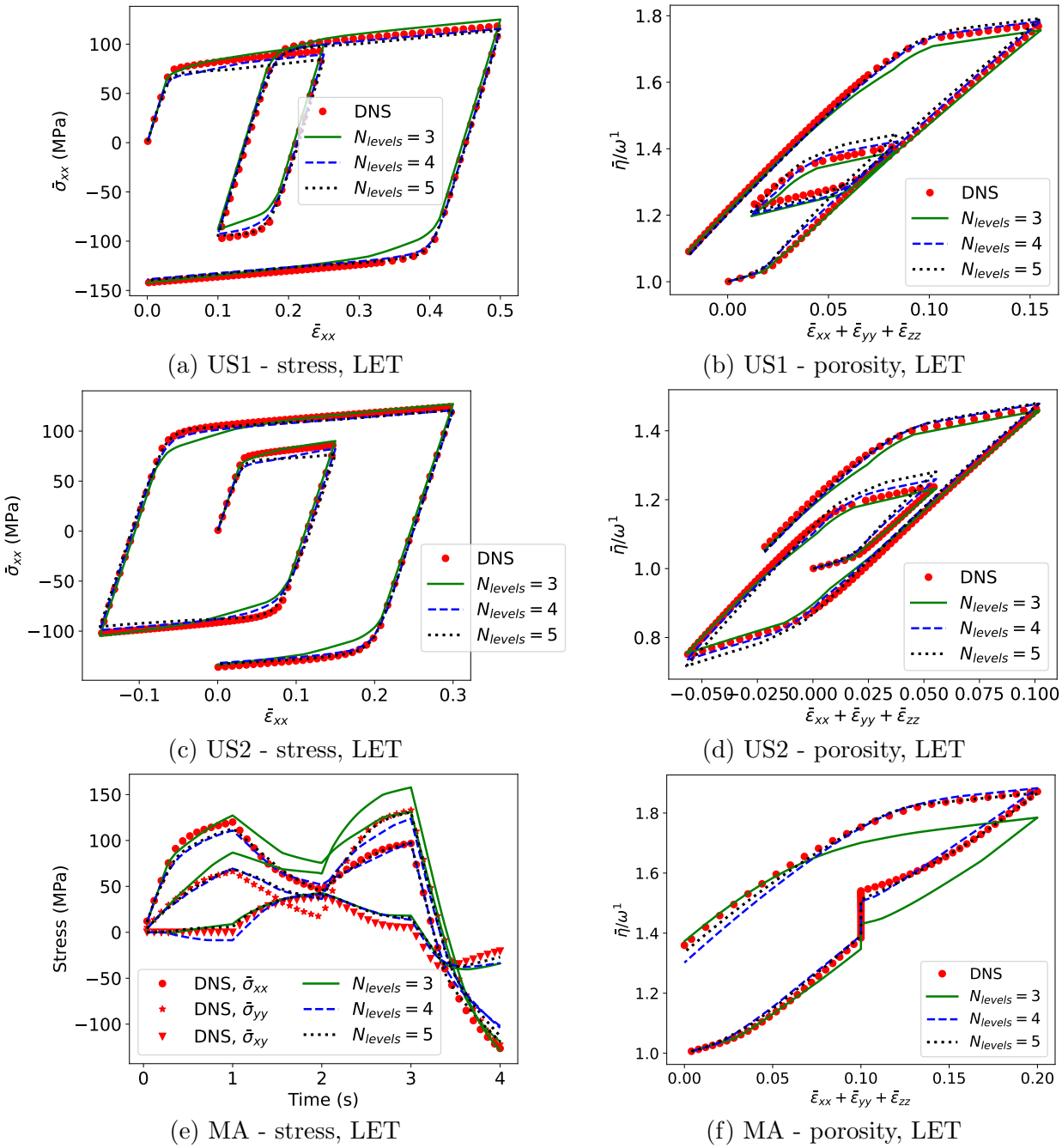
(d) US2 - porosity, NT

(e) MA - stress, NT

(f) MA - porosity, NT

Figure 9: 3D-RVE, NT - comparison of the results predicted by the material network with different architectures and by the direct numerical simulation (DNS) under different loading conditions: US1 (a, b), US2 (c, d), and MA (e,f).

where T_{DNS} and T_{MN} denote the wall-clock times required to complete the DNS and MN predictions, respectively. The speedup obtained by the material network is shown in Fig. 13 for the online predictions reported in Figs. 7, 8, 9, 10, 11, and 12. It can be seen that the material network predictions are much faster than the DNS, especially for the three-dimensional microstructure. Additional, a higher speedup is obtained with a smaller number of the material nodes since the resolution involves fewer degrees of freedom and less evaluations of the local constitutive behavior.



(a) US1 - stress, LET

(b) US1 - porosity, LET

(c) US2 - stress, LET

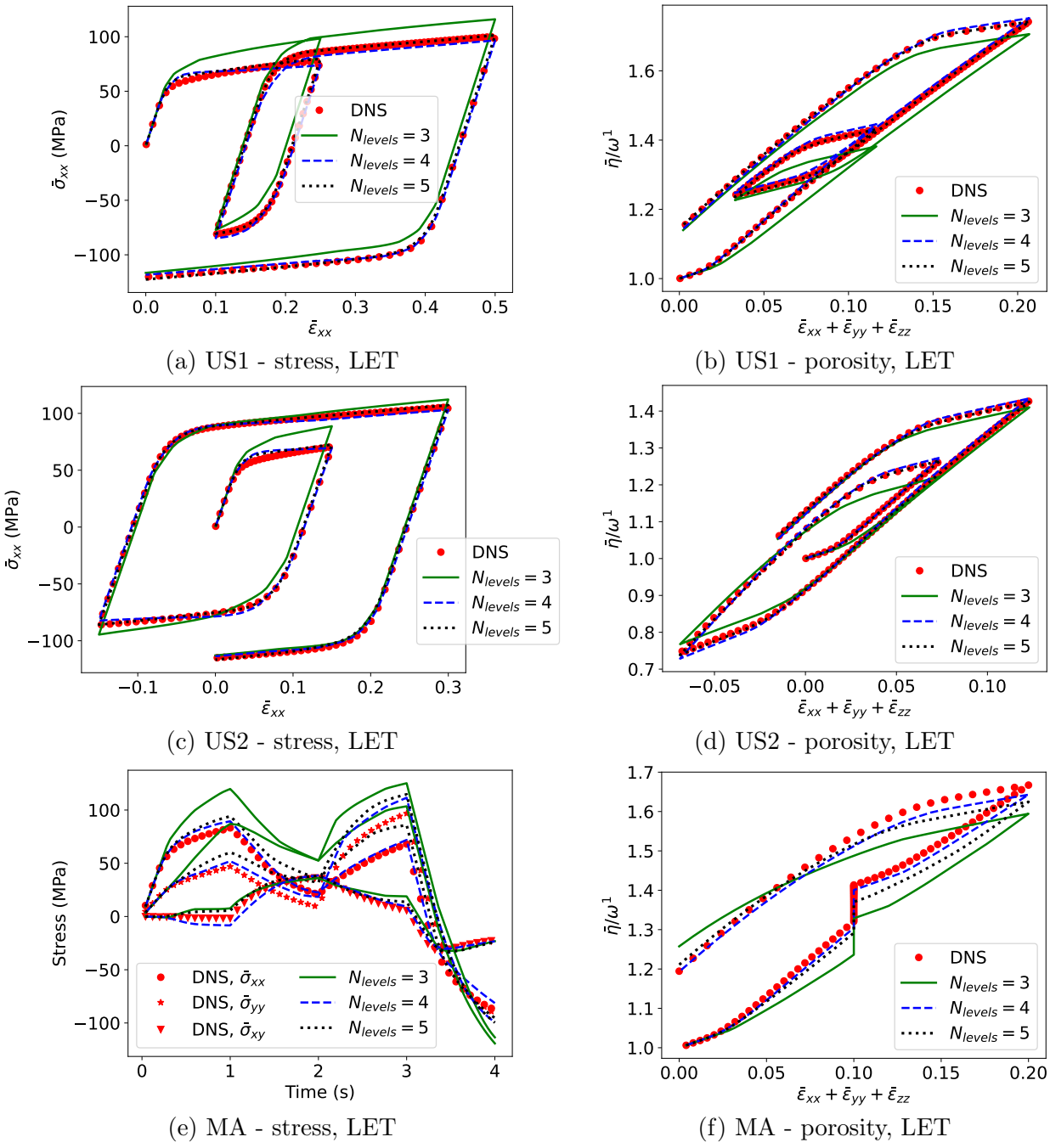
(d) US2 - porosity, LET

(e) MA - stress, LET

(f) MA - porosity, LET

Figure 10: 2D-RVE1, LET - comparison of the results predicted by the material network with different architectures and by the direct numerical simulation (DNS) under different loading conditions: US1 (a, b), US2 (c, d), and MA (e,f).

Although the material networks require the creation of an offline dataset and a training stage, their use provides a promising computational efficiency. Especially, when the NT procedure is employed, the material networks are trained only once with the proportional loading paths and can subsequently be used as a predictive model for different loading paths. The use of the NT procedure to train a material network allows improving its accuracy in comparison with using the LET procedure. However, as opposed to other materials networks [28, 29, 32, 33], the material



(a) US1 - stress, LET

(b) US1 - porosity, LET

(c) US2 - stress, LET

(d) US2 - porosity, LET

(e) MA - stress, LET

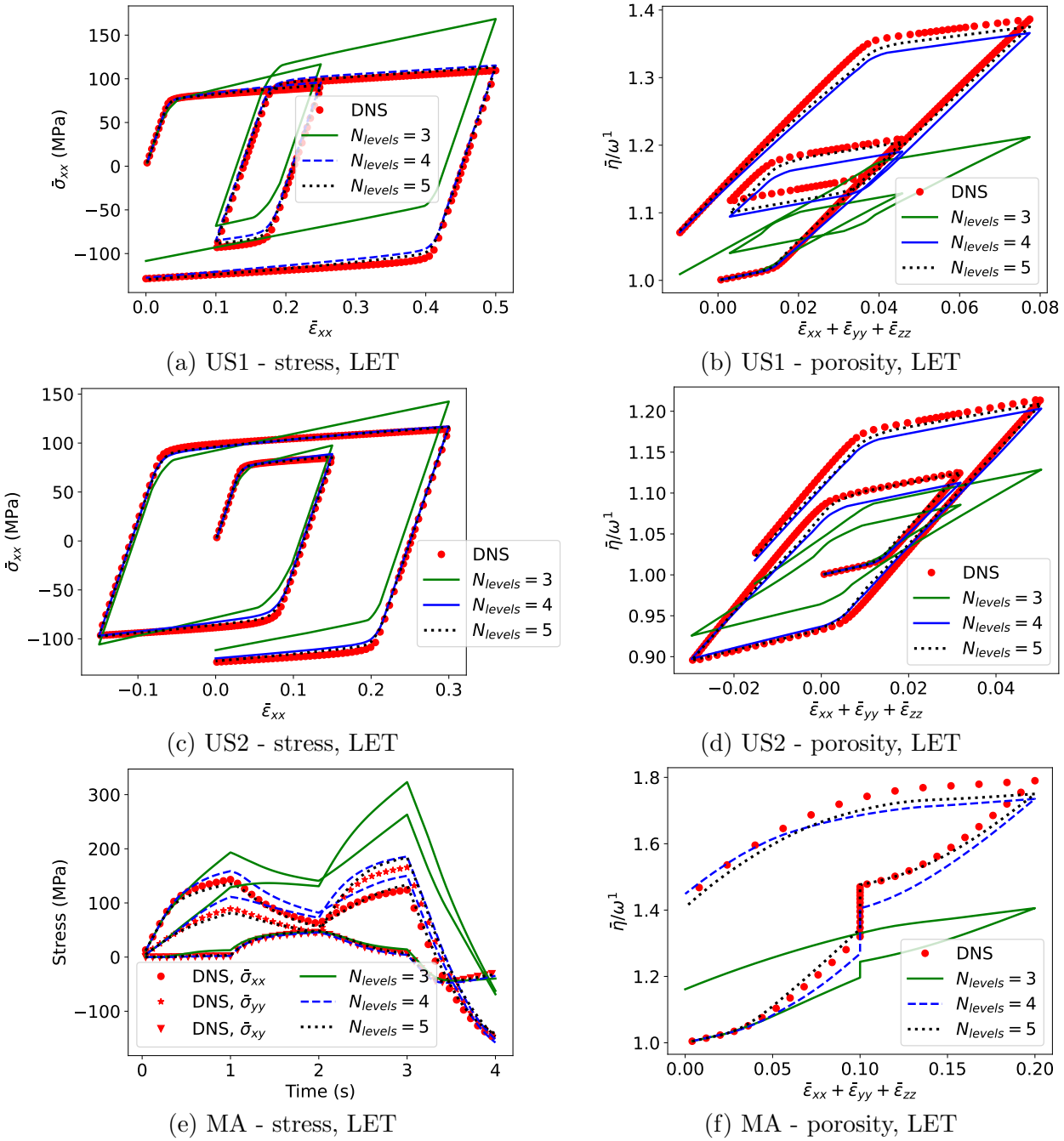
(f) MA - porosity, LET

Figure 11: 2D-RVE2, LET - comparison of the results predicted by the material network with different architectures and by the direct numerical simulation (DNS) under different loading conditions: US1 (a, b), US2 (c, d), and MA (e,f).

behaviors of the physical phases are used for the NT training, although extrapolation capabilities with respect to the material parameters set will be demonstrated in Section 5.6.

5.5. Thermodynamics of the interaction-based material network

In this section, we show that the trained material network respects the first and second laws of thermodynamics. We consider the material network with $N_{\text{levels}} = 5$ trained for 3D-RVE using



(a) US1 - stress, LET

(b) US1 - porosity, LET

(c) US2 - stress, LET

(d) US2 - porosity, LET

(e) MA - stress, LET

(f) MA - porosity, LET

Figure 12: 3D-RVE, LET - comparison of the results predicted by the material network with different architectures and by the direct numerical simulation (DNS) under different loading conditions: US1 (a, b), US2 (c, d), and MA (e,f).

the NT procedure as reported in Section 5.3.1. This material network is loaded by two other more complex loading scenarios:

- US3: uniaxial tensions following the x-direction are conducted with two prescribed $\bar{\varepsilon}_{xx}$ loading paths as illustrated in Fig. 14(a). All shear components of $\bar{\boldsymbol{\varepsilon}}$ are set to 0 and the remaining components are found by constraining the uniaxial stress condition.

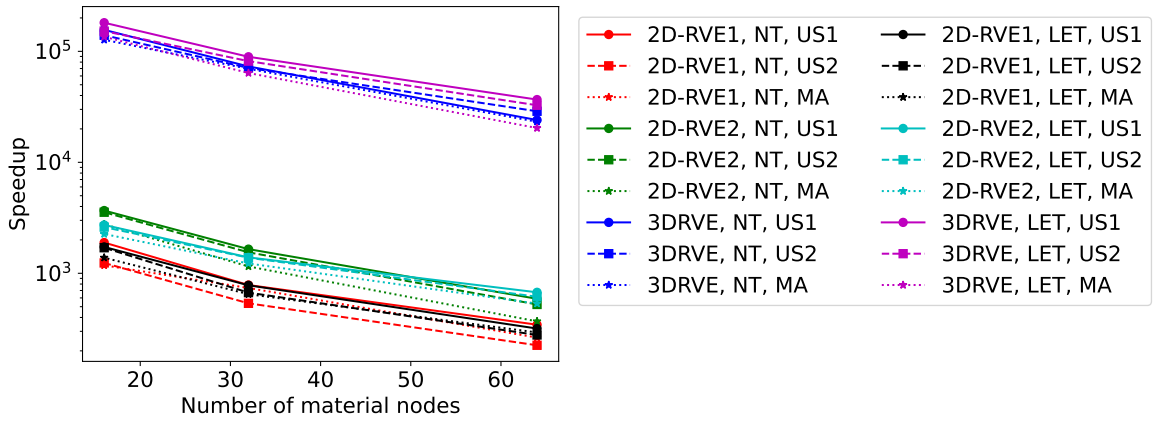
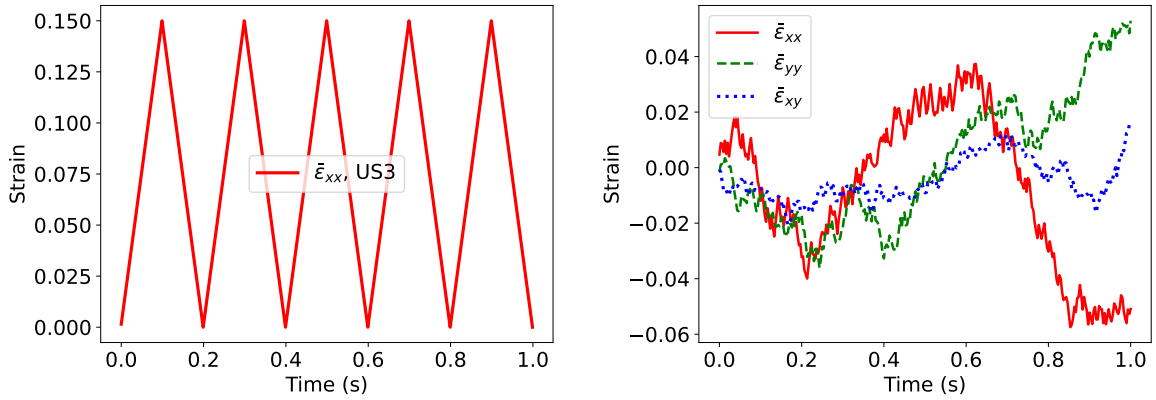


Figure 13: Speedup achieved with the material networks trained by the NT and LET procedures for the online predictions reported in Figs. 7, 8, 9, 10, 11, and 12.

- RANDOM: a multiaxial strain loading path is prescribed by a non-proportional loading-unloading paths of $\bar{\varepsilon}_{xx}$, $\bar{\varepsilon}_{yy}$ and $\bar{\varepsilon}_{xy}$ while the other strain components are set to 0, as illustrated in Fig. 14(b).



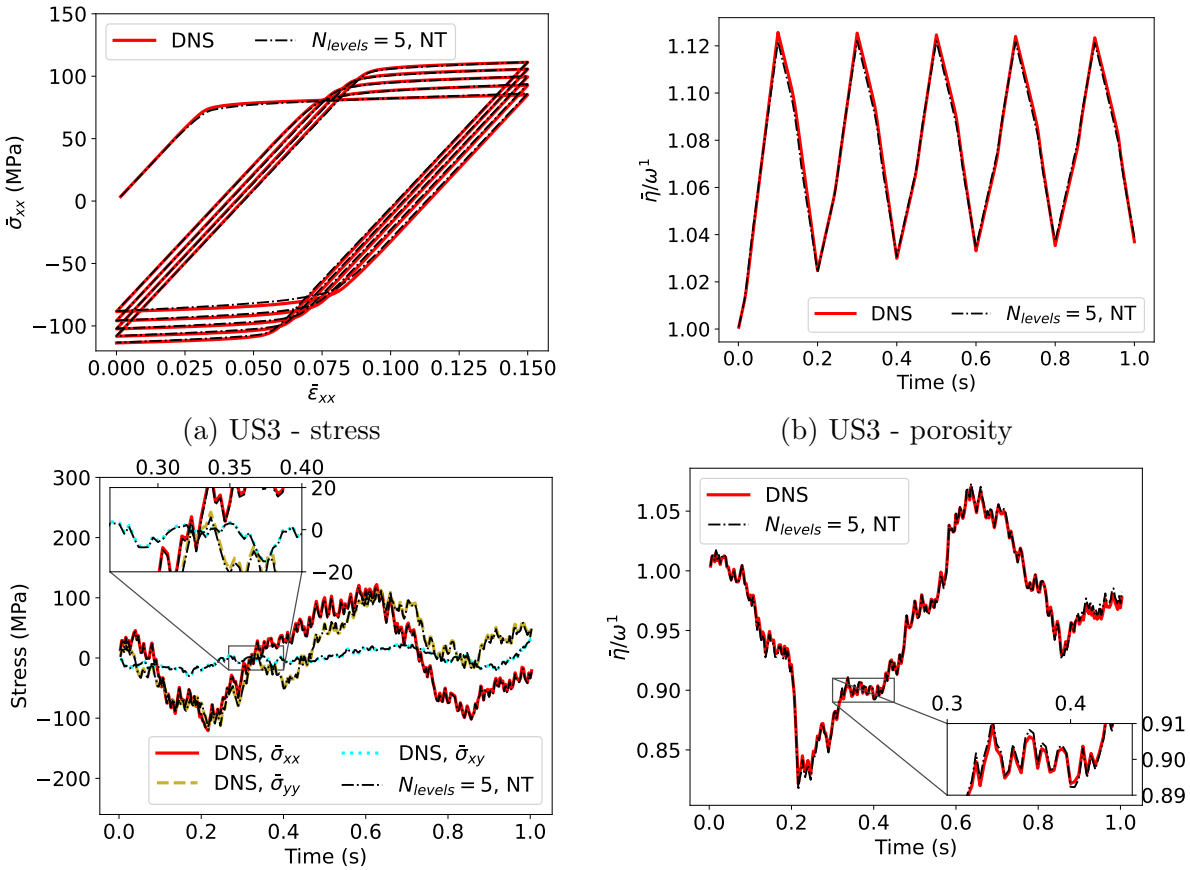
(a) US3 path (b) RANDOM path

Figure 14: Complex loading paths: (a) uniaxial stress condition (US3) with prescribed cyclic loading-unloading path for $\bar{\varepsilon}_{xx}$ and other strain components obtained to reach the uniaxial stress condition and (b) random loading condition (RANDOM) with random paths for $\bar{\varepsilon}_{xx}$, $\bar{\varepsilon}_{yy}$, and $\bar{\varepsilon}_{xy}$ and other strain components set to 0.

As shown in Fig. 15, good predictions in terms of the evolutions of the homogenized stress and of the porosity are also obtained with the trained material network compared to the corresponding DNS results using the microscopic problem. In particular, with the US3 loading condition (see Figs. 15(a) and (b)), the accuracy is maintained through multiple cycles.

In Fig. 16, the energy balance and the non-negative dissipation respectively following the first and second laws of thermodynamics are investigated for this trained material network with the US3 and RANDOM loading conditions. Considering a time step $[t_l \; t_{l+1}]$ with $l = 0, 1, \ldots$, on the one hand, the applied increment of the total mechanical energy reads

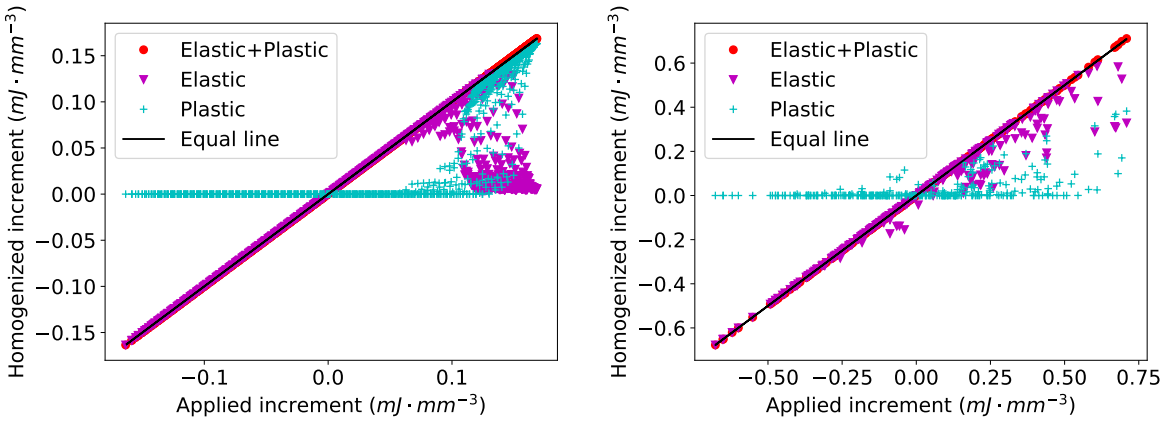
$$\Delta \mathcal{W}_l = \int_{t_l}^{t_{l+1}} \bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}} \, dt \, . \tag{95}$$



(a) US3 - stress (b) US3 - porosity

(c) RANDOM - stress (d) RANDOM - porosity

Figure 15: 3D-RVE, NT, $N_{\text{levels}} = 5$ - comparison of the results predicted by the material network and by the direct numerical simulation (DNS) under different loading conditions: US3 (a, b) and RANDOM (c, d).



(a) US3 (b) RANDOM

Figure 16: 3D-RVE, NT, $N_{\text{levels}} = 5$ - homogenized increments of the total mechanical energy, elastic energy, and plastic energy versus the applied increment of the total mechanical energy: US3 (a) and RANDOM (b).

On the other hand, the increment of the total mechanical energy in each material node is estimated

by

$$\Delta \mathcal{W}_l^i = \int_{t_l}^{t_{l+1}} \bar{\boldsymbol{\sigma}}^i : \dot{\bar{\boldsymbol{\varepsilon}}}^i \, dt \quad \text{for } i \in \mathcal{K} \,. \tag{96}$$

In each time step $[t_l \; t_{l+1}]$, the increment of the total mechanical energy $\Delta \mathcal{W}_l^i$ can also be decomposed into a reversible part (so-called elastic energy) and an irreversible part (so-called plastic energy). From these local values at material nodes, the homogenized increments of the total mechanical energy, elastic energy, and plastic energy are estimated using Eq. (38) over the whole material network, *i.e.* by using $\mathcal{C} = \mathcal{K}$. As shown in Fig. 16, the energy balance and the non-negative plastic dissipation of the trained material network with the US3 and RANDOM loading conditions in all time steps are demonstrated.

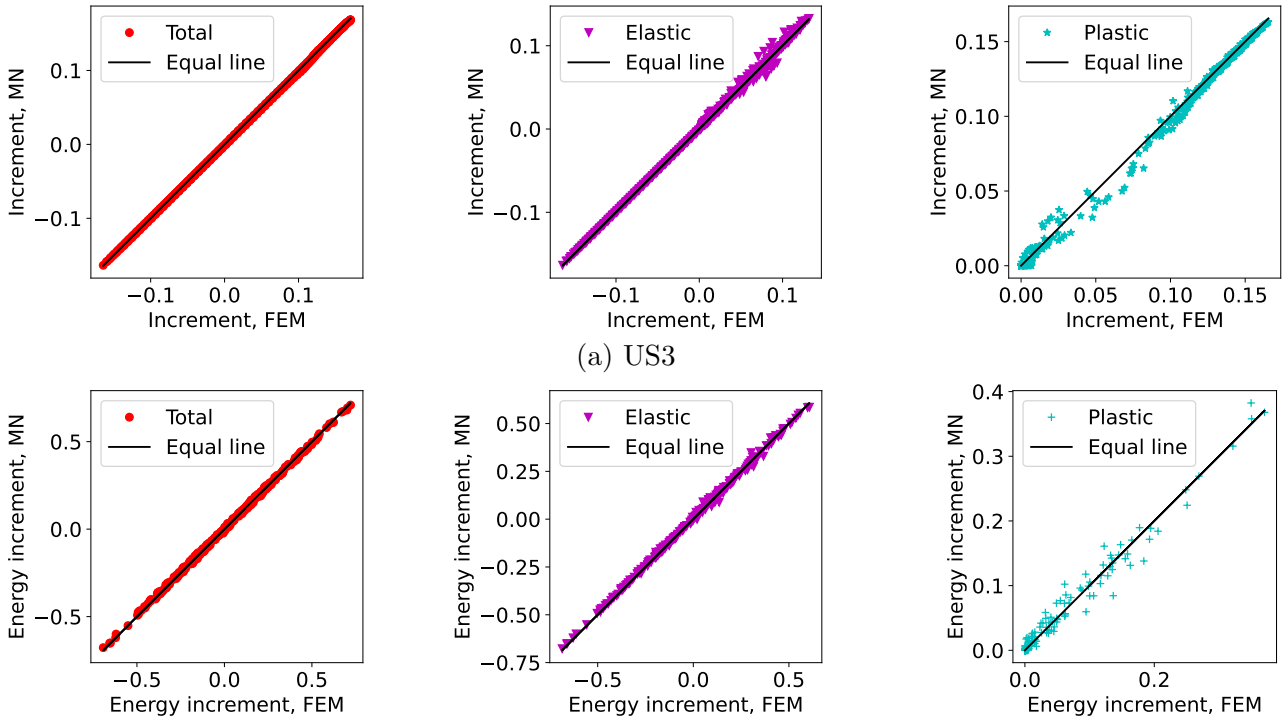


Figure 17: 3D-RVE, NT, $N_{\text{levels}} = 5$ - comparison of the homogenized increments of the total mechanical energy, elastic energy, and plastic energy predicted by the material network (MN) and by the direct numerical simulation (DNS) under different loading conditions: US3 (a) and RANDOM (b).

Figure 17 compares the homogenized increments of the total mechanical energy, elastic energy, and plastic energy predicted by the material network and the corresponding results predicted by the microscopic problem (denoted by DNS) in all time steps. For both loading cases, the increments of the total mechanical energy obtained by the material network and by DNS are in good agreement, which results from the enforcement of the Hill-Mandel condition (15) in the interaction-based MN definition. The increments of elastic energy and plastic energy obtained by the material network and by DNS are slightly different. This was expected since these equivalences are not enforced during the scale transition. However, the difference between the elastic and plastic energies predicted by the material network and by DNS can be improved by considering the plastic energy as a property of interest z when training in Eqs. (63, 64).

5.6. *Extrapolation for new materials of the interaction-based material network trained by nonlinear training procedure*

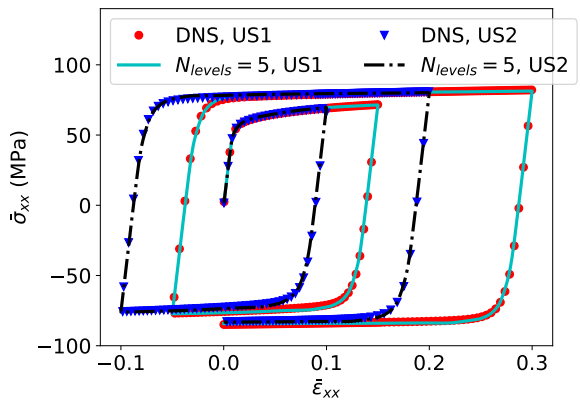
In the nonlinear training (NT) procedure, we consider given constitutive laws governing the physical phases of the microstructure. As a result, the training database is provided by the full-field microscopic simulations with nonlinear paths using these material laws. In this section, we show that the trained material network can extrapolate the RVE response to new material parameters sets with high accuracy. We consider two new sets of material properties respectively denoted as NEW-MAT1 and NEW-MAT2, as reported in Tab. 4. Two different cases are considered:

- NEW-MAT1 is used for the material network with $N_{\text{levels}} = 5$ trained for the 2D-RVE2 with the NT procedure as reported in Section 5.3.1. This trained material network is loaded with the US1 and US2 loading paths as shown in Fig. 6(a).
- NEW-MAT2 is used for the material network with $N_{\text{levels}} = 5$ trained for the 3D-RVE with the NT procedure as reported in Section 5.3.1. This trained material network is loaded with the US3 and RANDOM loading paths as shown in Fig. 14.

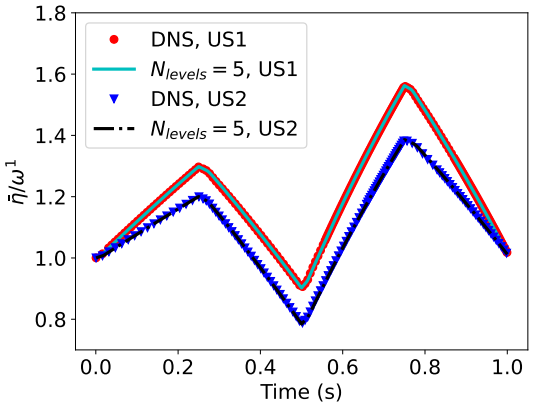
Table 4: New elastoplastic parameters of the solid phase.

Elastoplastic parameters	NEW-MAT1	NEW-MAT2
Young modulus (E [GPa])	10	5
Poisson ratio (ν [-])	0.3	0.25
Isotropic flow stress $\sigma_y(\gamma)$ [MPa] (γ - equivalent plastic strain)	$100\,(1+20\gamma)^{0.1}$	$120\,(1+20\gamma)^{0.15}$

The evolutions of the homogenized stress and of the porosity are reported respectively in Figs. 18 and 19. The corresponding results using the microscopic problem (denoted by direct numerical simulation -DNS) are also reported for comparison purpose. It is shown that the prediction remains accurate when extrapolating beyond the material laws used during the offline training. In a NT procedure, the material network is trained not only with elastic behavior but also with nonlinear behavior. The latter allows the trained material network to better capture the network interactions in comparison with the one trained with a LET procedure.

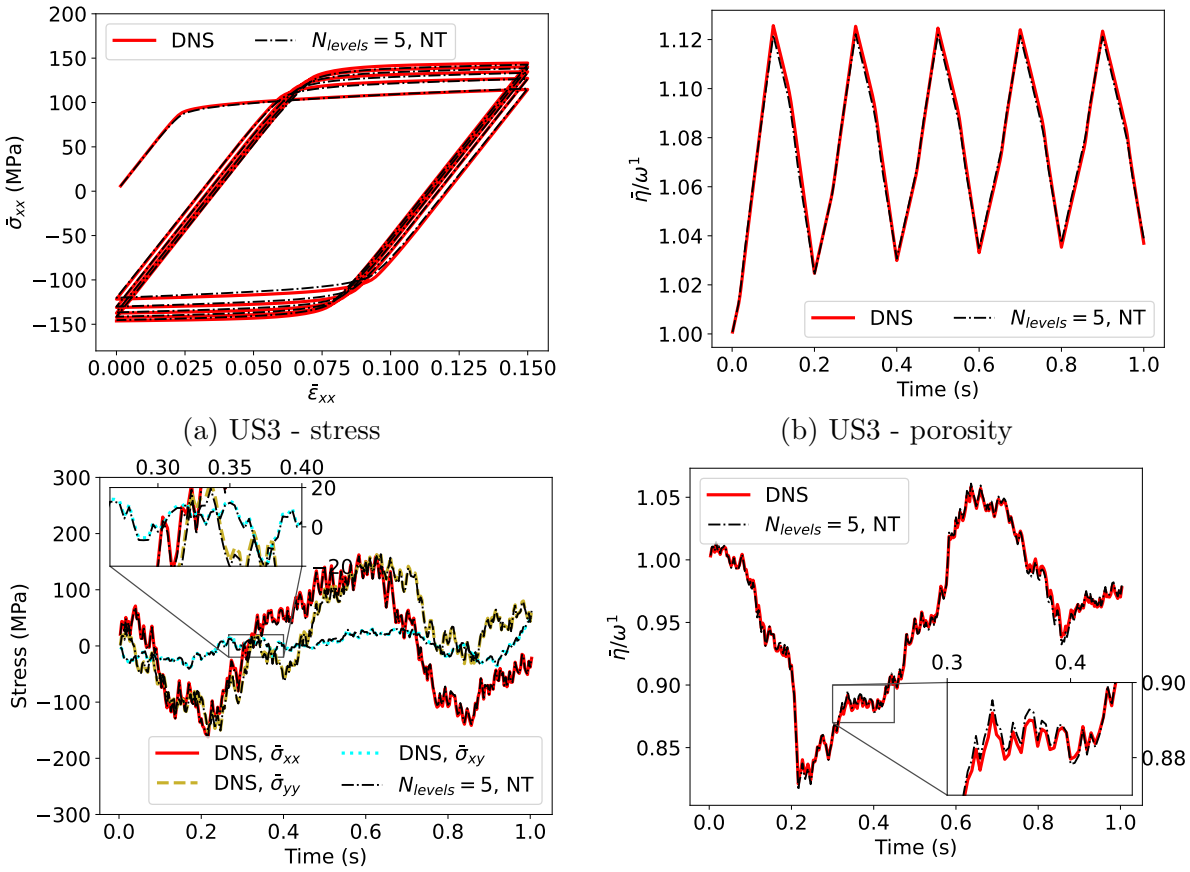


(a) US1 & US2 - stress



(b) US1 & US2 - porosity

Figure 18: 2D-RVE2, NT with NEW-MAT1 - comparison of the results predicted by the material network and by the direct numerical simulation (DNS) under the US1 and US2 loading conditions.



(a) US3 - stress

(b) US3 - porosity

(c) RANDOM - stress

(d) RANDOM - porosity

Figure 19: 3D-RVE, NT with NEW-MAT2 - comparison of the results predicted by the material network and by the direct numerical simulation (DNS) under different loading conditions: US3 (a, b) and RANDOM (c, d).

5.7. Multiscale simulation of a notched sample

The trained material networks can be used as a surrogate of a microscopic model in the corresponding FE^2 simulation. In this section, the simulation of a notched sample under tensile loading is considered. The result obtained with a fully coupled FE^2 simulation is compared to the ones obtained using the material network as a constitutive law in the macro-scale simulation.

The multiscale setting of the notched sample is illustrated in Fig. 20(a). The sample is loaded on its top edge under the prescribed displacement depicted in Fig. 20(b). Because of the geometrical and loading symmetries, only one quarter of the sample is modeled with symmetrical boundary conditions. The microscopic problem considers the single-hole microstructure (2D-RVE1) reported in Fig. 3(a). The corresponding material networks with N_{levels} equal to 3, 4, and 5 were trained using the NT and LET procedures and can be readily used in this section as surrogates of the microscopic problem.

The result in terms of the reaction force versus prescribed displacement is reported in Fig. 21. Fig. 21(a) shows the comparison between the FE-MN result using the material network trained by the NT procedure and the DNS result. While the FE-MN model using the material network with N_{levels} equal to 3 underestimates the reference solution, the cases of N_{levels} equal to 4 and 5 provide good predictions. A similar remark is observed in Fig. 21(b) in which the FE-MN result using the material network trained by the LET procedure is compared to the DNS result. However, a lower accuracy is achieved with the FE-MNs using the material network trained by the LET procedure

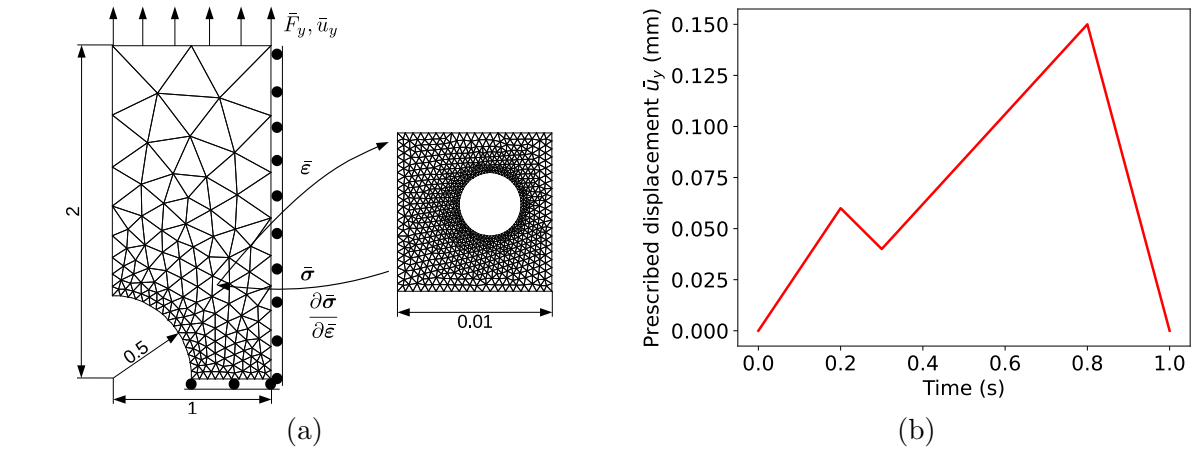


Figure 20: Notched sample: (a) multiscale setting and (b) prescribed displacement at the top boundary of the sample. The microstructure volume element in Fig. 20(a) corresponds to the single-hole microstructure shown in Fig. 3(a).

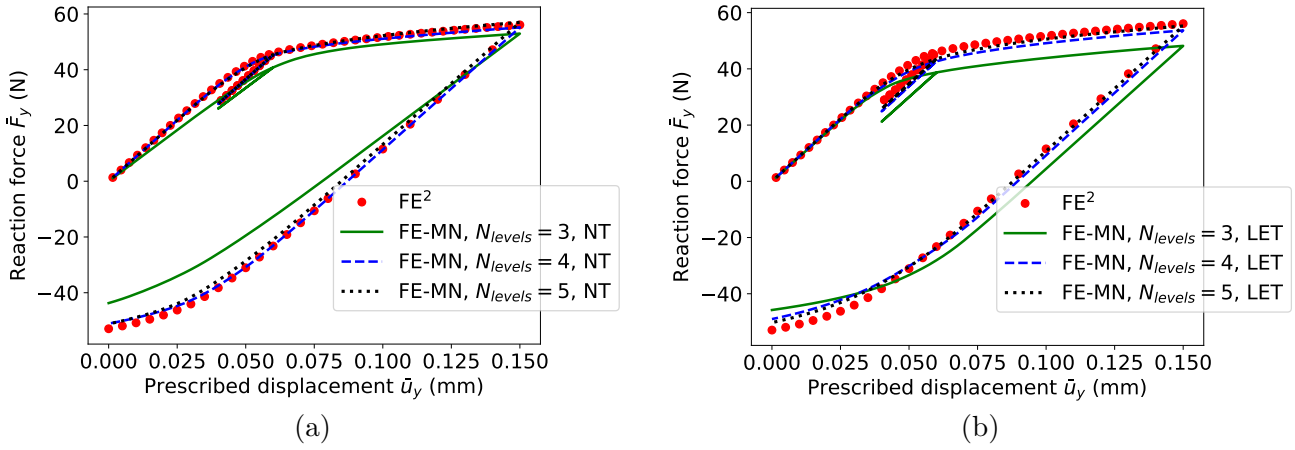


Figure 21: Notched sample - comparison of the reaction force versus prescribed displacement curves between the FE^2 and the finite element simulation (abbreviated by FE-MN) using the material networks trained by: (a) the NT procedure and (b) the LET procedure as surrogates.

with a similar network architecture.

Figure 22 compares the distributions of the macroscopic strain component $\bar{\varepsilon}_{yy}$ and stress component $\bar{\sigma}_{yy}$ obtained with the FE^2 simulation and with the simulations using the material networks trained by the NT procedure as surrogates. It can be seen that the models considering N_{levels} equal to 4 and 5 as surrogates predict the strain and stress distributions in good agreement with the FE^2 simulation while the case of N_{levels} equal to 3 does not. In particular, the case with N_{levels} equal to 5 and FE^2 yield almost the same results. Although the strain is out of the offline sampling space, in which the strain tensor components were limited to 0.2, the prediction remains accurate when extrapolating beyond this offline sampling space, see Fig. 22(a) where the maximal value of $\bar{\varepsilon}_{yy}$ reaches 0.4.

Figure 23 compares the distributions of the macroscopic strain component $\bar{\varepsilon}_{yy}$ and stress component $\bar{\sigma}_{yy}$ obtained with the FE^2 simulation and with the simulations using the material networks trained by the LET procedure as surrogates. It can be seen that the model considering N_{levels} equal to 5 as surrogates predicts the strain and stress distributions in good agreement with the FE^2 simulation while the other cases do not. Overall, the predictions by the material networks trained

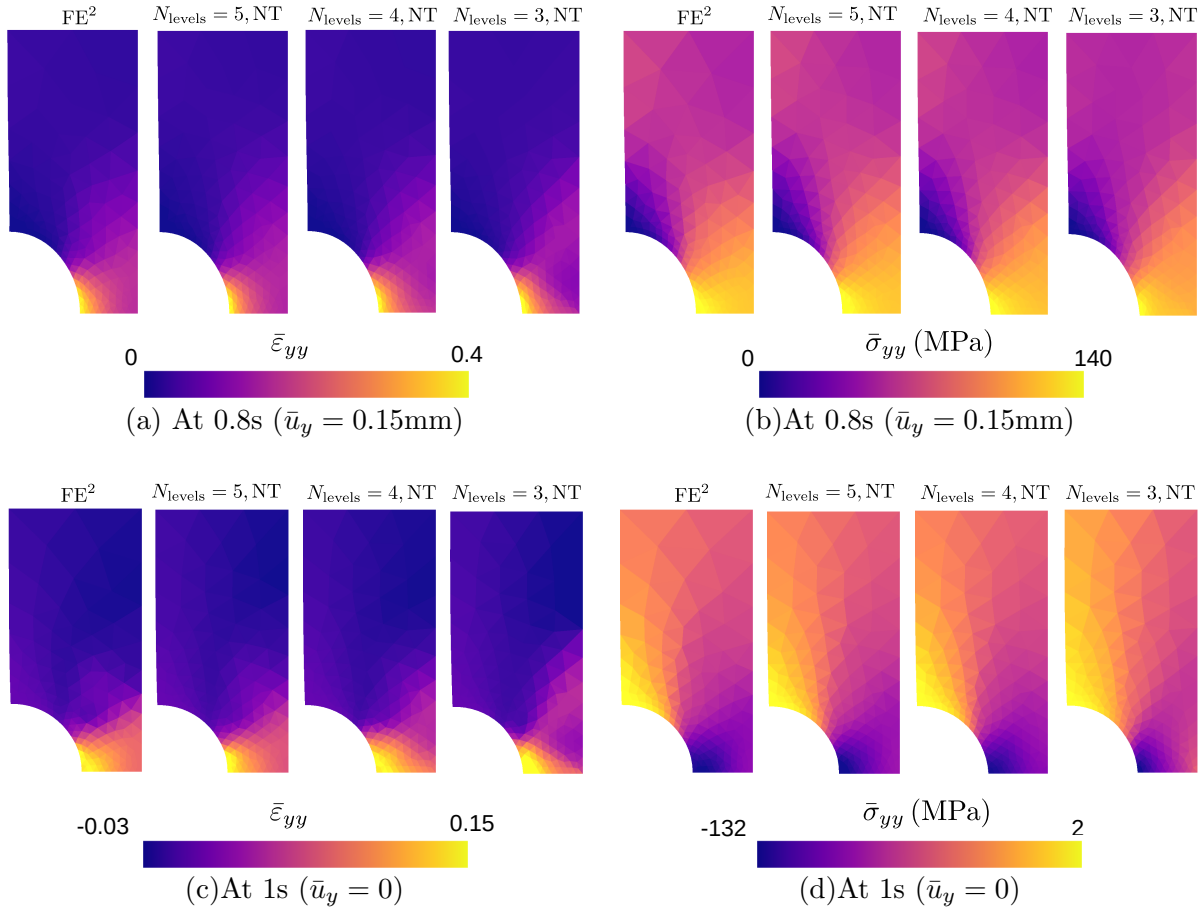


Figure 22: Notched sample - comparison between the FE^2 and the finite element simulation using the material networks trained by the NT procedure as surrogates: (a, c) the distribution of the macroscopic strain $\bar{\varepsilon}_{yy}$, and (b, d) the distribution of the macroscopic stress $\bar{\sigma}_{yy}$ at 0.8s (corresponding to $\bar{u}_y = 0.15$mm) and at 1s (corresponding to $\bar{u}_y = 0$mm), where the specimen is totally unloaded.

by the LET procedure are less accurate than the ones by the material networks trained by the NT procedure.

The computational cost breakdown for multiscale analyzes with the different methodologies is reported in Tab. 5. The computation time of the FE^2 simulation took around 24h using 300 processors on a cluster. The computational efficiency of the finite element simulations using material networks as surrogates is demonstrated. The computational times required for training and online simulations are relatively small compared to the one of the FE^2 simulation.

Table 5: Notched sample - computational cost of the multiscale simulations.

	Offline sampling	Training	Online
FE^2	-	-	7200 hour-cpu
$N_{\text{levels}} = 3$, NT	3.5 hour-cpu	2.3 hour-cpu	0.17 hour-cpu
$N_{\text{levels}} = 4$, NT	3.5 hour-cpu	6.3 hour-cpu	0.31 hour-cpu
$N_{\text{levels}} = 5$, NT	3.5 hour-cpu	19.1 hour-cpu	1.1 hour-cpu
$N_{\text{levels}} = 3$, LET	0.04 hour-cpu	0.16 hour-cpu	0.2 hour-cpu
$N_{\text{levels}} = 4$, LET	0.04 hour-cpu	0.4 hour-cpu	0.4 hour-cpu
$N_{\text{levels}} = 5$, LET	0.04 hour-cpu	1.1 hour-cpu	1 hour-cpu

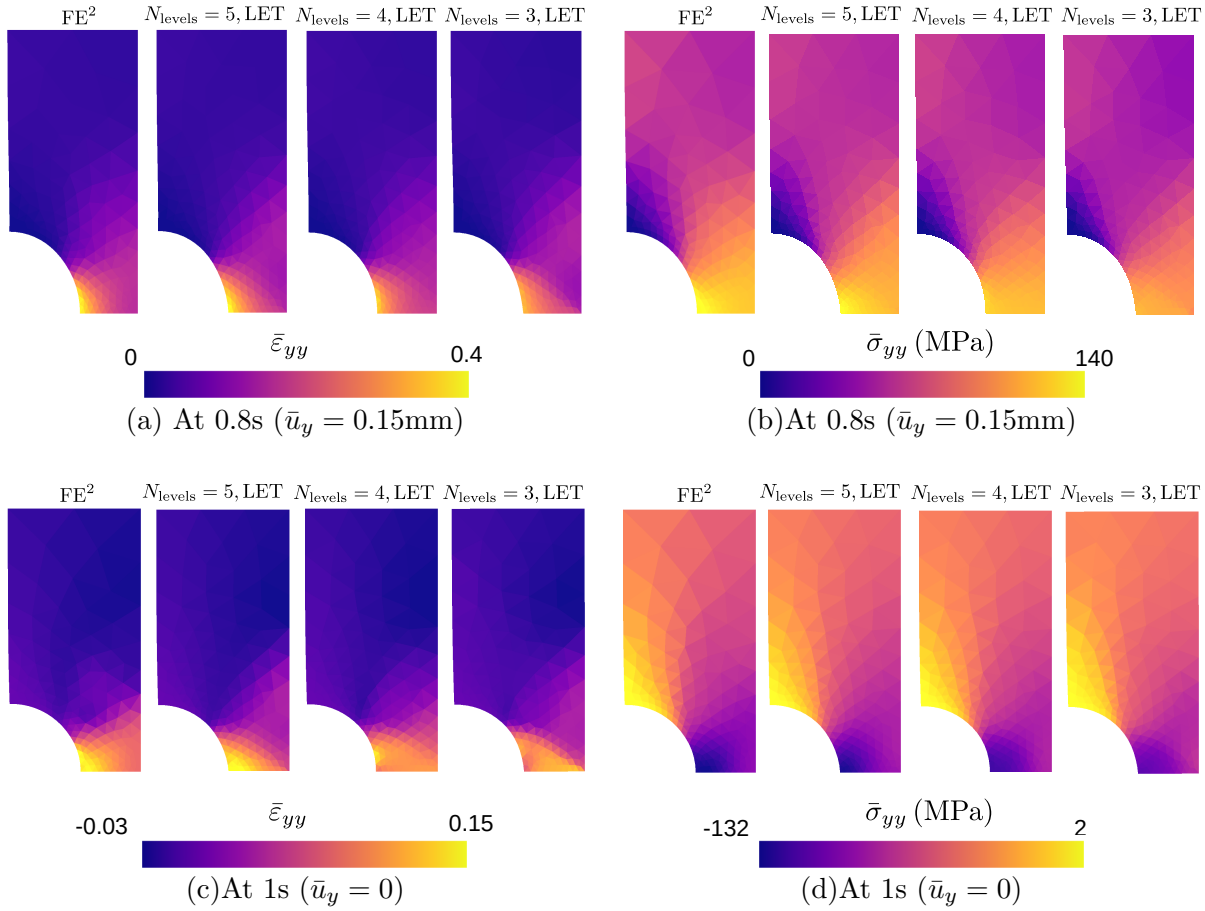


Figure 23: Notched sample - comparison between the FE^2 and the finite element simulation using the material networks trained by the LET procedure as surrogates: (a, c) the distribution of the macroscopic strain $\bar{\varepsilon}_{yy}$, and (b, d) the distribution of the macroscopic stress $\bar{\sigma}_{yy}$ at 0.8s (corresponding to $\bar{u}_y = 0.15$mm) and at 1s (corresponding to $\bar{u}_y = 0$mm), where the specimen is totally unloaded.

6. Conclusion

A general framework for the interaction-based material networks is developed. Based on the network interaction, a material network is viewed as a trainable system involving fitting parameters, which consist of not only the weights of the material nodes but also the parameters that characterize the network interactions. The latter involve not only the directions of the interaction mechanism but also the contribution of the material nodes in each mechanism. We show a complete machine learning process to identify these fitting parameters by minimizing the difference between the predictions of the material network and the corresponding full-field simulation in the context of both the nonlinear training, which is based on the nonlinear response in terms of the stress-strain sequences and the microstructure evolution, and of the linear elastic training strategy, which is based on the elastic response only.

The efficiency and accuracy of the proposed framework are demonstrated on the prediction of the highly nonlinear behaviors of porous microstructure. Once trained, the material network can be used to predict the homogenized response with good accuracy and at much lower computational time as compared to the use of a direct finite element simulation of microscopic BVP. A trained material network can be used as a surrogate of the microscopic model within a FE^2 simulation. It is noted that the proposed framework is not limited to the porous microstructure but can be applied

to a general one.

For a material network constructed by a sufficient number of interaction mechanisms and trained by the nonlinear training procedure, its online prediction extrapolates well the behavior since an accurate prediction can be obtained beyond the training sampling space and beyond the material laws used during training, although only proportional loading paths and a single material model parameters set were using during the nonlinear training. The linear elastic training strategy can be used to train the interaction-based material network at a lower computational cost in comparison with the nonlinear training strategy. However, less accurate predictions could be obtained with this training strategy.

The proposed framework opens new possibilities in the design and concurrent simulation of material systems involving multiple scales, in which case the offline training dataset can be obtained from experiments. In the future, the framework should be extended to more complex scenarios, e.g. failure under coalescence of voids, in which the challenges come from not only the microscopic simulations but also from the modeling of the crack propagation at the macroscopic scale. The proposed framework could extend to the finite strain regime, in which the physical phases obey finite strain constitutive models.

Acknowledgments

Computational resources have been provided by the supercomputing facilities of the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (FRS-FNRS).

Data availability

The raw/processed data required to reproduce these findings is available on [38] under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

Appendix A. Scale transition in computational homogenization

To close the formulation statement of the microscopic problem governed by Eqs. (3, 6), a microscopic boundary condition must be introduced to apply the macroscopic deformation over the RVE through a downscaling process while the macroscopic stress $\bar{\boldsymbol{\sigma}}$ and the corresponding tangent operator $\bar{\mathbf{C}}$ are obtained through an upscaling process.

Appendix A.1. Downscaling

In a kinematically driven framework, at a macroscopic material point, the macroscopic strain $\bar{\boldsymbol{\varepsilon}}$ is known. Assuming the separation of scales [5], the microscopic displacement $\mathbf{u}$ can be decomposed as

$$\mathbf{u} = \bar{\boldsymbol{\varepsilon}} \cdot \mathbf{x} + \mathbf{w} \tag{A.1}$$

where $\mathbf{w}$ is the fluctuation field which is present due to the local inhomogeneities. Once the macroscopic strain $\bar{\boldsymbol{\varepsilon}}$ is known, the strain averaging allows defining the deformation state of the microscopic problem as

$$\bar{\boldsymbol{\varepsilon}} = \langle \boldsymbol{\varepsilon} \rangle_V \ , \tag{A.2}$$

with the volume average operator (7). As a result of Eqs. (3, A.1), one has the decomposition of the local strain

$$\boldsymbol{\varepsilon} = \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} \otimes^s \mathbf{w} \,. \tag{A.3}$$

Using Eq. (A.3) and the Gauss theorem, Eq. (A.2) leads to a constraint of the fluctuation field

$$\int_{\partial V} \mathbf{w} \otimes^s \mathbf{n} \, dA = \mathbf{0} \,, \tag{A.4}$$

where $\mathbf{n}$ is the outward unit normal to the RVE boundary ∂V.

Appendix A.2. Upscaling

The microscopic problem governed by Eqs. (3, 6) coupled with the boundary condition (A.4) can be solved using the finite element method [39]. For upscaling the microscopic response to the macroscopic scale, the energy consistency between the different scales, which corresponds to the Hill-Mandel condition, has to be satisfied, yielding

$$\bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}} \quad = \quad \langle \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} \rangle_V \,, \tag{A.5}$$

in which the macroscopic mechanical power $\bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}}$ is equal to the volume averaging of its respective microscopic counterpart $\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}$. Using Eq. (A.3), Eq. (A.5) can be rewritten as

$$(\bar{\boldsymbol{\sigma}} - \langle \boldsymbol{\sigma} \rangle_V) : \dot{\bar{\boldsymbol{\varepsilon}}} - \langle \boldsymbol{\sigma} : (\dot{\mathbf{w}} \otimes \boldsymbol{\nabla}) \rangle_V = 0 \,. \tag{A.6}$$

Since the last equation should be true for all deformation states, one has the stress averaging relation

$$\bar{\boldsymbol{\sigma}} = \langle \boldsymbol{\sigma} \rangle_V \,, \tag{A.7}$$

from which not only the macroscopic stress but also the tangent operator $\bar{\mathbf{C}} = \dfrac{\partial \bar{\boldsymbol{\sigma}}}{\partial \bar{\boldsymbol{\varepsilon}}}$ can be computed [37]. Using Eq. (A.7) and the Gauss theorem, Eq. (A.6) becomes

$$\int_V \boldsymbol{\sigma} : (\dot{\mathbf{w}} \otimes \boldsymbol{\nabla}) \, dV = 0 \,. \tag{A.8}$$

It is noted that Eq. (A.8) is also the weak form of the microscopic BVP, implying that the finite element solution always satisfies the Hill-Mandel condition.

In general, Eq. (A.4) is satisfied a priori by defining specific microscopic boundary conditions. Several commonly applied boundary conditions are known as the linear displacement boundary condition, periodic boundary condition, and the minimal kinematic boundary condition [39]. In this work, the periodic boundary condition is adopted by constraining the periodicity of the fluctuation field on the boundary ∂V as

$$\mathbf{w}\left(\mathbf{x}^+\right) = \mathbf{w}\left(\mathbf{x}^-\right) \forall \mathbf{x}^+ \in \partial V^+ \text{ and corresponding } \mathbf{x}^- \in \partial V^- \,, \tag{A.9}$$

where the boundary ∂V is separated into non-overlapping opposite surfaces ∂V^+ and ∂V^-. The application of the periodic boundary condition (A.9) for an arbitrary mesh topology follows from the interpolation method proposed in [40].

For a given $\bar{\boldsymbol{\varepsilon}}$ history, the microscopic problem governed by Eqs. (3, 6) coupled with the boundary condition (A.9) is iteratively solved as detailed in [37]. Finally, the extraction of the macroscopic

stress $\bar{\boldsymbol{\sigma}}$ follows from Eq. (A.7). The estimation of the homogenized tangent operator $\bar{\mathbf{C}}$ can be found in the unified framework proposed in [37].

Appendix B. Limitations of the hierarchical material networks based on laminate building blocks in the presence of a void phase.

Let us consider a rotation-free hierarchical material network constructed by a general perfect binary tree of depth K ($K \geq 1$) and consisting of 2^{K-1} leaves as proposed in [32], in which laminates with a variable direction of lamination are used. A similar material network proposed in [28] can be equivalently formed by considering rotation matrices for laminates with a fixed direction of lamination. In this section, we show that such a hierarchical material network results generally in zero-homogenized stress in the plane of a laminate made of the other laminates having a void phase, and that this limits the training ability of the material network.

Appendix B.1. Laminate building block at the first interaction level

Let us consider a 2-phase laminate of 2 materials indexed from 0 to 1 stacking following the direction $\mathbf{n}$. We demote by $\mathcal{V}_1^j$ the list of the 2 material nodes belonging to the first interaction level laminate j. We can define an in-plane unit vector $\boldsymbol{\xi}$, which is parallel to the interface between these phases and satisfies $\boldsymbol{\xi} \cdot \mathbf{n} = 0$, and the out-of-plane unit vector $\boldsymbol{\chi} = \mathbf{n} \times \boldsymbol{\xi}$. As a result, the orthogonal triple $(\mathbf{n}, \boldsymbol{\xi}, \boldsymbol{\chi})$ forms a local coordinate system of the laminate. The volume fractions of phases are $f_1^{j,i}$ with $i = 0, 1$, satisfying $f_1^{j,0} + f_1^{j,1} = 1$. Each phase obeys a linear elastic law as

$$\boldsymbol{\sigma}^i = \mathbf{C}^i : \boldsymbol{\varepsilon}^i \text{ with } i = 0,\, 1\,, \tag{B.1}$$

in which the index i is used to indicate a quantity belonging to the phase i of the laminate, and $\boldsymbol{\varepsilon}$, $\boldsymbol{\sigma}$, and $\mathbf{C}$ denote respectively the strain tensor, the stress tensor, and the elastic tangent operator.

The homogenized stress-strain relation over this laminate can be expressed as

$$\mathcal{H}_{\mathcal{V}_1^j}(\boldsymbol{\sigma}) = \sum_{i=0}^{1} f_1^{j,i} \boldsymbol{\sigma}^i \text{ and } \mathcal{H}_{\mathcal{V}_1^j}(\boldsymbol{\varepsilon}) = \sum_{i=0}^{1} f_1^{j,i} \boldsymbol{\varepsilon}^i\,. \tag{B.2}$$

The stress balance at the interfaces of these phases reads

$$\left(\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1\right) \cdot \mathbf{n} = \mathbf{0}\,. \tag{B.3}$$

In the local coordinate $\left(\mathbf{n}^j,\, \boldsymbol{\xi}^j,\, \boldsymbol{\chi}^j\right)$ of the laminate, Eq. (B.3) can be equivalently rewritten as

$$\begin{cases} \left(\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1\right) : \left(\mathbf{n}^j \otimes \mathbf{n}^j\right) = 0 \\ \left(\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1\right) : \left(\boldsymbol{\xi}^j \otimes \mathbf{n}^j\right) = \left(\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1\right) : \left(\mathbf{n}^j \otimes \boldsymbol{\xi}^j\right) = 0\,, \text{ and} \\ \left(\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1\right) : \left(\boldsymbol{\chi}^j \otimes \mathbf{n}^j\right) = \left(\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1\right) : \left(\mathbf{n}^j \otimes \boldsymbol{\chi}^j\right) = 0\,, \end{cases} \tag{B.4}$$

in which the symmetric properties of the stress tensor is used. The strain compatibility condition between two phases reads

$$\begin{cases} \left(\boldsymbol{\varepsilon}^0 - \boldsymbol{\varepsilon}^1\right) : \left(\boldsymbol{\xi}^j \otimes \boldsymbol{\xi}^j\right) = 0 \\ \left(\boldsymbol{\varepsilon}^0 - \boldsymbol{\varepsilon}^1\right) : \left(\boldsymbol{\xi}^j \otimes \boldsymbol{\chi}^j\right) = \left(\boldsymbol{\varepsilon}^0 - \boldsymbol{\varepsilon}^1\right) : \left(\boldsymbol{\chi}^j \otimes \boldsymbol{\xi}^j\right) = 0\,, \text{ and} \\ \left(\boldsymbol{\varepsilon}^0 - \boldsymbol{\varepsilon}^1\right) : \left(\boldsymbol{\chi}^j \otimes \boldsymbol{\chi}^j\right) = 0\,. \end{cases} \tag{B.5}$$

Equations (B.1, B.2, B.4, B.5) can be solved, and the homogenized response can be found, see [35] for details.

Considering a 2-phase laminate formed in a local coordinate system $\left(\mathbf{n}^j, \boldsymbol{\xi}^j, \boldsymbol{\chi}^j\right)$ and the phase indexed by 1 being void. Since the stress always vanish in the void part, *i.e.* $\boldsymbol{\sigma}^1 = \mathbf{0}$, Eq. (B.3) leads to

$$\boldsymbol{\sigma}^0 \cdot \mathbf{n}^j = \mathbf{0}\,. \tag{B.6}$$

Clearly, these stress states are not three-dimensional. As a result of Eq. (B.4), the homogenized stress of this voided laminate can be rewritten as

$$\mathcal{H}_{\mathcal{V}_1^j}\left(\boldsymbol{\sigma}\right) = \sum_{i=0}^{1} f_1^{j,i}\boldsymbol{\sigma}^i = a\boldsymbol{\xi}^j \otimes \boldsymbol{\xi}^j + b\left(\boldsymbol{\xi}^j \otimes \boldsymbol{\chi}^j + \boldsymbol{\chi}^j \otimes \boldsymbol{\xi}^j\right) + c\boldsymbol{\chi}^j \otimes \boldsymbol{\chi}^j\,, \tag{B.7}$$

in which a, b, c are respectively the components of the homogenized stress $\mathcal{H}_{\mathcal{V}_1^j}\left(\boldsymbol{\sigma}\right)$ in the local coordinate of the laminate.

Appendix B.2. Laminate building block at the second interaction level

Considering a laminate of 2 phases whose homogenized stress-strain relations are obtained from 2 voided laminates, this case corresponding to a material network of 2^2 leaves and is defined by the interaction list $\mathcal{V}_2^l$. This laminate is formed in a local coordinate system $(\mathbf{n}, \boldsymbol{\xi}, \boldsymbol{\chi})$. The voided laminates indexed by 0 and 1 are built from the interaction sets $\mathcal{V}_1^j$ and $\mathcal{V}_1^{j+1}$, and are defined in the $\left(\mathbf{n}^j, \boldsymbol{\xi}^j, \boldsymbol{\chi}\right)$, ... and $\left(\mathbf{n}^{j+1}, \boldsymbol{\xi}^{j+1}, \boldsymbol{\chi}\right)$ referentials, respectively. Indeed the 2 laminates are always co-planar so that one direction (the normal $\boldsymbol{\chi}$ to the plane) is unique. Eq. (B.7) is then written for each voided laminate as

$$\mathcal{H}_{\mathcal{V}_1^{j+k}}\left(\boldsymbol{\sigma}\right) \;=\; a^k\boldsymbol{\xi}^{j+k} \otimes \boldsymbol{\xi}^{j+k} + b^k\left(\boldsymbol{\xi}^{j+k} \otimes \boldsymbol{\chi} + \boldsymbol{\chi} \otimes \boldsymbol{\xi}^{j+k}\right) + c^k\boldsymbol{\chi} \otimes \boldsymbol{\chi} \text{ for } k = 0,\, 1\,. \tag{B.8}$$

In this context, Eq. (B.3) rewritten at the interaction level 2 becomes

$$\left(a^0\boldsymbol{\xi}^j + b^0\boldsymbol{\chi}\right)\left(\boldsymbol{\xi}^j \cdot \mathbf{n}\right) - \left(a^1\boldsymbol{\xi}^{j+1} + b^1\boldsymbol{\chi}\right)\left(\boldsymbol{\xi}^{j+1} \cdot \mathbf{n}\right) = \mathbf{0}\,, \tag{B.9}$$

since $\boldsymbol{\chi} \cdot \mathbf{n} = 0$. Applying the dot product of Eq. (B.9) with $\mathbf{n}^{j+1}$ and $\mathbf{n}^j$ and using $\boldsymbol{\chi} \cdot \mathbf{n}^j = \boldsymbol{\xi}^j \cdot \mathbf{n}^j = \boldsymbol{\chi} \cdot \mathbf{n}^{j+1} = \boldsymbol{\xi}^{j+1} \cdot \mathbf{n}^{j+1} = 0$, one has

$$\begin{cases} a^0\left(\boldsymbol{\xi}^j \cdot \mathbf{n}^{j+1}\right)\left(\boldsymbol{\xi}^j \cdot \mathbf{n}\right) = 0 \text{ and} \\ a^1\left(\boldsymbol{\xi}^{j+1} \cdot \mathbf{n}^j\right)\left(\boldsymbol{\xi}^{j+1} \cdot \mathbf{n}\right) = 0 \end{cases}\,. \tag{B.10}$$

Since the laminate directions in a hierarchical material network are randomly initialized when training [28, 29, 32, 35], one has generally $\boldsymbol{\xi}^j \cdot \mathbf{n}^{j+1} \neq 0$, $\boldsymbol{\xi}^j \cdot \mathbf{n} \neq 0$, $\boldsymbol{\xi}^{j+1} \cdot \mathbf{n}^j \neq 0$, and $\boldsymbol{\xi}^{j+1} \cdot \mathbf{n} \neq 0$, leading to

$$a^0 = a^1 = 0\,. \tag{B.11}$$

In this case, the homogenized stress of this laminate is expressed by

$$\mathcal{H}_{\mathcal{V}_2^l}\left(\boldsymbol{\sigma}\right) = \sum_{k=0}^{1} f_2^{l,k}\mathcal{H}_{\mathcal{V}_1^{j+k}}\left(\boldsymbol{\sigma}\right) = \sum_{k=0}^{1} f_2^{l,k}b^k\left(\boldsymbol{\xi}^{j+k} \otimes \boldsymbol{\chi} + \boldsymbol{\chi} \otimes \boldsymbol{\xi}^{j+k}\right) + \left(\sum_{k=0}^{1} f_2^{l,k}c^i\right)\boldsymbol{\chi} \otimes \boldsymbol{\chi}\,, \tag{B.12}$$

which consists of only the out-of-plane components. As a result of Eq. (B.12), all in-plane stress components of a laminate of a laminate are equal to zero at the second level.

Appendix B.3. Training of a binary laminate-based material network

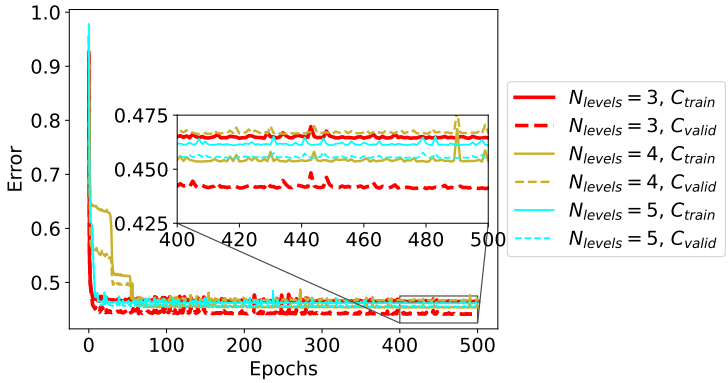


Figure B.24: Material networks based on a binary architecture using laminate building blocks with voids - histories of the training error C_{train} and validation error C_{valid} for 3D-RVE reported in Fig. 3(c).

In a general three-dimensional case, a hierarchical material network considering one of its phases being void can still stand a certain amount of stress. However, because of the fact that the in-plane stress components of a laminate of a voided laminate are equal to zero the training cannot be optimized, as illustrated by the histories of the training error and validation error shown in Fig. B.24. In these material networks, voids are considered as a physical phase without any stiffness. The training is performed with the offline data reported in the LET training for the case of 3D-RVE in Section 5.3.2. It is shown that the training and validation errors stabilize at relative high values (> 0.4 for all cases) while in Fig. 5 much smaller values (< 0.04 for all cases) were reported with the new interaction-based approach. This poor training capability of the material networks based on a binary architecture using laminate building blocks with voids can be explained by too restrictive stress state inside each material node.

Appendix C. Material network evaluation

In this section, the iterative procedure proposed in [35] to solve the governing equation (37) of the material network $\mathcal{M}$ indexed in $\mathcal{K}$ is summarized in the context of the small strain setting. Assuming its weights and its network interactions have been determined, the unknowns involve only the M vectors $\mathbf{a}^j$ considered in Eq. (22). These unknown vectors are collected in a column vector denoted by $\mathbf{U}$ as

$$\mathbf{U} = \left[\left(\mathbf{a}^j\right)^T \text{ for } j = 0, \dots, M-1 \right]^T . \tag{C.1}$$

The state of the material network is driven by the homogenized strain history $\bar{\boldsymbol{\varepsilon}}(t)$. By knowing the values of $\mathbf{U}(t - \Delta t)$ and the history data at all the material nodes $\mathbf{q}^i(t - \Delta t)$, $\forall i \in \mathcal{K}$, at the previous converged solution (at time $t - \Delta t$), the response of the material network at time t with a time step Δt for the current value of $\bar{\boldsymbol{\varepsilon}}(t)$ is first computed. Then the outputs consist of the homogenized stress $\bar{\boldsymbol{\sigma}}(t)$ and the tangent operator of the homogenized solution

$$\bar{\mathbf{C}}(t) = \frac{\partial \bar{\boldsymbol{\sigma}}(t)}{\partial \bar{\boldsymbol{\varepsilon}}(t)} . \tag{C.2}$$

The computation of $\bar{\mathbf{C}}$ is mandatory when integrating the material network in a multiscale finite element setting as a surrogate model predicting the local stress-strain behavior. In the following, a quantity at time t is written without time argument for ease of readability. The iterative resolution follows a downscaling procedure, a nonlinear system resolution, and an upscaling procedure as follows:

Appendix C.1. Downscaling

The homogenized strain $\bar{\boldsymbol{\varepsilon}}$ is downscaled. With two arbitrary vectors $\mathbf{a}$ and $\mathbf{n}$, we can define an operator $\mathbf{H}(\bullet)$ to convert a dyadic tensor product into a matrix-vector multiplication as

$$\mathrm{vec}^{\varepsilon}(\mathbf{a} \otimes^{s} \mathbf{n}) = \mathbf{H}(\mathbf{n})\,\mathbf{a} = \mathbf{H}(\mathbf{a})\,\mathbf{n}\,, \tag{C.3}$$

where the operator $\mathrm{vec}^{\varepsilon}(\bullet)$ is defined by

$$\mathrm{vec}^{\varepsilon}(\boldsymbol{\varepsilon}) = \begin{bmatrix} \varepsilon_{00} & \varepsilon_{11} & \varepsilon_{22} & 2\varepsilon_{01} & 2\varepsilon_{02} & 2\varepsilon_{12} \end{bmatrix}^{T}, \tag{C.4}$$

and the operator $\mathbf{H}(\bullet)$ is defined by

$$\mathbf{H}(\mathbf{n}) = \begin{bmatrix} n_0 & 0 & 0 \\ 0 & n_1 & 0 \\ 0 & 0 & n_2 \\ n_1 & n_0 & 0 \\ n_2 & 0 & n_0 \\ 0 & n_2 & n_1 \end{bmatrix}. \tag{C.5}$$

Consequently, the interaction mapping described by Eq. (22) can be rewritten for a material node $i \in \mathcal{K}$ under the vector form using the unknown vector $\mathbf{U}$ defined by Eq. (C.1), yielding

$$\mathrm{vec}^{\varepsilon}(\boldsymbol{\varepsilon}^{i}) = \mathrm{vec}^{\varepsilon}(\bar{\boldsymbol{\varepsilon}}) + \mathbf{D}^{i}\mathbf{U}\ \forall i \in \mathcal{K}\,, \tag{C.6}$$

where $\mathbf{D}^{i}$ is a row-block matrix depending only on the weights and interaction directions,

$$\mathbf{D}^{i} = \left[\mathbf{Z}^{i,j} \text{ for } j = 0, \ldots, M-1\right] \text{ with } \mathbf{Z}^{i,j} = \begin{cases} \alpha^{i,j}\mathbf{H}(\mathbf{n}^{j}) & \text{if } i \in \mathcal{V}^{j} \\ \mathbf{0}_{6\times 3} & \text{if } i \notin \mathcal{V}^{j} \end{cases}. \tag{C.7}$$

In the last equation, $\mathbf{0}_{6\times 3}$ denotes a 6×3 zero matrix.

From the local deformation gradient estimated by Eq. (C.6) at each material node, the material constitutive law associated to this node as defined by Eq. (20) is used to estimate the local stress $\boldsymbol{\sigma}^{i}$ and the internal variables $\mathbf{q}^{i}$ with an appropriate integration scheme in the time interval $[t-\Delta t\,, t]$ from the current strain $\boldsymbol{\varepsilon}^{i}$ and the internal variables of the previous converged solution $\mathbf{q}^{i}(t-\Delta t)$:

$$\begin{cases} \boldsymbol{\sigma}^{i} & = \hat{\mathfrak{P}}^{p_i}\left(\boldsymbol{\varepsilon}^{i}, \mathbf{q}^{i}(t-\Delta t)\right) \\ \mathbf{q}^{i} & = \hat{\mathfrak{Q}}^{p_i}\left(\boldsymbol{\varepsilon}^{i}, \mathbf{q}^{i}(t-\Delta t)\right) \end{cases} \forall i \in \mathcal{K}\,, \tag{C.8}$$

where t and Δt denote respectively the current time and time step. The local tangent operator $\mathbf{C}^{i} = \frac{\partial \boldsymbol{\sigma}^{i}}{\partial \boldsymbol{\varepsilon}^{i}}$ is also computed.

Appendix C.2. Nonlinear system resolution

The system of equations (37) is iteratively solved. Using the matrix $\mathbf{D}^i$ expressed in Eq. (C.7), the residual vector of the system of $3M$ equations described by the system (37) can be rewritten as

$$\mathbf{r} = \sum_{i \in \mathcal{K}} W^i \left(\mathbf{D}^i\right)^T \mathrm{vec}^\sigma \left(\boldsymbol{\sigma}^i\right) , \tag{C.9}$$

where the operator $\mathrm{vec}^\sigma(\bullet)$ is defined by

$$\mathrm{vec}^\sigma \left(\bar{\boldsymbol{\sigma}}\right) = \begin{bmatrix} \sigma_{00} & \sigma_{11} & \sigma_{22} & \sigma_{01} & \sigma_{02} & \sigma_{12} \end{bmatrix}^T . \tag{C.10}$$

The convergence is achieved if the following condition is satisfied:

$$||\mathbf{r}||_\infty < \tau^{\mathrm{abs}} \text{ or } ||\mathbf{r}||_\infty < \tau^{\mathrm{rel}} ||\mathbf{r}_0||_\infty , \tag{C.11}$$

where $||\bullet||_\infty$ represents the infinity norm operator, τ^{abs} and τ^{rel} are respectively the absolute and relative tolerances, and $\mathbf{r}_0$ is the initial residual. In this work, $\tau^{\mathrm{abs}} = 10^{-12}$ and $\tau^{\mathrm{rel}} = 10^{-6}$ are considered.

If the condition (C.11) is not satisfied, the unknown vector $\mathbf{U}$ is corrected with

$$\delta\mathbf{U} = -\mathbf{K}^{-1}\mathbf{r} \text{ and } \mathbf{U} \leftarrow \mathbf{U} + \delta\mathbf{U} , \tag{C.12}$$

where $\mathbf{K}$ is the Jacobian matrix

$$\mathbf{K} = \frac{\partial \mathbf{r}}{\partial \mathbf{U}} . \tag{C.13}$$

Using Eqs. (C.6, C.8, C.9), the estimation of $\mathbf{K}$ following Eq. (C.13) is given by

$$\mathbf{K} = \sum_{i \in \mathcal{K}} W^i \left(\mathbf{D}^i\right)^T \mathrm{mat}\left(\mathbf{C}^i\right) \mathbf{D}^i , \tag{C.14}$$

where $\mathrm{mat}(\bullet)$ is the matrix representation operator of a fourth order tensor as

$$\mathrm{mat}\left(\mathbf{C}\right) = \begin{bmatrix} C_{0000} & C_{0011} & C_{0022} & C_{0001} & C_{0002} & C_{0012} \\ C_{1100} & C_{1111} & C_{1122} & C_{1101} & C_{1102} & C_{1112} \\ C_{2200} & C_{2211} & C_{2222} & C_{2201} & C_{2202} & C_{2212} \\ C_{0100} & C_{0111} & C_{0122} & C_{0101} & C_{0102} & C_{0112} \\ C_{0200} & C_{0211} & C_{0222} & C_{0201} & C_{0202} & C_{0212} \\ C_{1200} & C_{1211} & C_{1222} & C_{1201} & C_{1202} & C_{1212} \end{bmatrix} . \tag{C.15}$$

Appendix C.3. Upscaling

The homogenized stress tensor $\bar{\boldsymbol{\sigma}}$ is computed by Eq. (19). If the material network is considered in an iterative multiscale simulation as a constitutive law, the homogenized tangent operator $\bar{\mathbf{C}}$ also needs to be estimated following Eq. (C.2). Using Eqs. (19, C.6), Eq. (C.2) yields

$$\mathrm{mat}\left(\bar{\mathbf{C}}\right) = \frac{1}{\sum_{i \in \mathcal{K}} W^i} \sum_{i \in \mathcal{K}} W^i \mathrm{mat}\left(\mathbf{C}^i\right) + \mathbf{Y} \frac{\partial \mathbf{U}}{\partial \mathrm{vec}^\varepsilon \left(\bar{\boldsymbol{\varepsilon}}\right)} , \tag{C.16}$$

where

$$\mathbf{Y} = \frac{1}{\sum_{i\in\mathcal{K}} W^i} \sum_{i\in\mathcal{K}} W^i \mathrm{mat}\left(\mathbf{C}^i\right) \mathbf{D}^i \,. \tag{C.17}$$

In order to estimate $\bar{\mathbf{C}}$ following Eq. (C.16), the matrix $\dfrac{\partial \mathbf{U}}{\partial \mathrm{vec}^{\varepsilon}\left(\bar{\boldsymbol{\varepsilon}}\right)}$ needs to be computed. To this end, the consistency of Eq. (C.9) is expressed as

$$\delta\mathbf{r} = \sum_{i\in\mathcal{K}} W^i \left(\mathbf{D}^i\right)^T \mathrm{mat}\left(\mathbf{C}^i\right) \left(\mathrm{vec}^{\varepsilon}\left(\delta\bar{\boldsymbol{\varepsilon}}\right) + \mathbf{D}^i \delta\mathbf{U}\right) = \mathbf{0} \,. \tag{C.18}$$

This last equation allows the computation of $\dfrac{\partial \mathbf{U}}{\partial \mathrm{vec}^{\varepsilon}\left(\bar{\boldsymbol{\varepsilon}}\right)}$ by

$$\frac{\partial \mathbf{U}}{\partial \mathrm{vec}^{\varepsilon}\left(\bar{\boldsymbol{\varepsilon}}\right)} = -\mathbf{K}^{-1}\mathbf{M} \,, \tag{C.19}$$

where $\mathbf{K}$ is given by Eq. (C.14) and

$$\mathbf{M} = \sum_{i\in\mathcal{K}} W^i \left(\mathbf{D}^i\right)^T \mathrm{mat}\left(\mathbf{C}^i\right) \,. \tag{C.20}$$

Equation (C.19) corresponds to a set of linear systems, one for each column of the right hand side matrix, and whose matrix $\mathbf{K}$ was previously factorized in order to solve the system of Eq. (C.12). The resolution of the set of linear systems (C.19) is then performed using this factorized matrix at a reduced computational time.

Appendix C.4. Summary

For a material network $\mathcal{M}$ indexed in $\mathcal{K}$, and assuming its weights and its network interactions are known, the resolution framework [35] for estimating its response is summarized as follows:

(I) Initialization:

- (i) assemble $\mathbf{U}$ following Eq. (C.1);
- (ii) initialize $\mathbf{U} = \mathbf{0}$;
- (iii) assemble $\mathbf{D}^i$ $\forall i \in \mathcal{K}$ following Eq. (C.7).

(II) Evaluation at time t with a time step Δt for the current value of $\bar{\boldsymbol{\varepsilon}}\left(t\right)$, and knowing the values of $\mathbf{U}\left(t-\Delta t\right)$ and the history data at material nodes $\mathbf{q}^i\left(t-\Delta t\right)$, $\forall i \in \mathcal{K}$, at the previous converged solution (at time $t-\Delta t$):

- (i) initialize $\mathbf{U}(t) = \mathbf{U}\left(t-\Delta t\right)$ and $\mathbf{q}^i\left(t\right) = \mathbf{q}^i\left(t-\Delta t\right)$, $\forall i \in \mathcal{K}$;
- (ii) downscale $\bar{\boldsymbol{\varepsilon}}(t)$ and $\mathbf{U}(t)$ following Eq. (C.6);
- (iii) evaluate the local constitutive law at all material nodes following Eq. (C.8), yielding $\boldsymbol{\sigma}^i\left(t\right)$, $\mathbf{q}^i\left(t\right)$, and $\mathbf{C}^i\left(t\right)$, $\forall i \in \mathcal{K}$;
- (iv) evaluate the residual $\mathbf{r}$ following Eq. (C.9);
- (v) **if** the convergence criterion following Eq. (C.11) is achieved **go to** (viii); **else go to** (vi);
- (vi) correct the value of $\mathbf{U}(t)$ following Eq. (C.12);
- (vii) go to (ii);

(viii) compute the homogenized stress $\bar{\boldsymbol{\sigma}}(t)$ following Eq. (19);
(ix) compute the homogenized tangent operator $\bar{\mathbf{C}}(t)$ following Eq. (C.16) if required;
(x) store $\mathbf{U}(t)$ and history data $\mathbf{q}^i(t)$ at all material nodes for the next step;
(xi) exit.

Appendix D. Computation of the gradient of the loss function

In this section, the computation of the gradient of the cost function $C^{[l]}(\mathbf{X})$ given by Eq. (68) with respect to the fitting parameter $\mathbf{X}$ is detailed. In the following, the index l is omitted for simplicity. One has

$$\mathbf{g} = \frac{\partial \epsilon^{\text{rel}}\left(\widehat{\boldsymbol{\mathcal{S}}}, \bar{\boldsymbol{\mathcal{S}}}\right)}{\partial \mathbf{X}} + \sum_{q=0}^{L_{\bar{\mathcal{Z}}}-1} \lambda^q \frac{\partial \epsilon^{\text{rel}}\left(\widehat{\mathcal{Z}}^q, \bar{\mathcal{Z}}^q\right)}{\partial \mathbf{X}} \,. \tag{D.1}$$

Using Eq. (69), the last equation can be rewritten as

$$\begin{aligned} \mathbf{g} = \frac{1}{\text{len}\left(\widehat{\boldsymbol{\mathcal{S}}}\right)} \Bigg(& \sum_{k=0}^{\text{len}\left(\widehat{\boldsymbol{\mathcal{S}}}\right)-1} \frac{1}{||\hat{\boldsymbol{\sigma}}_k||_2} \frac{\partial ||\hat{\boldsymbol{\sigma}}_k - \bar{\boldsymbol{\sigma}}_k||_2}{\partial \bar{\boldsymbol{\sigma}}_k} : \frac{\partial \bar{\boldsymbol{\sigma}}_k}{\partial \mathbf{X}} \\ & + \sum_{q=0}^{L_{\bar{\mathcal{Z}}}-1} \lambda^q \sum_{k=0}^{\text{len}\left(\widehat{\boldsymbol{\mathcal{S}}}\right)-1} \frac{1}{||\hat{z}_k^q||_2} \frac{\partial ||\hat{z}_k^q - \bar{z}_k^q||_2}{\partial \bar{z}_k^q} \frac{\partial \bar{z}_k^q}{\partial \mathbf{X}} \Bigg) \,, \end{aligned} \tag{D.2}$$

where $\dfrac{\partial ||\hat{\boldsymbol{\sigma}}_k - \bar{\boldsymbol{\sigma}}_k||_2}{\partial \bar{\boldsymbol{\sigma}}_k}$ and $\dfrac{\partial ||\hat{z}_k^q - \bar{z}_k^q||_2}{\partial \bar{z}_k^q}$ are known using Eq. (70) as

$$\frac{\partial ||\bullet||_2}{\partial \bullet} = \begin{cases} \dfrac{|\bullet|}{\bullet} & \text{if } \bullet \text{ is a scalar}\,, \\ (||\bullet||_2)^{-1} \bullet & \text{if } \bullet \text{ is a vector}\,, \text{ and} \\ (||\bullet||_2)^{-1} \bullet & \text{if } \bullet \text{ is a second-order tensor}\,, \end{cases} \tag{D.3}$$

for an arbitrary quantity $\bullet$. As a result, to compute $\mathbf{g}$, we have to compute

$$\frac{\partial \bar{\boldsymbol{\sigma}}_k}{\partial \mathbf{X}} \text{ and } \frac{\partial \bar{z}_k^q}{\partial \mathbf{X}} \text{ for } k = 0, \ldots, \text{len}\left(\widehat{\boldsymbol{\mathcal{S}}}\right) - 1 \text{ and } q = 0, \ldots, L_{\bar{\mathcal{Z}}} - 1 \,. \tag{D.4}$$

Since $\mathbf{X}$ consists of $\widetilde{\mathbf{W}}$, $\widetilde{\mathbf{A}}$, and $\widetilde{\mathbf{N}}$ as described in Section 3.3, one needs to estimate

$$\begin{aligned} & \frac{\partial \bar{\boldsymbol{\sigma}}_k}{\partial \widetilde{\mathbf{W}}}, \frac{\partial \bar{\boldsymbol{\sigma}}_k}{\partial \widetilde{\mathbf{A}}}, \frac{\partial \bar{\boldsymbol{\sigma}}_k}{\partial \widetilde{\mathbf{N}}}, \frac{\partial \bar{z}_k^q}{\partial \widetilde{\mathbf{W}}}, \frac{\partial \bar{z}_k^q}{\partial \widetilde{\mathbf{A}}}, \text{ and } \frac{\partial \bar{z}_k^q}{\partial \widetilde{\mathbf{N}}} \\ & \text{for } k = 0, \ldots, \text{len}\left(\widehat{\boldsymbol{\mathcal{S}}}\right) - 1 \text{ and } q = 0, \ldots, L_{\bar{\mathcal{Z}}} - 1 \,. \end{aligned} \tag{D.5}$$

Without loss of generality, we consider a quantity $\bar{\mathbf{v}}$ which can be extracted from the material network $\mathcal{M}$ using Eq. (38) as

$$\bar{\mathbf{v}} = \mathcal{H}_{\mathcal{C}}(\mathbf{v}) \,, \tag{D.6}$$

in which $\mathcal{C}$ can be $\mathcal{K}$ for averaging over the whole material networks or $\mathcal{K}^p$ if only the average on the phase indexed by p is of interest. At a material node i, the quantity $\mathbf{v}^i$ is defined $\forall i \in \mathcal{C}$ and its tangent with respect to the local strain, *i.e.* $\mathbf{G}^i = \dfrac{\partial \mathbf{v}^i}{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^i\right)}$, is known. The computation procedure for

$$\frac{\partial \bar{\mathbf{v}}}{\partial \widetilde{\mathbf{W}}}\,, \frac{\partial \bar{\mathbf{v}}}{\partial \widetilde{\mathbf{A}}}\,, \text{ and } \frac{\partial \bar{\mathbf{v}}}{\partial \widetilde{\mathbf{N}}}\,, \tag{D.7}$$

is detailed here below. The results of Eq. (D.5) can be directly obtained by substituting $\bar{\mathbf{v}}$ by either $\mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}\right)$ or z^q and respectively either setting $\mathcal{C} = \mathcal{K}$ or particularizing $\mathcal{C}$ depending on the context. Moreover, we consider the following cases:

- $\mathbf{v}^i = \mathrm{vec}^{\sigma}\left(\boldsymbol{\sigma}^i\right)$, leading to $\mathbf{G}^i = \mathrm{mat}\left(\mathbf{C}^i\right)$.
- $\mathbf{v}^i = 1 + \mathrm{tr}\left(\boldsymbol{\varepsilon}^i\right)$, which is used to quantify the porosity evolution in a porous microstructure, leading to

$$\mathbf{G}^i = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{D.8}$$

Appendix D.1. Computation of $\dfrac{\partial \bar{\mathbf{v}}}{\partial \widetilde{\mathbf{W}}}$

For each component $Z^l \in \widetilde{\mathbf{W}}$ $\forall l \in \mathcal{K}$, one has the following chain rule

$$\frac{\partial \bar{\mathbf{v}}}{\partial Z^l} = \sum_{i \in \mathcal{C}} \frac{\partial \bar{\mathbf{v}}}{\partial W^i} \frac{\partial W^i}{\partial Z^l}\,. \tag{D.9}$$

As a result, the computations of $\dfrac{\partial \bar{\mathbf{v}}}{\partial W^i}$ and of $\dfrac{\partial W^i}{\partial Z^l}$ are required. Using Eq. (50), one has

$$\frac{\partial W^i}{\partial Z^l} = \begin{cases} \dfrac{\omega^p}{\sum_{c \in \mathcal{K}^p} f\left(Z^c\right)} \left(1 - \dfrac{f\left(Z^i\right)}{\sum_{c \in \mathcal{K}^p} f\left(Z^c\right)}\right) f'\left(Z^l\right) & \text{if } l = i\,, \\ -\omega^p \dfrac{f\left(Z^i\right)}{\left(\sum_{c \in \mathcal{K}^p} f\left(Z^c\right)\right)^2} f'\left(Z^l\right) & \text{if } l \neq i \text{ and } l \in \mathcal{K}^p\,, \text{ and} \\ 0 & \text{otherwise}\,, \end{cases} \tag{D.10}$$

for $i \in \mathcal{K}^p$ and $p = 0, \ldots, P-1$,

where f' is the derivative of the activation function defined by Eq. (51), yielding

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{relu}^{\mathrm{s}}\left(x\right) = \frac{\mathrm{e}^{sx}}{1 + \mathrm{e}^{sx}}\,. \tag{D.11}$$

Following Eqs. (38, D.6), one has

$$\frac{\partial \bar{\mathbf{v}}}{\partial W^i} = \begin{cases} \dfrac{\mathbf{v}^i - \bar{\mathbf{v}}}{\sum_{c \in \mathcal{C}} W^c} + \dfrac{1}{\sum_{c \in \mathcal{C}} W^c} \sum_{l \in \mathcal{C}} W^l \mathbf{G}^l \dfrac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^l\right)}{\partial W^i} & \text{if } i \in \mathcal{C}\,, \\ \dfrac{1}{\sum_{c \in \mathcal{C}} W^c} \sum_{l \in \mathcal{C}} W^l \mathbf{G}^l \dfrac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^l\right)}{\partial W^i} & \text{otherwise}\,. \end{cases} \tag{D.12}$$

Eqs. (33, C.6) lead to

$$\frac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^{l}\right)}{\partial W^{i}} = \sum_{j \in \mathcal{T}^{l}} \frac{\partial \alpha^{l,j}}{\partial W^{i}} \mathrm{vec}^{\varepsilon}\left(\mathbf{a}^{j} \otimes^{s} \mathbf{n}^{j}\right) + \mathbf{D}^{l} \frac{\partial \mathbf{U}}{\partial W^{i}}\,, \tag{D.13}$$

where $\dfrac{\partial \alpha^{l,j}}{\partial W^{i}} = 0$ for $l \notin \mathcal{V}^{j}$ or $i \notin \mathcal{V}^{j}$ and $\dfrac{\partial \alpha^{l,j}}{\partial W^{i}}$ for $l \in \mathcal{V}^{j}$ and $i \in \mathcal{V}^{j}$ are estimated using Eq. (54) as

$$\frac{\partial \alpha^{l,j}}{\partial W^{i}} = \begin{cases} 0 & \text{if } \mathrm{loc}\left(l, \mathcal{V}^{j}\right) = 0\,, \\ -\dfrac{\alpha^{\mathcal{V}_0^j, j}}{W^{l}} \dfrac{\beta^{\mathrm{loc}\left(l, \mathcal{V}^{j}\right)-1, j}}{\sum_{c=0}^{L_j-2} \beta^{c,j}} & \text{if } \mathrm{loc}\left(l, \mathcal{V}^{j}\right) > 0 \text{ and } \mathrm{loc}\left(i, \mathcal{V}^{j}\right) = 0\,, \\ \dfrac{W^{\mathcal{V}_0^j} \alpha^{\mathcal{V}_0^j, j}}{\left(W^{l}\right)^{2}} \dfrac{\beta^{\mathrm{loc}\left(l, \mathcal{V}^{j}\right)-1, j}}{\sum_{c=0}^{L_j-2} \beta^{c,j}} & \text{if } \mathrm{loc}\left(l, \mathcal{V}^{j}\right) > 0 \text{ and } \mathrm{loc}\left(i, \mathcal{V}^{j}\right) > 0 \text{ and } l = i\,, \text{ and} \\ 0 & \text{otherwise}\,, \end{cases} \tag{D.14}$$

for $l \in \mathcal{V}^{j}$ and $i \in \mathcal{V}^{j}$.

Finally, $\dfrac{\partial \mathbf{U}}{\partial \mathbf{W}} = \left[\dfrac{\partial \mathbf{U}}{\partial W^{i}} \forall i \in \mathcal{K}\right]$ in Eq. (D.13) needs to be estimated and is detailed in Appendix D.4.

Appendix D.2. Computation of $\dfrac{\partial \bar{\mathbf{v}}}{\partial \widetilde{\mathbf{A}}}$

For each element $\boldsymbol{\beta}^{j} \in \widetilde{\mathbf{A}}$, one has from Eq. (D.6) the following relation

$$\frac{\partial \bar{\mathbf{v}}}{\partial \boldsymbol{\beta}^{j}} = \frac{1}{\sum_{c \in \mathcal{C}} W^{c}} \sum_{i \in \mathcal{C}} W^{i} \mathbf{G}^{i} \frac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^{i}\right)}{\partial \boldsymbol{\beta}^{j}}\,. \tag{D.15}$$

Using Eq. (33, C.6), one has

$$\frac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^{i}\right)}{\partial \boldsymbol{\beta}^{j}} = \frac{\partial \alpha^{i,j}}{\partial \boldsymbol{\beta}^{j}} \mathrm{vec}^{\varepsilon}\left(\mathbf{a}^{j} \otimes^{s} \mathbf{n}^{j}\right) + \mathbf{D}^{i} \frac{\partial \mathbf{U}}{\partial \boldsymbol{\beta}^{j}}\,, \tag{D.16}$$

where $\dfrac{\partial \alpha^{i,j}}{\partial \boldsymbol{\beta}^{j}}$ is computed using Eq. (54) as

$$\frac{\partial \alpha^{l,j}}{\partial \beta^{r,j}} = \begin{cases} 0 & \text{if } l \notin \mathcal{V}^{j} \text{ or } \mathrm{loc}\left(l, \mathcal{V}^{j}\right) = 0 \\ -\dfrac{W^{\mathcal{V}_0^j} \alpha^{\mathcal{V}_0^j, j}}{W^{l}} \dfrac{1}{\sum_{c=0}^{L_j-2} \beta^{c,j}} \left(1 - \dfrac{\beta^{r,j}}{\sum_{c=0}^{L_j-2} \beta^{c,j}}\right) & \text{if } \mathrm{loc}\left(l, \mathcal{V}^{j}\right) = r+1 \\ \dfrac{W^{\mathcal{V}_0^j} \alpha^{\mathcal{V}_0^j, j}}{W^{l}} \dfrac{\beta^{\mathrm{loc}\left(l, \mathcal{V}^{j}\right)-1, j}}{\left(\sum_{c=0}^{L_j-2} \beta^{c,j}\right)^{2}} & \text{if } \mathrm{loc}\left(l, \mathcal{V}^{j}\right) \neq r+1 \end{cases} \tag{D.17}$$

for $l \in \mathcal{V}^{j}$ and $r = 0, \ldots, L_j - 2$,

Finally, $\dfrac{\partial \mathbf{U}}{\partial \widetilde{\mathbf{A}}} = \left[\dfrac{\partial \mathbf{U}}{\partial \boldsymbol{\beta}^{j}} \text{ for } j = 0, \ldots, M-1\right]$ needs to be estimated and is detailed in Appendix D.4.

Appendix D.3. Computation of $\dfrac{\partial \bar{\mathbf{v}}}{\partial \widetilde{\mathbf{N}}}$

For each element $\mathbf{e}^j \in \widetilde{\mathbf{N}}$, one has from Eq. (D.6)

$$\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{e}^j} = \frac{1}{\sum_{c\in\mathcal{C}} W^c} \sum_{i\in\mathcal{C}} W^i \mathbf{G}^i \frac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^i\right)}{\partial \mathbf{e}^j} \,. \tag{D.18}$$

Using Eq. (33, C.6), one has

$$\frac{\partial \mathrm{vec}^{\varepsilon}\left(\boldsymbol{\varepsilon}^i\right)}{\partial \mathbf{e}^j} = \alpha^{i,j} \mathbf{H}\left(\mathbf{a}^j\right) \frac{\partial \mathbf{n}^j}{\partial \mathbf{e}^j} + \mathbf{D}^i \frac{\partial \mathbf{U}}{\partial \mathbf{e}^j} \,, \tag{D.19}$$

where the operator $\mathbf{H}(\bullet)$ is given in Eq. (C.5), and $\dfrac{\partial \mathbf{n}^j}{\partial \mathbf{e}^j}$ is computed using Eq. (57) as

$$\frac{\partial \mathbf{n}^j}{\partial \mathbf{e}^j} = \begin{cases} \begin{bmatrix} -2\pi \sin\left(2\pi\varphi^j\right) & 2\pi \cos\left(\varphi^j\right) & 0 \end{bmatrix}^T & \text{in plane strain} \,, \\ \begin{bmatrix} -2\pi \sin\left(2\pi\varphi^j\right) \sin\left(\pi\theta^j\right) & 2\pi \cos\left(2\pi\varphi^j\right) \sin\left(\pi\theta^j\right) & 0 \\ \pi \cos\left(2\pi\varphi^j\right) \cos\left(\pi\theta^j\right) & \pi \sin\left(2\pi\varphi^j\right) \cos\left(\pi\theta^j\right) & -\pi \sin\left(\pi\theta^j\right) \end{bmatrix}^T & \text{otherwise} \,, \end{cases} \tag{D.20}$$

Finally, $\dfrac{\partial \mathbf{U}}{\partial \widetilde{\mathbf{N}}} = \left[\dfrac{\partial \mathbf{U}}{\partial \mathbf{e}^j} \text{ for } j = 0, \ldots, M-1 \right]$ needs to be estimated and is detailed in Appendix D.4.

Appendix D.4. Computation of $\dfrac{\partial \mathbf{U}}{\partial \mathbf{W}}$, $\dfrac{\partial \mathbf{U}}{\partial \widetilde{\mathbf{A}}}$, *and* $\dfrac{\partial \mathbf{U}}{\partial \widetilde{\mathbf{N}}}$

Finally, we must compute $\dfrac{\partial \mathbf{U}}{\partial W^i}$, $\dfrac{\partial \mathbf{U}}{\partial \boldsymbol{\beta}^j}$, $\dfrac{\partial \mathbf{U}}{\partial \mathbf{e}^j}$ for $i \in \mathcal{K}$ and $j = 0, \ldots, M-1$. For this purpose, Eq. (C.9) is first rewritten as

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}^0 \\ \mathbf{r}^1 \\ \vdots \\ \mathbf{r}^{M-1} \end{bmatrix} = \mathbf{0} \,, \tag{D.21}$$

where

$$\mathbf{r}^j = \left(\sum_{i\in\mathcal{V}^j} W^i \boldsymbol{\sigma}^i \alpha^{i,j} \right) \cdot \mathbf{n}^j = \mathbf{0} \text{ with } j = 0 \,, \ldots \,, M-1 \,. \tag{D.22}$$

The consistency of Eq. (D.21) for a given $\bar{\boldsymbol{\varepsilon}}$ reads

$$\delta \mathbf{r} = \mathbf{K} \delta \mathbf{U} + \sum_{i\in\mathcal{K}} \frac{\partial \mathbf{r}}{\partial W^i} \delta W^i + \sum_{j=0}^{M-1} \left(\frac{\partial \mathbf{r}}{\partial \boldsymbol{\beta}^j} \delta \boldsymbol{\beta}^j + \frac{\partial \mathbf{r}}{\partial \mathbf{e}^j} \delta \mathbf{e}^j \right) = \mathbf{0} \,, \tag{D.23}$$

leading to

$$\begin{bmatrix} \dfrac{\partial \mathbf{U}}{\partial \mathbf{W}} & \dfrac{\partial \mathbf{U}}{\partial \widetilde{\mathbf{A}}} & \dfrac{\partial \mathbf{U}}{\partial \widetilde{\mathbf{N}}} \end{bmatrix} = -\mathbf{K}^{-1} \begin{bmatrix} \dfrac{\partial \mathbf{r}}{\partial \mathbf{W}} & \dfrac{\partial \mathbf{r}}{\partial \widetilde{\mathbf{A}}} & \dfrac{\partial \mathbf{r}}{\partial \widetilde{\mathbf{N}}} \end{bmatrix} \,, \tag{D.24}$$

where

$$\frac{\partial \mathbf{r}}{\partial \mathbf{W}} = \left[\frac{\partial \mathbf{r}}{\partial W^l} \text{ for } l \in \mathcal{K} \right], \tag{D.25}$$

$$\frac{\partial \mathbf{r}}{\partial \widetilde{\mathbf{A}}} = \left[\frac{\partial \mathbf{r}}{\partial \tilde{\boldsymbol{\alpha}}^j} \text{ for } j = 0, \ldots, M-1 \right], \text{ and} \tag{D.26}$$

$$\frac{\partial \mathbf{r}}{\partial \widetilde{\mathbf{N}}} = \left[\frac{\partial \mathbf{r}}{\partial \mathbf{e}^j} \text{ for } j = 0, \ldots, M-1 \right]. \tag{D.27}$$

Equation (D.24) is a set of linear systems, one per column of the right hand side matrix, and the matrix $\mathbf{K}$ was previously factorized in order to solve the system of Eq. (C.12). The resolution of the set of linear systems (C.19) is then performed using this factorized matrix at a reduced computational time.

The following elements of the right hand side of Eq. (D.24) are computed as

$$\frac{\partial \mathbf{r}^j}{\partial W^i} = \frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} W^i} + \sum_{k \in \mathcal{T}^i} \frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} \tilde{\boldsymbol{\alpha}}^k} \frac{\partial \tilde{\boldsymbol{\alpha}}^k}{\partial W^i}, \tag{D.28}$$

$$\frac{\partial \mathbf{r}^j}{\partial \boldsymbol{\beta}^k} = \frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} \tilde{\boldsymbol{\alpha}}^k} \frac{\partial \tilde{\boldsymbol{\alpha}}^k}{\partial \boldsymbol{\beta}^k}, \text{ and} \tag{D.29}$$

$$\frac{\partial \mathbf{r}^j}{\partial \mathbf{e}^k} = \frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} \mathbf{n}^k} \frac{\partial \mathbf{n}^k}{\partial \mathbf{e}^k}, \tag{D.30}$$

where $\dfrac{\partial \tilde{\boldsymbol{\alpha}}^k}{\partial W^i}$, $\dfrac{\partial \tilde{\boldsymbol{\alpha}}^k}{\partial \boldsymbol{\beta}^k}$ and $\dfrac{\partial \mathbf{n}^k}{\partial \mathbf{e}^k}$ are computed by Eqs. (D.14, D.17, D.20), respectively, and the following terms

$$\frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} W^i} \quad \frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} \tilde{\boldsymbol{\alpha}}^k} \quad \text{and} \quad \frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} \mathbf{n}^k}, \text{ for } j = 0, \ldots, M-1, i \in \mathcal{K}, \text{ and } k = 0, \ldots, M-1, \tag{D.31}$$

are computed from Eq. (D.22) as follows:

- For each element $W^i \in \mathbf{W}$, one has (no sum intended)

$$\frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} W^i} = \begin{cases} \alpha^{i,j} \boldsymbol{\sigma}^i \cdot \mathbf{n}^j & \text{if } i \in \mathcal{V}^j, \text{ and} \\ 0 & \text{otherwise}. \end{cases} \tag{D.32}$$

- For each element $\alpha^{i,k} \in \tilde{\boldsymbol{\alpha}}^k$, one has (no sum intended)

$$\frac{\tilde{\partial} \mathbf{r}^j}{\tilde{\partial} \alpha^{i,k}} = \begin{cases} W^i \boldsymbol{\sigma}^i \cdot \mathbf{n}^j + W^i \alpha^{i,j} \left[\mathbf{C}^i \vdots \mathbf{n}^j \otimes \left(\mathbf{a}^k \otimes^s \mathbf{n}^k \right) \right] & \text{if } j = k \text{ and } i \in \mathcal{V}^j, \\ W^i \alpha^{i,j} \left[\mathbf{C}^i \vdots \mathbf{n}^j \otimes \left(\mathbf{a}^k \otimes^s \mathbf{n}^k \right) \right] & \text{if } j \neq k \text{ and } i \in \mathcal{V}^j, \text{ and} \\ 0 & \text{otherwise}. \end{cases} \tag{D.33}$$

- For each element $\mathbf{n}^k$, one has (no sum on j or k intended)

$$\frac{\tilde{\partial}\mathbf{r}^j}{\tilde{\partial}\mathbf{n}^k} = \begin{cases} \sum_{i\in\mathcal{V}^j} W^i\alpha^{i,j}\boldsymbol{\sigma}^i + \sum_{i\in\mathcal{V}^j} W^i\alpha^{i,j}\mathbf{C}^i \vdots \mathbf{n}^j \otimes \left(\alpha^{i,k}\mathbf{a}^k \otimes \mathbf{I}\right) & \text{if } j = k\,, \text{ and} \\ \sum_{i\in\mathcal{V}^j\cap\mathcal{V}^k} W^i\alpha^{i,j}\mathbf{C}^i \vdots \mathbf{n}^j \otimes \left(\alpha^{i,k}\mathbf{a}^k \otimes \mathbf{I}\right) & \text{if } j \neq k\,. \end{cases} \tag{D.34}$$

Appendix E. Optimization iterations

A stochastic gradient descent algorithm with adaptive moment estimation (Adam) [36] is adopted for the parameters update as follows:

- **initialize** the fitting parameters $\mathbf{X}_0$ following uniform distributions as follows:
  - generate $Z^l \in U\,(0.1, 1)\ \forall l \in \mathcal{K}$ and normalize these values as

$$W^i \leftarrow \omega^p \frac{Z^i}{\sum_{l\in\mathcal{K}^p} Z^l} \text{ for } i \in \mathcal{K}^p \text{ and } p = 0, \ldots, P-1\,, \tag{E.1}$$

    and then

$$Z^l = f^{-1}\left(\frac{N^p W^l}{\omega^p}\right) \text{ for } l \in \mathcal{K}^p \text{ and } p = 0, \ldots, P-1\,, \tag{E.2}$$

    where N^p is the number of material nodes governed by the constituent p and f^{-1} is the inverse of the activation function f;
  - generate $\boldsymbol{\beta}^j \in [U\,(-1, 1)]^{L_j - 1}$ for $j = 0, \ldots, M-1$; and
  - generate $\mathbf{e}^j \in [U\,(0, 1)]^d$ for $j = 0, \ldots, M-1$ and d being the number of components in $\mathbf{e}^j$;
- **set** the number of epochs N_{epochs} and initial learning rate l_r^0;
- **initialize** the first moment vector $\mathbf{m}_0 = \mathbf{0}$ and the second moment vector $\mathbf{s}_0 = \mathbf{0}$;
- **for** i **from** 0 **to** $N_{\text{epochs}} - 1$:
  - **update** learning rate l_r^{i+1} following a decay plan;
  - **shuffle** N_{train} training samples;
  - **for** l **from** 0 **to** $N_{train} - 1$:
    - **compute** $C_i^{[l]} = C^{[l]}\,(\mathbf{X}_i)$ for sample l following Eq. (68);
    - **compute** $\mathbf{g}_i^{[l]} = \mathbf{g}^{[l]}\,(\mathbf{X}_i)$ for sample l following Eq. (71) as detailed in Appendix D;

∗ **update** $\mathbf{X}_{i+1}$:

$$\mathbf{m}_{i+1} = \kappa_1 \mathbf{m}_i + (1-\kappa_1)\,\mathbf{g}_i^{[l]}\,, \tag{E.3}$$

$$\mathbf{s}_{i+1} = \kappa_2 \mathbf{s}_i + (1-\kappa_2)\left(\mathbf{g}_i^{[l]}\right)^2\,, \tag{E.4}$$

$$\hat{\mathbf{m}}_{i+1} = \frac{\mathbf{m}_{i+1}}{1-\kappa_1^{i+1}}\,, \tag{E.5}$$

$$\hat{\mathbf{s}}_{i+1} = \frac{\mathbf{s}_{i+1}}{1-\kappa_2^{i+1}}\,, \text{ and} \tag{E.6}$$

$$\mathbf{X}_{i+1} = \mathbf{X}_i - l_r^{i+1}\frac{\hat{\mathbf{m}}_{i+1}}{\sqrt{\hat{\mathbf{s}}_{i+1}}+\gamma}\,, \tag{E.7}$$

where the values $\kappa_1 = 0.9$, $\kappa_2 = 0.999$, and $\gamma = 10^{-8}$ proposed in [36] are used;

– **compute** the total loss C_{train} and the validation error C_{valid} following Eq. (72) respectively using the training dataset and the validation dataset for monitoring.

- exit.

The value N_{epochs} can be adapted on-the-fly using an early stopping criterion. $N_{\text{epochs}} = 500$ is set since the loss is almost unchanged after this value. It is often recommended to lower the learning rate by a decay plan during the training progress in order to obtain a better convergence near a minimum. The following decay plan is considered

$$l_r^{i+1} = l_r^0 0.5^{\frac{i}{100}}\,, \tag{E.8}$$

where $l_r^0 = 0.002$.

Appendix F. Homogenized elastic tensor and its gradient with respect to the fitting parameters

For the ease of implementation, the Voigt's notations of stress and strain tensors respectively corresponding to the bijective operators $\text{vec}^{\varepsilon}(\bullet)$ and $\text{vec}^{\sigma}(\bullet)$ are introduced in order to convert the strain and stress tensors into vectors as follows:

- For the strain tensor $\boldsymbol{\varepsilon}$:

$$\text{vec}^{\varepsilon}(\boldsymbol{\varepsilon}) = \begin{bmatrix}\varepsilon_{00} & \varepsilon_{11} & \varepsilon_{22} & 2\varepsilon_{01} & 2\varepsilon_{02} & 2\varepsilon_{12}\end{bmatrix}^T\,. \tag{F.1}$$

In particular, when a plane strain state is employed, one has its reduced form

$$\text{vec}^{\varepsilon,2D}(\boldsymbol{\varepsilon}) = \begin{bmatrix}\varepsilon_{00} & \varepsilon_{11} & 2\varepsilon_{01}\end{bmatrix}^T\,. \tag{F.2}$$

- For the stress tensor $\bar{\boldsymbol{\sigma}}$:

$$\text{vec}^{\sigma}(\bar{\boldsymbol{\sigma}}) = \begin{bmatrix}\sigma_{00} & \sigma_{11} & \sigma_{22} & \sigma_{01} & \sigma_{02} & \sigma_{12}\end{bmatrix}^T\,. \tag{F.3}$$

When a plane strain state is employed, one has its reduced form

$$\text{vec}^{\sigma,2D}(\bar{\boldsymbol{\sigma}}) = \begin{bmatrix}\sigma_{00} & \sigma_{11} & \sigma_{01}\end{bmatrix}^T\,. \tag{F.4}$$

As a result, the material tensor $\mathbf{C}$ can also be written in a matrix form of 36 components to preserve the tensor operations through matrix-vector multiplications as

$$\mathrm{mat}\left(\mathbf{C}\right) = \begin{bmatrix} C_{0000} & C_{0011} & C_{0022} & C_{0001} & C_{0002} & C_{0012} \\ C_{1100} & C_{1111} & C_{1122} & C_{1101} & C_{1102} & C_{1112} \\ C_{2200} & C_{2211} & C_{2222} & C_{2201} & C_{2202} & C_{2212} \\ C_{0100} & C_{0111} & C_{0122} & C_{0101} & C_{0102} & C_{0112} \\ C_{0200} & C_{0211} & C_{0222} & C_{0201} & C_{0202} & C_{0212} \\ C_{1200} & C_{1211} & C_{1222} & C_{1201} & C_{1202} & C_{1212} \end{bmatrix}, \tag{F.5}$$

where $\mathrm{mat}\left(\bullet\right)$ is the operator for this tensor-matrix conversion. Moreover, when a plane strain state is employed, one has its reduced form

$$\mathrm{mat}^{2D}\left(\mathbf{C}\right) = \begin{bmatrix} C_{0000} & C_{0011} & C_{0001} \\ C_{1100} & C_{1111} & C_{1101} \\ C_{0100} & C_{0111} & C_{0101} \end{bmatrix}. \tag{F.6}$$

The computation of the homogenized elastic tensor $\bar{\mathbf{C}}$ for given P elastic tensors $\mathbf{C}^{(0)}, \ldots, \mathbf{C}^{(P-1)}$ and given fitting parameters $\mathbf{X}$ is obtained as follows:

(i) The P elastic tensors are assigned to the material nodes;

(ii) A mode-wise strain sequence is defined as

$$\bar{\boldsymbol{\mathcal{E}}} = \begin{cases} \begin{bmatrix} \bar{\boldsymbol{\varepsilon}}_0 & \bar{\boldsymbol{\varepsilon}}_1 & \bar{\boldsymbol{\varepsilon}}_2 & \bar{\boldsymbol{\varepsilon}}_3 & \bar{\boldsymbol{\varepsilon}}_4 & \bar{\boldsymbol{\varepsilon}}_5 \end{bmatrix} & \text{for a general 3-dimensional strain state}, \\ \begin{bmatrix} \bar{\boldsymbol{\varepsilon}}_0 & \bar{\boldsymbol{\varepsilon}}_1 & \bar{\boldsymbol{\varepsilon}}_3 \end{bmatrix} & \text{for a plane strain state}, \end{cases} \tag{F.7}$$

where

$$\bar{\boldsymbol{\varepsilon}}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \bar{\boldsymbol{\varepsilon}}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \bar{\boldsymbol{\varepsilon}}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{F.8}$$

$$\bar{\boldsymbol{\varepsilon}}_3 = \begin{bmatrix} 0 & 0.5 & 0 \\ 0.5 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \bar{\boldsymbol{\varepsilon}}_4 = \begin{bmatrix} 0 & 0 & 0.5 \\ 0 & 0 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}, \text{ and } \bar{\boldsymbol{\varepsilon}}_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0.5 & 0 \end{bmatrix};$$

(iii) The corresponding response of the material network in terms of the homogenized stress sequence is estimated as

$$\bar{\boldsymbol{\mathcal{S}}} = \begin{cases} \begin{bmatrix} \bar{\boldsymbol{\sigma}}_0 & \bar{\boldsymbol{\sigma}}_1 & \bar{\boldsymbol{\sigma}}_2 & \bar{\boldsymbol{\sigma}}_3 & \bar{\boldsymbol{\sigma}}_4 & \bar{\boldsymbol{\sigma}}_5 \end{bmatrix} & \text{for a general 3-dimensional strain state}, \\ \begin{bmatrix} \bar{\boldsymbol{\sigma}}_0 & \bar{\boldsymbol{\sigma}}_1 & \bar{\boldsymbol{\sigma}}_3 \end{bmatrix} & \text{for a plane strain state}. \end{cases} \tag{F.9}$$

(iv) The homogenized elastic tensor $\bar{\mathbf{C}}$ is given by

– For a general 3-dimensional strain state:

$$\mathrm{mat}\left(\bar{\mathbf{C}}\right) = \begin{bmatrix} \mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}_0\right) & \mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}_1\right) & \mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}_2\right) & \mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}_3\right) & \mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}_4\right) & \mathrm{vec}^{\sigma}\left(\bar{\boldsymbol{\sigma}}_5\right) \end{bmatrix}. \tag{F.10}$$

– For a plane strain state:

$$\mathrm{mat}^{2D}\left(\bar{\mathbf{C}}\right) = \left[\mathrm{vec}^{\sigma,2D}\left(\bar{\boldsymbol{\sigma}}_0\right) \quad \mathrm{vec}^{\sigma,2D}\left(\bar{\boldsymbol{\sigma}}_1\right) \quad \mathrm{vec}^{\sigma,2D}\left(\bar{\boldsymbol{\sigma}}_3\right)\right] , \tag{F.11}$$

in which the components corresponding to the out-of-plane deformation mode are ignored.

(v) The derivatives of $\mathrm{mat}\left(\bar{\mathbf{C}}\right)$ with respect to the fitting parameters $\mathbf{X}$ can also be estimated. For each component X_j of $\mathbf{X}$, since the terms $\dfrac{\partial \bar{\boldsymbol{\sigma}}_i}{\partial X_j}$ with $i = 0, 1, 3$ for a plane strain state and $i = 0, \ldots, 5$ for a general 3-dimensional strain state, are computed by the procedure detailed in Appendix D, consequently, $\dfrac{\partial \bar{\mathbf{C}}}{\partial X_j}$ can be easily obtained as follows:

– For a general 3-dimensional strain state:

$$\mathrm{mat}\left(\frac{\partial \bar{\mathbf{C}}}{\partial X_j}\right) = \left[\mathrm{vec}^{\sigma}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_0}{\partial X_j}\right) \ \mathrm{vec}^{\sigma}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_1}{\partial X_j}\right) \ \mathrm{vec}^{\sigma}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_2}{\partial X_j}\right) \ \mathrm{vec}^{\sigma}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_3}{\partial X_j}\right) \ \mathrm{vec}^{\sigma}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_4}{\partial X_j}\right) \ \mathrm{vec}^{\sigma}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_5}{\partial X_j}\right)\right] . \tag{F.12}$$

– For a plane strain state:

$$\mathrm{mat}^{2D}\left(\frac{\partial \bar{\mathbf{C}}}{\partial X_j}\right) = \left[\mathrm{vec}^{\sigma,2D}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_0}{\partial X_j}\right) \quad \mathrm{vec}^{\sigma,2D}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_1}{\partial X_j}\right) \quad \mathrm{vec}^{\sigma,2D}\left(\frac{\partial \bar{\boldsymbol{\sigma}}_3}{\partial X_j}\right)\right] . \tag{F.13}$$

Appendix G. J_2 plasticity model and implicit time integration

For small strain plasticity, the total strain $\boldsymbol{\varepsilon}$ is decomposed following the rate form as

$$\dot{\boldsymbol{\varepsilon}} = \dot{\boldsymbol{\varepsilon}}^e + \dot{\boldsymbol{\varepsilon}}^p , \tag{G.1}$$

where $\boldsymbol{\varepsilon}^e$ and $\boldsymbol{\varepsilon}^p$ are the elastic and plastic parts respectively. The elastic constitutive equation reads

$$\boldsymbol{\sigma} = \mathbb{H} : \left(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p\right) \ \text{ with } \mathbb{H}_{ijkl} = \frac{E\nu}{(1+\nu)(1-2\nu)}\delta_{ij}\delta_{kl} + \frac{E}{2(1+\nu)}\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) , \tag{G.2}$$

where $\mathbb{H}$ is the fourth order Hooke tensor, E and ν are the Young's modulus and Poisson's ratio, respectively. The boundary of the elastic domain is described by the von Mises yield surface

$$f_y = \sigma^{\mathrm{VM}} - \sigma_y\left(\gamma\right) \leq 0 , \tag{G.3}$$

where $\sigma^{\mathrm{VM}} = \sqrt{\frac{3}{2}\mathrm{dev}\left(\boldsymbol{\sigma}\right) : \mathrm{dev}\left(\boldsymbol{\sigma}\right)}$ is the von Mises equivalent stress with $\mathrm{dev}\left(\bullet\right)$ being the deviatoric operator, *i.e* $\mathrm{dev}\left(\mathbf{A}\right) = \mathbf{A} - \dfrac{A_{ii}}{3}\mathbf{I}$ for an arbitrary symmetric second order tensor $\mathbf{A}$, σ_y is the isotropic yield stress, which is a function of γ being the equivalent plastic strain. The evolution of

γ is estimated from the plastic deformation tensor $\boldsymbol{\varepsilon}^p$ as

$$\dot{\gamma} = \sqrt{\frac{2}{3}\dot{\boldsymbol{\varepsilon}}^p : \dot{\boldsymbol{\varepsilon}}^p}\,. \tag{G.4}$$

The evolution of the plastic strain is governed by an associated flow rule

$$\dot{\boldsymbol{\varepsilon}}^p = \dot{\gamma}\frac{\partial f_y}{\partial \boldsymbol{\sigma}}\,. \tag{G.5}$$

The plastic evolution has to fulfill the Karush-Kuhn-Tucker condition

$$f_y \leq 0\,, \dot{\gamma} \geq, \text{ and } f_y\dot{\gamma} = 0\,. \tag{G.6}$$

The internal state is defined by $\mathbf{q} = [\boldsymbol{\varepsilon}^p\,, \gamma]$.

From the previous state $(\boldsymbol{\varepsilon}_n,\ \mathbf{q}_n)$ at time t_n, for given $\boldsymbol{\varepsilon}_{n+1}$, the current stress tensor $\boldsymbol{\sigma}_{n+1}$, the current tangent operator $\mathbf{C}_{n+1}$, and the current internal state $\mathbf{q}_{n+1}$ at time t_{n+1} are computed using a predictor-corrector scheme as follows:

(i) *Elastic predictor*:

- Assume: $\boldsymbol{\varepsilon}^{ppr} = \boldsymbol{\varepsilon}^p_n$ and $\gamma^{pr} = \gamma_n$.
- Compute predictor stress following Eq. (G.2): $\boldsymbol{\sigma}^{pr} = \mathbb{H} : (\boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}^{ppr})$.
- Estimate the yield condition (G.3) as

$$f_y^{pr} = \sigma^{\mathrm{VM}pr} - \sigma_y\left(\gamma_n\right) \text{ with } \sigma^{\mathrm{VM}pr} = \sqrt{\frac{3}{2}\mathrm{dev}\left(\boldsymbol{\sigma}^{pr}\right) : \mathrm{dev}\left(\boldsymbol{\sigma}^{pr}\right)}\,. \tag{G.7}$$

  * If $f_y^{pr} \leq 0$: $\boldsymbol{\sigma}_{n+1} = \boldsymbol{\sigma}^{pr}$, $\mathbf{C}_{n+1} = \mathbb{H}$, $\boldsymbol{\varepsilon}^p_{n+1} = \boldsymbol{\varepsilon}^{ppr}$, $\gamma_{n+1} = \gamma^{pr}$, and go to (iii).
  * If $f_y^{pr} > 0$: go to (ii) for plastic corrector.

(ii) *Plastic corrector*: The plastic flow rule (G.5) can be integrated using the implicit radial return mapping, leading to

$$\Delta\boldsymbol{\varepsilon}^p_{n+1} = \int_{t_n}^{t_{n+1}} \dot{\gamma}\frac{3\mathrm{dev}\left(\boldsymbol{\sigma}\right)}{2\sigma^{\mathrm{VM}}}\,dt \approx \Delta\gamma_{n+1}\frac{3\mathrm{dev}\left(\boldsymbol{\sigma}_{n+1}\right)}{2\sigma^{\mathrm{VM}}_{n+1}}\,, \tag{G.8}$$

where

$$\Delta\boldsymbol{\varepsilon}^p{}_{n+1} = \boldsymbol{\varepsilon}^p_{n+1} - \boldsymbol{\varepsilon}^p_n \text{ and } \Delta\gamma_{n+1} = \gamma_{n+1} - \gamma_n\,. \tag{G.9}$$

As a result, Eq. (G.2) becomes

$$\boldsymbol{\sigma}_{n+1} = \boldsymbol{\sigma}^{pr} - \mathbb{H} : \Delta\boldsymbol{\varepsilon}^p_{n+1} \text{ and } \mathrm{dev}\left(\boldsymbol{\sigma}_{n+1}\right) = \mathrm{dev}\left(\boldsymbol{\sigma}^{pr}\right) - 2G\Delta\gamma_{n+1}\frac{3\mathrm{dev}\left(\boldsymbol{\sigma}_{n+1}\right)}{2\sigma^{\mathrm{VM}}_{n+1}} \tag{G.10}$$

where $G = \frac{E}{2(1+\nu)}$. The last equation leads to

$$\mathrm{dev}\left(\boldsymbol{\sigma}_{n+1}\right) = \frac{\mathrm{dev}\left(\boldsymbol{\sigma}^{pr}\right)}{1 + \dfrac{3G\Delta\gamma_{n+1}}{\sigma^{\mathrm{VM}}_{n+1}}}\,, \tag{G.11}$$

and consequently

$$\sigma_{n+1}^{\mathrm{VM}} = \sigma^{\mathrm{VM}pr} - 3G\Delta\gamma_{n+1}\,. \tag{G.12}$$

As a result, the yield condition (G.3) becomes

$$\sigma^{\mathrm{VM}pr} - 3G\Delta\gamma_{n+1} - \sigma_y\left(\gamma_n + \Delta\gamma_{n+1}\right) = 0\,, \tag{G.13}$$

which is an equation in $\Delta\gamma_{n+1}$. Once $\Delta\gamma_{n+1}$ is known from the solution of Eq. (G.13) and using Eqs. (G.8, G.11), one has

$$\Delta\boldsymbol{\varepsilon}_{n+1}^{p} = \Delta\gamma_{n+1}\mathbb{N}^{\mathrm{pr}}\,, \tag{G.14}$$

where

$$\mathbb{N}^{\mathrm{pr}} = \frac{3\mathrm{dev}\left(\boldsymbol{\sigma}^{pr}\right)}{2\sigma^{\mathrm{VM}pr}}\,. \tag{G.15}$$

Finally, Eqs. (G.9, G.10) lead to

$$\begin{aligned}
\boldsymbol{\varepsilon}_{n+1}^{p} &= \boldsymbol{\varepsilon}_{n}^{p} + \Delta\boldsymbol{\varepsilon}_{n+1}^{p}\,, && \text{(G.16)}\\
\gamma_{n+1} &= \gamma_n + \Delta\gamma_{n+1}\,, \text{ and} && \text{(G.17)}\\
\boldsymbol{\sigma}_{n+1} &= \boldsymbol{\sigma}^{pr} - \mathbb{H} : \Delta\boldsymbol{\varepsilon}_{n+1}^{p}\,. && \text{(G.18)}
\end{aligned}$$

The tangent operator $\mathbf{C}_{n+1} = \dfrac{\partial\boldsymbol{\sigma}_{n+1}}{\partial\boldsymbol{\varepsilon}_{n+1}}$ is estimated using Eq. (G.18) as

$$\mathbf{C}_{n+1} = \frac{\partial\boldsymbol{\sigma}_{n+1}}{\partial\boldsymbol{\varepsilon}_{n+1}} = \mathbb{H} - \mathbb{H} : \frac{\partial\Delta\boldsymbol{\varepsilon}_{n+1}^{p}}{\partial\boldsymbol{\varepsilon}_{n+1}}\,. \tag{G.19}$$

The term $\dfrac{\partial\Delta\boldsymbol{\varepsilon}_{n+1}^{p}}{\partial\boldsymbol{\varepsilon}_{n+1}}$ is estimated using Eq. (G.14) by

$$\frac{\partial\Delta\boldsymbol{\varepsilon}_{n+1}^{p}}{\partial\boldsymbol{\varepsilon}_{n+1}} = \frac{\partial\Delta\gamma_{n+1}}{\partial\boldsymbol{\varepsilon}_{n+1}}\mathbb{N}^{\mathrm{pr}} + \Delta\gamma_{n+1}\frac{\partial\mathbb{N}^{\mathrm{pr}}}{\partial\boldsymbol{\varepsilon}_{n+1}}\,. \tag{G.20}$$

As a result, the computation of $\mathbf{C}_{n+1}$ following Eq. (G.19) requires the terms $\dfrac{\partial\Delta\gamma_{n+1}}{\partial\boldsymbol{\varepsilon}_{n+1}}$ and $\dfrac{\partial\mathbb{N}^{\mathrm{pr}}}{\partial\boldsymbol{\varepsilon}_{n+1}}$ to be estimated. The two terms are computed as follows:

– The consistency of Eq. (G.13) reads

$$\mathbb{N}^{\mathrm{pr}} : \mathbb{H} : \delta\boldsymbol{\varepsilon}_{n+1} - \left(3G + \frac{\sigma_y\left(\gamma_{n+1}\right)}{\partial\gamma_{n+1}}\right)\delta\Delta\gamma_{n+1} = 0\,. \tag{G.21}$$

As a result, one has

$$\frac{\partial\Delta\gamma_{n+1}}{\partial\boldsymbol{\varepsilon}_{n+1}} = \frac{\mathbb{N}^{\mathrm{pr}} : \mathbb{H}}{3G + H_{n+1}}\,, \tag{G.22}$$

where $H_{n+1} = \dfrac{\sigma_y(\gamma_{n+1})}{\partial\gamma_{n+1}}$ is the isotropic hardening modulus, which is known from material setting.

– The term $\dfrac{\partial\mathbb{N}^{\mathrm{pr}}}{\partial\boldsymbol{\varepsilon}_{n+1}}$ is estimated using Eq. (G.15) as

$$\frac{\partial\mathbb{N}^{\mathrm{pr}}}{\partial\boldsymbol{\varepsilon}_{n+1}} = \frac{1}{\sigma^{\mathrm{VM}pr}}\left(\frac{3\mathrm{dev}(\boldsymbol{\mathcal{I}})}{2} - \mathbb{N}^{\mathrm{pr}}\otimes\mathbb{N}^{\mathrm{pr}}\right) : \mathbb{H}\,, \tag{G.23}$$

where $\mathrm{dev}(\boldsymbol{\mathcal{I}})$ denotes the deviatoric part of the fourth order unit tensor $\boldsymbol{\mathcal{I}}$.

(iii) *End.*

Appendix H. Material sampling in the plane strain state

The generation procedure in [35] is summarized. For plane strain problems, P elastic matrices $\mathbf{L}^{(i)}$ with $i = 0, \ldots, P-1$ in each sample $\mathbb{X}^{[l]}$ are directly generated. The orthotropic elastic material of each phase is written as

$$\mathbf{L}^{(i)} = \begin{bmatrix} C_{00}^{(i)} & C_{01}^{(i)} & 0 \\ C_{01}^{(i)} & C_{11}^{(i)} & 0 \\ 0 & 0 & C_{22}^{(i)} \end{bmatrix} \text{ with } i = 0, \ldots, P-1\,. \tag{H.1}$$

It is noted that the conditions $C_{00}^{(i)}C_{11}^{(i)} - \left(C_{01}^{(i)}\right)^2 > 0$ and $C_{22}^{(i)} > 0$, with $i = 0, \ldots, P-1$, are enforced to obtain a positive definite matrix.

Since the set $[\mathbf{L}^{(0)}, \ldots, \mathbf{L}^{(P-1)}]$ and the set $[\gamma\mathbf{L}^{(0)}, \ldots, \gamma\mathbf{L}^{(P-1)}]$ for $\gamma \neq 0$ yield the same results, in order to avoid this scaling issue, we first generate

$$C_{00}^{(0)} = 1 \text{ and } \ln\left(\frac{C_{00}^{(i)}}{C_{00}^{(i-1)}}\right) \in U(-1,1) \text{ with } i = 1, \ldots, P-1\,, \tag{H.2}$$

while all others components are then generated as

$$\ln\left(\frac{C_{11}^{(i)}}{C_{00}^{(i)}}\right) \in U(-1,1)\,,\ \frac{C_{01}^{(i)}}{\sqrt{C_{00}^{(i)}C_{11}^{(i)}}} \in U(0,0.9)\,,\ \text{and } \ln\left(\frac{C_{22}^{(i)}}{\sqrt{C_{00}^{(i)}C_{11}^{(i)}}}\right) \in U(-1,1)\,, \tag{H.3}$$
$$\text{for } i = 0, \ldots, P-1,$$

where $U(a,b)$ denotes a uniform distribution in the range $[a,b]$.

References

[1] J. Michel, H. Moulinec, P. Suquet, Effective properties of composite materials with periodic microstructure: a computational approach, Computer Methods in Applied Mechanics and Engineering 172 (1) (1999) 109–143 (1999). doi:https://doi.org/10.1016/S0045-7825(98)00227-8.
URL https://www.sciencedirect.com/science/article/pii/S0045782598002278

[2] F. Feyel, J.-L. Chaboche, Fe2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre sic/ti composite materials, Computer Methods in Applied Mechanics and Engineering 183 (3) (2000) 309–330 (2000). doi:https://doi.org/10.1016/S0045-7825(99)00224-8.
URL https://www.sciencedirect.com/science/article/pii/S0045782599002248

[3] C. Miehe, A. Koch, Computational micro-to-macro transitions of discretized microstructures undergoing small strains, Archive of Applied Mechanics 72 (4) (2002) 300–317 (2002).

[4] V. Kouznetsova, W. Brekelmans, F. Baaijens, An approach to micro-macro modeling of heterogeneous materials, Computational mechanics 27 (1) (2001) 37–48 (2001).

[5] M. G. Geers, V. G. Kouznetsova, W. Brekelmans, Multi-scale computational homogenization: Trends and challenges, Journal of computational and applied mathematics 234 (7) (2010) 2175–2182 (2010).

[6] M. G. Geers, V. G. Kouznetsova, K. Matouš, J. Yvonnet, Homogenization methods and multiscale modeling: nonlinear problems, Encyclopedia of Computational Mechanics Second Edition (2017) 1–34 (2017).

[7] K. Matouš, M. G. Geers, V. G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, Journal of Computational Physics 330 (2017) 192 – 220 (2017). doi:https://doi.org/10.1016/j.jcp.2016.10.070.
URL http://www.sciencedirect.com/science/article/pii/S0021999116305782

[8] J. Yvonnet, E. Monteiro, Q.-C. He, Computational homogenization method and reduced database model for hyperelastic heterogeneous structures, International Journal for Multiscale Computational Engineering 11 (3) (2013).

[9] J. Hernández, J. Oliver, A. Huespe, M. Caicedo, J. Cante, High-performance model reduction techniques in computational multiscale homogenization, Computer Methods in Applied Mechanics and Engineering 276 (2014) 149 – 189 (2014). doi:https://doi.org/10.1016/j.cma.2014.03.011.
URL http://www.sciencedirect.com/science/article/pii/S0045782514000978

[10] Z. Liu, M. Bessa, W. K. Liu, Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials, Computer Methods in Applied Mechanics and Engineering 306 (2016) 319 – 341 (2016). doi:https://doi.org/10.1016/j.cma.2016.04.004.
URL http://www.sciencedirect.com/science/article/pii/S0045782516301499

[11] J.-C. Michel, P. Suquet, A model-reduction approach in micromechanics of materials preserving the variational structure of constitutive relations, Journal of the Mechanics and Physics of Solids 90 (2016) 254 – 285 (2016). doi:https://doi.org/10.1016/j.jmps.2016.02.005.
URL http://www.sciencedirect.com/science/article/pii/S0022509616300928

[12] D. Soldner, B. Brands, R. Zabihyan, P. Steinmann, J. Mergheim, A numerical study of different projection-based model reduction techniques applied to computational homogenisation, Computational mechanics 60 (4) (2017) 613–625 (2017).

[13] M. J. Zahr, P. Avery, C. Farhat, A multilevel projection-based model order reduction framework for nonlinear dynamic multiscale problems in structural and solid mechanics, International Journal for Numerical Methods in Engineering 112 (8) (2017) 855–881 (2017). arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.5535, doi:10.1002/nme.5535.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.5535

[14] J. Oliver, M. Caicedo, A. Huespe, J. Hernández, E. Roubin, Reduced order modeling strategies for computational multiscale fracture, Computer Methods in Applied Mechanics and Engineering 313 (2017) 560 – 595 (2017). doi:https://doi.org/10.1016/j.cma.2016.09.039.
URL http://www.sciencedirect.com/science/article/pii/S0045782516303322

[15] S. Wulfinghoff, F. Cavaliere, S. Reese, Model order reduction of nonlinear homogenization problems using a hashinshtrikman type finite element method, Computer Methods in Applied Mechanics and Engineering 330 (2018) 149 – 179 (2018). doi:https://doi.org/10.1016/j.cma.2017.10.019.
URL http://www.sciencedirect.com/science/article/pii/S0045782517306904

[16] B. A. Le, J. Yvonnet, Q. C. He, Computational homogenization of nonlinear elastic materials using neural networks, International Journal for Numerical Methods in Engineering 104 (12) (2015) 1061–1084 (Dec. 2015). doi:10.1002/nme.4953.

[17] M. Bessa, R. Bostanabad, Z. Liu, A. Hu, D. W. Apley, C. Brinson, W. Chen, W. Liu, A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality, Computer Methods in Applied Mechanics and Engineering 320 (2017) 633 – 667 (2017). doi:https://doi.org/10.1016/j.cma.2017.03.037.
URL http://www.sciencedirect.com/science/article/pii/S0045782516314803

[18] F. Fritzen, M. Fernndez, F. Larsson, On-the-fly adaptivity for nonlinear twoscale simulations using artificial neural networks and reduced order modeling, Frontiers in Materials 6 (2019) 75 (2019). doi:10.3389/fmats.2019.00075.
URL https://www.frontiersin.org/article/10.3389/fmats.2019.00075

[19] H. Yang, X. Guo, S. Tang, W. K. Liu, Derivation of heterogeneous material laws via data-driven principal component expansions, Computational Mechanics 64 (2) (2019) 365–379 (2019).

[20] M. B. Gorji, M. Mozaffar, J. N. Heidenreich, J. Cao, D. Mohr, On the potential of recurrent neural networks for modeling path dependent plasticity, Journal of the Mechanics and Physics of Solids 143 (2020) 103972 (2020). doi:https://doi.org/10.1016/j.jmps.2020.103972.
URL http://www.sciencedirect.com/science/article/pii/S0022509620302076

[21] D. Huang, J. N. Fuhg, C. Weienfels, P. Wriggers, A machine learning based plasticity model using proper orthogonal decomposition, Computer Methods in Applied Mechanics and Engineering 365 (2020) 113008 (2020). doi:https://doi.org/10.1016/j.cma.2020.113008.
URL http://www.sciencedirect.com/science/article/pii/S0045782520301924

[22] M. Mozaffar, R. Bostanabad, W. Chen, K. Ehmann, J. Cao, M. A. Bessa, Deep learning predicts path-dependent plasticity, Proceedings of the National Academy of Sciences 116 (52) (2019) 26414–26420 (2019). arXiv:https://www.pnas.org/content/116/52/26414.full.pdf, doi:10.1073/pnas.1911815116.
URL https://www.pnas.org/content/116/52/26414

[23] F. Ghavamian, A. Simone, Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network, Computer Methods in Applied Mechanics and Engineering 357 (2019) 112594 (2019). doi:https://doi.org/10.1016/j.cma.2019.112594.
URL http://www.sciencedirect.com/science/article/pii/S0045782519304700

[24] L. Wu, V. D. Nguyen, N. G. Kilingar, L. Noels, A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths, Computer Methods in Applied Mechanics and Engineering 369 (2020) 113234 (2020). doi:https://doi.org/10.1016/j.cma.2020.113234.
URL http://www.sciencedirect.com/science/article/pii/S0045782520304199

[25] H. J. Logarzo, G. Capuano, J. J. Rimoli, Smart constitutive laws: Inelastic homogenization through machine learning, Computer Methods in Applied Mechanics and Engineering 373 (2021) 113482 (2021). doi:https://doi.org/10.1016/j.cma.2020.113482.
URL http://www.sciencedirect.com/science/article/pii/S0045782520306678

[26] F. Masi, I. Stefanou, P. Vannucci, V. Maffi-Berthier, Thermodynamics-based artificial neural networks for constitutive modeling, Journal of the Mechanics and Physics of Solids 147 (2021) 104277 (2021). doi:https://doi.org/10.1016/j.jmps.2020.104277.
URL https://www.sciencedirect.com/science/article/pii/S0022509620304841

[27] J. N. Fuhg, C. Bhm, N. Bouklas, A. Fau, P. Wriggers, M. Marino, Model-data-driven constitutive responses: Application to a multiscale computational framework, International Journal of Engineering Science 167 (2021) 103522 (2021). doi:https://doi.org/10.1016/j.ijengsci.2021.103522.
URL https://www.sciencedirect.com/science/article/pii/S0020722521000690

[28] Z. Liu, C. Wu, M. Koishi, A deep material network for multiscale topology learning and accelerated nonlinear modeling of heterogeneous materials, Computer Methods in Applied Mechanics and Engineering 345 (2019) 1138 – 1168 (2019). doi:https://doi.org/10.1016/j.cma.2018.09.020.
URL http://www.sciencedirect.com/science/article/pii/S0045782518304729

[29] Z. Liu, C. Wu, Exploring the 3d architectures of deep material network in data-driven multiscale mechanics, Journal of the Mechanics and Physics of Solids 127 (2019) 20 – 46 (2019). doi:https://doi.org/10.1016/j.jmps.2019.03.004.
URL http://www.sciencedirect.com/science/article/pii/S0022509618310688

[30] Z. Liu, Deep material network with cohesive layers: Multi-stage training and interfacial failure analysis, Computer Methods in Applied Mechanics and Engineering 363 (2020) 112913 (2020). doi:https://doi.org/10.1016/j.cma.2020.112913.
URL http://www.sciencedirect.com/science/article/pii/S0045782520300967

[31] Z. Liu, Cell division in deep material networks applied to multiscale strain localization modeling, Computer Methods in Applied Mechanics and Engineering 384 (2021) 113914 (2021). doi:https://doi.org/10.1016/j.cma.2021.113914.
URL https://www.sciencedirect.com/science/article/pii/S0045782521002516

[32] S. Gajek, M. Schneider, T. Bhlke, On the micromechanics of deep material networks, Journal of the Mechanics and Physics of Solids 142 (2020) 103984 (2020). doi:https://doi.org/10.1016/j.jmps.2020.103984.
URL http://www.sciencedirect.com/science/article/pii/S0022509620302192

[33] S. Gajek, M. Schneider, T. Böhlke, An fedmn method for the multiscale analysis of short fiber reinforced plastic components, Computer Methods in Applied Mechanics and Engineering 384 (2021) 113952 (2021). doi:https://doi.org/10.1016/j.cma.2021.113952.
URL https://www.sciencedirect.com/science/article/pii/S0045782521002899

[34] L. Wu, L. Adam, L. Noels, Micro-mechanics and data-driven based reduced order models for multi-scale analyses of woven composites, Composite Structures 270 (2021) 114058 (2021). doi:https://doi.org/10.1016/j.compstruct.2021.114058.
URL https://www.sciencedirect.com/science/article/pii/S0263822321005183

[35] V. D. Nguyen, L. Noels, Micromechanics-based material networks revisited from the interaction viewpoint; robust and efficient implementation for multi-phase composites, European Journal of Mechanics - A/Solids (2021) 104384 (2021). doi:https://doi.org/10.1016/j.euromechsol.2021.104384.
URL https://www.sciencedirect.com/science/article/pii/S0997753821001431

[36] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[37] V.-D. Nguyen, L. Wu, L. Noels, Unified treatment of microscopic boundary conditions and efficient algorithms for estimating tangent operators of the homogenized behavior in the computational homogenization method, Comput. Mech. 59 (3) (2017) 483505 (Mar. 2017). doi:10.1007/s00466-016-1358-z.
URL https://doi.org/10.1007/s00466-016-1358-z

[38] V.-D. Nguyen, L. Noels, Data of Interaction-based material network: a general framework for (porous) microstructured materials (2021). doi:https://doi.org/10.5281/zenodo.5568833.

[39] C. Miehe, Computational micro-to-macro transitions for discretized micro-structures of heterogeneous materials at finite strains based on the minimization of averaged incremental energy, Computer Methods in Applied Mechanics and Engineering 192 (5) (2003) 559 – 591 (2003). doi:https://doi.org/10.1016/S0045-7825(02)00564-9.
URL http://www.sciencedirect.com/science/article/pii/S0045782502005649

[40] V.-D. Nguyen, E. Bchet, C. Geuzaine, L. Noels, Imposing periodic boundary condition on arbitrary meshes by polynomial interpolation, Computational Materials Science 55 (2012) 390 – 406 (2012). doi:https://doi.org/10.1016/j.commatsci.2011.10.017.
URL http://www.sciencedirect.com/science/article/pii/S0927025611005866