Learning constitutive models from microstructural simulations via a non-intrusive reduced basis method

Theron Guo[a],[*], Ondřej Rokoš[b], Karen Veroy[a]

[a] *Centre for Analysis, Scientific Computing and Applications, Department of Mathematics and Computer Science, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands*
[b] *Mechanics of Materials, Department of Mechanical Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands*

Abstract

In order to optimally design materials, it is crucial to understand the structure–property relations in the material by analyzing the effect of microstructure parameters on the macroscopic properties. In computational homogenization, the microstructure is thus explicitly modeled inside the macrostructure, leading to a coupled two-scale formulation. Unfortunately, the high computational costs of such multiscale simulations often render the solution of design, optimization, or inverse problems infeasible. To address this issue, we propose in this work a non-intrusive reduced basis method to construct inexpensive surrogates for parametrized microscale problems; the method is specifically well-suited for multiscale simulations since the coupled simulation is decoupled into two independent problems: (1) solving the microscopic problem for different (loading or material) parameters and learning a surrogate model from the data; and (2) solving the macroscopic problem with the learned material model. The proposed method has three key features. First, the microscopic stress field can be fully recovered, which is useful for instance for revealing local stress concentrations inside the microstructure. Second, the method is able to accurately predict the stress field for a wide range of material parameters; furthermore, the derivatives of the effective stress with respect to the material parameters are available and can be readily utilized in solving optimization problems. Finally, it is more data efficient, i.e. requiring less training data, as compared to directly performing a regression on the effective stress. To construct the surrogate model, first, a proper orthogonal decomposition is performed on precomputed microscopic stress field snapshots to find a reduced basis for the stress. Second, a regression is employed to infer the coefficients of the reduced basis approximation for any arbitrary parameter value, thus enabling a rapid online evaluation of the microscopic stress. Equipped with the stress field, the effective stress and its partial derivatives can then be derived analytically. For the microstructures in the two test problems considered, the mean approximation error of the effective stress is as low as 0.1% despite using a relatively small training dataset. Embedded into the macroscopic problem, the reduced order model leads to an online computational speed up of approximately three orders of magnitude while maintaining a high accuracy as compared to the FE^2 solver.

* Corresponding author.
E-mail addresses: t.guo@tue.nl (T. Guo), o.rokos@tue.nl (O. Rokoš), k.p.veroy@tue.nl (K. Veroy).

1. Introduction

Simulation methods are widely utilized to bridge the understanding between macroscopic behavior and the microstructure of the material. This can be achieved by using multiscale methods such as computational homogenization (CH). In CH, the microstructure and macrostructure are both separately modeled, with the microstructure essentially replacing the constitutive model of the macroscopic problem. The microstructure is fully defined on a representative volume element (RVE), which can be highly heterogeneous and exert highly nonlinear behavior, while the macrostructure is then assumed to be homogeneous.

A comprehensive overview on CH is given in Geers et al. [1,2] and Matouš et al. [3]. The two-scale problem can be solved with either a nested finite element (FE) scheme, first proposed in Feyel [4], or by a combination of the FE method and fast Fourier transform (FFT) [5]. More recently, other mesh-free methods based on deep neural networks such as physics-informed neural networks [6,7] and the deep collocation method [8,9] have also been used for the solution of partial differential equations (PDEs).

Unfortunately, solving the two-scale system in CH is computationally expensive and can quickly exceed available resources, limiting the usage to mostly two-dimensional applications [10–13]. Moreover, the difficulty in obtaining the consistent effective stiffness from the microscale simulation presents an additional hurdle, as it is typically required for the macroscopic Newton solver [14]. Popular approaches to find the appropriate stiffness include the Lagrange multiplier method [15,16], condensation method [10], or the perturbation method [17,18]. However, all these methods introduce additional problems such as increased storage requirements or higher computational costs.

To overcome this computational burden, several different reduced order models have been proposed to speed up or to replace the microscopic simulation. These can be largely classified according to how they treat parameters, and according to whether or not they are intrusive. The ability to handle parametrized problems makes a method suitable for inverse tasks such as parameter identification, material design or optimization. Intrusive methods require the modification of the underlying PDE solver, often making them impractical to use. In contrast, non-intrusive (data-driven) methods can be wrapped around an already existent PDE solver, allowing for a simpler adoption. In the context of CH, data-driven methods typically generate datasets from microscopic simulations and learn an effective constitutive model from the data, essentially decoupling the two-scale problem.

One popular data-driven framework was introduced in Kirchdoerfer et al. [19] and extended in [20–22]. These works propose a distance minimizing scheme, with which pairs of strain and stress data can be directly utilized inside a simulation without the need of deriving an empirical constitutive model. This dataset can come either from experiments or (microscopic) simulations. However, due to the nature of this method, it is only able to describe and reconstruct a given material dataset and cannot predict new materials.

Recognizing that stress–strain data is generally difficult to obtain from experiments, a slightly different approach was introduced in Huang et al. [23,24], where the constitutive model is approximated by a neural network, which is then inversely trained from load–displacement data. Unfortunately, the trained parameters inside the neural network (weights and biases) do not represent actual physical parameters, and therefore the method can only describe the given dataset.

Several other works attempted to learn a constitutive model directly from pairs of stress and strain data with different neural network architectures [25–32]. After the pioneering works in Ghaboussi et al. [25], several works have applied deep neural networks (DNN) to different areas of constitutive modeling. In [27–29] different recurrent neural network architectures were proposed to learn an inelastic material model from stress and strain loading data that were obtained from RVE simulations. However, the models were trained only with deformation data and did not consider parameters, inhibiting their use for finding new materials and designs. Recent works have also attempted to overcome this challenge. In Le et al. [33], it was proposed to use neural networks to learn the effective potential of a hyperelastic material from which the effective stress and stiffness can be derived, while also including multiple microstructural parameters. In Mozaffar et al. [30] a recurrent neural network that can treat both temporal (strain path) and non-temporal (volume fraction) parameters simultaneously was proposed, and the stress prediction for a class of composite materials was illustrated. More recently, works [24,31,32] have also attempted to construct a neural network that learns constitutive models that fulfill the thermodynamical laws as defined in [34].

Although these neural network-based methods are capable of finding accurate surrogate models, they typically require huge amounts of data, which are often generated from simulations. For complex microstructures, running these simulations repeatedly might be computationally too expensive. Furthermore, these surrogate models generally only consider the effective quantities, which are obtained by averaging the microscopic quantities, and neglect all the

field information of the microscopic simulation. This means that the connection between macro- and microstructure is completely lost and local microscopic quantities can no longer be recovered.

Intrusive methods, on the other hand, still consider the microscopic PDE and therefore often require far less data due to the known physics. Instead of replacing the microscopic problem, the solution is accelerated. One notable method is the Transformation Field Analysis (TFA) [35], which is specifically designed for simulations involving inelastic materials. By considering the field of internal variables to be piecewise uniform, the computational cost of evaluating the nonlinear terms is greatly decreased. Moreover, the number of internal variables that needs to be tracked is also reduced. Later, the method was generalized to non-uniform fields and referred to as Nonuniform Transformation Field Analysis (NTFA) [36,37]. Even though successful in accelerating two-scale simulations, the reduced model is designed for a specific microscopic material model and therefore cannot handle material parameters [38].

Another approach called the Self-consistent Clustering Analysis (SCA) was proposed in Liu et al. [39] and extended in [40]. The idea of this method is to find material points within a microstructure that exhibit similar deformation behavior and to group them into clusters. It is then assumed that each cluster exhibits the same deformation, therefore reducing the number of independent points to the number of clusters. The compressed problem is then solved with a FFT approach. A similar approach was proposed in [41], however, in that work the problem is solved with a three-field Hashin–Shtrikman type variational formulation. Similarly to (N)TFA, this method does not offer a direct way of treating material parameters.

In the Proper Generalized Decomposition (PGD) [42,43], the solution to a PDE is approximated by a linear combination of terms that each consist of a separated functional representation for each independent variable (i.e. coordinates, time, and parameters). The PGD approach has been applied to nonlinear solid mechanics problems, see, e.g., [44]. It has also been combined with the LATIN multiscale method for non-linear problems, see, e.g., [43]. However, the method becomes increasingly more difficult for severe nonlinearities, especially in combination with varying parameters [42].

A related approach is the reduced basis (RB) method [45–47] which attempts to reduce the parametric problem by searching for the PDE solution on a reduced space spanned by global parameter-independent basis functions. The basis functions can be constructed by a proper orthogonal decomposition (POD) of a collection of high-fidelity simulations, which are often referred to as snapshots. The solution can then be found by either solving the reduced problem with a (Petrov–)Galerkin projection [48–53] or as proposed in more recent works by using a regression-based approach to find the basis coefficients [54–59]. Material parameters can naturally be dealt with in this framework.

The projection-based approach works especially well when the problem has an affine dependence on the parameters, allowing the decoupling into an offline (construction of the RB) and online stage (solving the reduced set of equations). However, this is not the case for the general solution of a nonlinear problem since the nonlinear terms will still depend on the full scale problem. To work around this issue, methods such as the empirical interpolation method (EIM) [60] and its discrete variant [61], or hyperreduction [52] can be used to approximate the nonlinear term and recover the affine dependence. In the context of CH, POD was first applied on a hyperelastic material in Yvonnet et al. [62]. However, in this work the Galerkin approach was directly applied without any special treatment of the nonlinear term, thus resulting in a rather modest speedup. In Hernández et al. [53], the computation of the nonlinear term was accelerated by a hyperreduction scheme and an algorithm to find appropriate integration points was proposed. In Soldner et al. [50], POD was combined with three different hyperreduction schemes and it was found that some approximations led to convergence problems in some cases. Another problem is that there does not exist a practical way of calculating the consistent effective stiffness [62], as this would require the full assembly of the microscopic tangent matrix. If the tangent matrix was approximated with a hyperreduction scheme, its symmetry might not be preserved and therefore an inconsistent effective stiffness obtained with a perturbation method might be preferable [50].

Instead of solving the reduced system, the regression-based approach infers the coefficients of the corresponding RB functions from a regression. In doing so, the online phase only consists of evaluating the coefficients from which the solution field can be directly obtained. Furthermore, the method becomes non-intrusive, similar to the other data-driven approaches. However, as observed in Hesthaven et al. [55], considerably more snapshots might be needed to construct an accurate regression as compared to an accurate basis. For that reason, active learning algorithms [54,59] or a bifidelity reconstruction [56] were recently proposed in order to minimize the number of snapshots needed.

While the projection-based method efficiently approximates only the field variable (here, the displacements), the regression-based approach could, in principle, be applied to any derived variable obtained in the high-fidelity solution (here, e.g., the stress). In [57], Swischuk et al. utilized a regression-based POD to directly find the strain field. In the case of CH, the microscopic displacements are not needed, but only the effective stress is required. Hence, in this work, we propose a regression-based RB surrogate model for the microscopic stress field, which is specifically geared towards multiscale simulations, to combine the advantages of intrusive and non-intrusive methods: (i) Due to the non-intrusiveness, this method can be easily implemented and the two-scale problem is decoupled into two independent problems, (ii) the approximation of the stress field allows a direct way of computing the effective stress and its derivative with respect to any (loading or material) parameter and (iii) even though the method requires more data as compared to the projection-based POD, it is more data efficient as compared to the pure data-driven methods. Furthermore, there are two reasons that favor the stress-based regression over the usual displacement-based approach:

1. If the displacements were approximated, the stresses would still need to be computed from the displacements. This means that the material model must be implemented again, or the displacement data has to be fed into the microstructural solver to return the stresses, thus negating the advantages of the non-intrusive method. Furthermore this would lead to unnecessary additional computation, leading to decreased efficiency.
2. Inelastic material models are defined in stress rate rather than stresses, including internal variables in their formulation that need to be tracked. It would thus not be possible to determine the stresses using only the displacement field since the evolution of the internal variables would not be known. As a result, the procedure would become highly inefficient or even inapplicable to history-dependent material behavior. On the other hand, works in DNN have shown that history-dependent behavior can be learned from stress and strain data alone without the need for internal variables [25–30,63].

The remainder of this paper is organized as follows. In Section 2, the two-scale theory of first-order computational homogenization is reviewed. Following that, Section 3 presents how the non-intrusive reduced order model is constructed. Moreover, an error estimation and a comparison to neural networks is provided. Two microstructural simulations and a two-scale problem are addressed in Section 4, testing the capabilities of the proposed framework. A summary of the findings and concluding remarks are presented in Section 5.

Notation. In this work, italic bold symbols are employed for coordinates and functions, such as the coordinates $\boldsymbol{X}$, the displacement field $\boldsymbol{u}$ or the stress field $\boldsymbol{P}$. Upright bold symbols are used for vectors or tensors representing algebraic or discretized quantities, such as the identity matrix $\mathbf{I}$, discrete stress field snapshots $\mathbf{P}$ or the snapshot tensor $\mathbf{S}$. When $\mathbf{A}$ is a Nth-order tensor with indices $x_1, x_2, \ldots, x_N$, then $\mathbf{A}_{x_1,x_2,\ldots,x_n}$ with $n \leq N$ denotes the same tensor with the first n indices held fixed and is of $(N - n)$th-order. As an example, when $\mathbf{A}$ is a matrix, then $\mathbf{A}_i$ denotes the ith row vector and A_{ij} the entry at the ith row and jth column. Macroscopic variables are denoted by an overline to distinguish them from microscopic quantities; otherwise, the notation described above is used for both micro- and macroscopic quantities.

2. Homogenization theory

2.1. Macroscopic simulation

Consider a solid body $\bar{\Omega}$ that is deformed under prescribed boundary conditions. Under the deformation, each point $\bar{\boldsymbol{X}}$ of the undeformed body is continuously mapped onto the coordinates $\bar{\boldsymbol{x}}$ of the deformed body. The macroscopic displacement for each point is defined as $\bar{\boldsymbol{u}} := \bar{\boldsymbol{x}} - \bar{\boldsymbol{X}}$, and the macroscopic deformation gradient then follows as

$$\bar{\boldsymbol{F}} := \frac{\partial \bar{\boldsymbol{x}}}{\partial \bar{\boldsymbol{X}}} = \mathbf{I} + \frac{\partial \bar{\boldsymbol{u}}}{\partial \bar{\boldsymbol{X}}}. \tag{1}$$

The governing partial differential equation (PDE) describing the deformation is given by the quasi-static linear momentum balance,

$$\begin{aligned}
\operatorname{Div} \bar{\boldsymbol{P}} + \bar{\boldsymbol{B}} &= \boldsymbol{0} && \text{on } \bar{\Omega}, \\
\bar{\boldsymbol{P}} \bar{\boldsymbol{N}} &= \bar{\boldsymbol{t}}_0 && \text{on } \partial \bar{\Omega}^N, \text{ and} \\
\bar{\boldsymbol{u}} &= \bar{\boldsymbol{u}}_0 && \text{on } \partial \bar{\Omega}^D,
\end{aligned} \tag{2}$$

where $\bar{\boldsymbol{P}}$ is the macroscopic first Piola–Kirchhoff (PK1) stress tensor, $\bar{\boldsymbol{B}}$ are the body forces, $\bar{\boldsymbol{N}}$ is the outward normal on the surface of the body, $\bar{\boldsymbol{t}}_0$ and $\bar{\boldsymbol{u}}_0$ are the prescribed traction and displacement, and $\partial\bar{\Omega}^N$, $\partial\bar{\Omega}^D$ denote the Neumann and Dirichlet boundaries with $\partial\bar{\Omega} = \partial\bar{\Omega}^N \cup \partial\bar{\Omega}^D$ and $\partial\bar{\Omega}^N \cap \partial\bar{\Omega}^D = \emptyset$. The weak form of Eq. (2) is then

$$\bar{G} := \int_{\bar{\Omega}} \frac{\partial\delta\bar{\boldsymbol{u}}}{\partial\bar{\boldsymbol{X}}} : \bar{\boldsymbol{P}}dV - \int_{\bar{\Omega}} \bar{\boldsymbol{B}} \cdot \delta\bar{\boldsymbol{u}}dV - \int_{\partial\bar{\Omega}^N} \bar{\boldsymbol{t}}_0 \cdot \delta\bar{\boldsymbol{u}}dA = 0, \qquad \forall\delta\bar{\boldsymbol{u}} \in H_0^1(\bar{\Omega}), \tag{3}$$

where $\delta\bar{\boldsymbol{u}} \in H_0^1(\bar{\Omega}) = \{\boldsymbol{v} \in H^1(\bar{\Omega}) \mid \boldsymbol{v} = 0 \text{ on } \partial\bar{\Omega}^D\}$ is a test function, and a solution for the displacement, $\bar{\boldsymbol{u}} \in H^1(\bar{\Omega})$, is sought that fulfills $\bar{\boldsymbol{u}} = \bar{\boldsymbol{u}}_0$ on $\partial\bar{\Omega}^D$. To solve for the displacements, $\bar{\boldsymbol{u}}$, the linearization of $\bar{G}$ in the direction $\Delta\bar{\boldsymbol{u}}$ around the current deformation $\bar{\boldsymbol{u}}$ is needed,

$$\begin{aligned}
D\bar{G} \cdot \Delta\bar{\boldsymbol{u}} &= \int_{\bar{\Omega}} \frac{\partial(\delta\bar{\boldsymbol{u}})}{\partial\bar{\boldsymbol{X}}} : \frac{\partial\bar{\boldsymbol{P}}}{\partial\bar{\boldsymbol{F}}} : (D\bar{\boldsymbol{F}} \cdot \Delta\bar{\boldsymbol{u}})dV \\
&= \int_{\bar{\Omega}} \frac{\partial(\delta\bar{\boldsymbol{u}})}{\partial\bar{\boldsymbol{X}}} : \bar{\boldsymbol{A}} : \frac{\partial\Delta\bar{\boldsymbol{u}}}{\partial\bar{\boldsymbol{X}}}dV,
\end{aligned} \tag{4}$$

where $\bar{\boldsymbol{A}} := \dfrac{\partial\bar{\boldsymbol{P}}}{\partial\bar{\boldsymbol{F}}}$ is the macroscopic stiffness tensor. To model a certain material behavior, a constitutive relation which connects $\bar{\boldsymbol{P}}$ and $\bar{\boldsymbol{A}}$ to the deformation $\bar{\boldsymbol{F}}$ has to be provided.

2.2. Microscopic simulation

The microscopic simulation is defined on a representative volume element (RVE) Ω. In first-order CH, it is assumed that the microscopic deformed points $\boldsymbol{x}$ are coupled to the macroscopic deformation by

$$\boldsymbol{x} := \bar{\boldsymbol{F}}\boldsymbol{X} + \boldsymbol{w}(\boldsymbol{X}), \tag{5}$$

where $\boldsymbol{w}(\boldsymbol{X})$ is a zero-mean microscopic fluctuation displacement field. Note that the macroscopic deformation gradient $\bar{\boldsymbol{F}}$ depends on the macroscopic coordinates $\bar{\boldsymbol{X}}$. For conciseness, this dependence is omitted here and the following equations are given for a fixed $\bar{\boldsymbol{X}}$. The microscopic displacement is then given as

$$\boldsymbol{u} := (\bar{\boldsymbol{F}} - \mathbf{I})\boldsymbol{X} + \boldsymbol{w}(\boldsymbol{X}), \tag{6}$$

and the microscopic deformation gradient as

$$\boldsymbol{F} := \frac{\partial\boldsymbol{x}}{\partial\boldsymbol{X}} = \bar{\boldsymbol{F}} + \frac{\partial\boldsymbol{w}}{\partial\boldsymbol{X}}. \tag{7}$$

Analogous to the above, the microscopic deformation is governed by the quasi-static linear momentum balance,

$$\begin{aligned}
\text{Div}\,\boldsymbol{P} &= \boldsymbol{0} && \text{on } \Omega, \\
\boldsymbol{P}\boldsymbol{N} &= \boldsymbol{t}_0 && \text{on } \partial\Omega^N, \text{ and} \\
\boldsymbol{u} &= \boldsymbol{u}_0 && \text{on } \partial\Omega^D,
\end{aligned} \tag{8}$$

where body forces are neglected for simplicity. The weak form is given as

$$G = \int_{\Omega} \frac{\partial\delta\boldsymbol{u}}{\partial\boldsymbol{X}} : \boldsymbol{P}dV - \int_{\partial\Omega^N} \boldsymbol{t}_0 \cdot \delta\boldsymbol{u}dA = 0, \qquad \forall\delta\boldsymbol{u} \in H_0^1(\Omega), \tag{9}$$

and the linearization of G in the direction $\Delta\boldsymbol{u}$ around the current state $\boldsymbol{u}$ is

$$\begin{aligned}
DG \cdot \Delta\boldsymbol{u} &= \int_{\Omega} \frac{\partial(\delta\boldsymbol{u})}{\partial\boldsymbol{X}} : \frac{\partial\boldsymbol{P}}{\partial\boldsymbol{F}} : (D\boldsymbol{F} \cdot \Delta\boldsymbol{u})dV \\
&= \int_{\bar{\Omega}} \frac{\partial(\delta\boldsymbol{u})}{\partial\boldsymbol{X}} : \boldsymbol{A} : \frac{\partial\Delta\boldsymbol{u}}{\partial\boldsymbol{X}}dV.
\end{aligned} \tag{10}$$

In contrast to the macroscale, the materials are fully specified by a constitutive law in the microscale. In this work, we consider a hyperelastic Neo-Hookean material model with strain energy density function

$$W(\boldsymbol{F}, \boldsymbol{\mu}) = C_1(\text{Tr}(\boldsymbol{C}) - 3 - 2\ln J) + D_1(J - 1)^2, \tag{11}$$
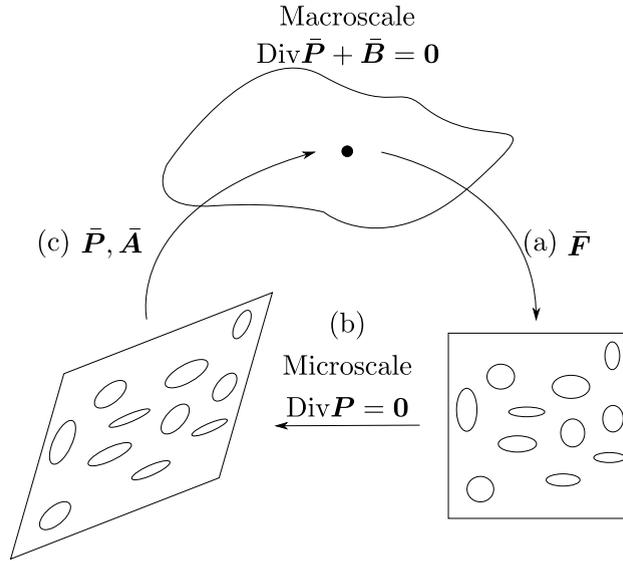
Fig. 1. Scale coupling. (a) The macroscale simulation transfers a macroscopic deformation gradient to the microscale. (b) The microscopic response is subsequently computed and (c) the effective stress and stiffness are transferred back to the macroscale. The macroscopic solver does not require the microscopic deformation, hence it is beneficial to directly construct a surrogate model for the microscopic stress.

where C_1 and D_1 are the material parameters stored in $\boldsymbol{\mu} = [C_1, D_1]^T$, Tr($\bullet$) denotes the trace of a tensor, $\boldsymbol{C} = \boldsymbol{F}^T \boldsymbol{F}$ the right Cauchy–Green deformation tensor and $J = \det(\boldsymbol{F})$ is the determinant of $\boldsymbol{F}$ respectively. The material parameters C_1 and D_1 are related to the Young's modulus E and Poisson ratio ν by

$$E = \frac{2C_1(3D_1 + 2C_1)}{C_1 + D_1}, \quad \nu = \frac{D_1}{2(C_1 + D_1)}. \tag{12}$$

The PK1 stress and stiffness tensor can then be derived from Eq. (11),

$$\boldsymbol{P} = \frac{\partial W}{\partial \boldsymbol{F}}, \quad \boldsymbol{A} = \frac{\partial \boldsymbol{P}}{\partial \boldsymbol{F}} = \frac{\partial^2 W}{\partial \boldsymbol{F}^2}. \tag{13}$$

2.3. Scale coupling

The aim of the microscopic simulation is to replace the macroscopic constitutive model $\bar{\boldsymbol{P}}(\bar{\boldsymbol{X}}, \bar{\boldsymbol{F}})$. For every quadrature point in the macrostructure, the macroscopic simulation transfers the macroscopic deformation gradient to the microscopic simulation, which in turn returns an effective stress and stiffness tensor; see Fig. 1 for a visualization. The displacements of both scales are coupled according to Eq. (6). Furthermore, the **Hill–Mandel Condition**,

$$\langle \boldsymbol{P} : \delta \boldsymbol{F} \rangle = \bar{\boldsymbol{P}} : \delta \bar{\boldsymbol{F}}, \tag{14}$$

states that the virtual work exerted on both scales has to be the same; here, $\langle \{\bullet\} \rangle = |\Omega|^{-1} \int_\Omega \{\bullet\} dV$ denotes the volumetric averaging operator with $|\Omega|$ the volume of the RVE Ω.

It can be shown that the condition is always fulfilled by introducing appropriate boundary conditions for the microscopic problem. One set of boundary conditions that fulfill Eq. (14) are linear displacement boundary conditions, where the fluctuation displacement $\boldsymbol{w}$ on the RVE boundary is assumed to be zero. With Eq. (5), the following Dirichlet boundary condition follows:

$$\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{X} = (\bar{\boldsymbol{F}} - \mathbf{I})\boldsymbol{X} \quad \text{on } \partial \Omega. \tag{15}$$

In [14], it is shown that, by prescribing Eq. (15) on the boundaries, the macroscopic deformation gradient is always equal to the averaged microscopic deformation gradient, i.e. $\bar{\boldsymbol{F}} = \langle \boldsymbol{F} \rangle$. It then follows that the averaged

microscopic stress has to be equal to the macroscopic stress:

$$\langle \boldsymbol{P} \rangle = \left\langle \frac{\partial W}{\partial \boldsymbol{F}} \right\rangle = \frac{\partial \bar{W}}{\partial \bar{\boldsymbol{F}}} : \frac{\partial \bar{\boldsymbol{F}}}{\partial \boldsymbol{F}} = \frac{\partial \bar{W}}{\partial \bar{\boldsymbol{F}}} = \bar{\boldsymbol{P}}. \tag{16}$$

However, the averaged stiffness tensor is generally not equal to the macroscopic stiffness tensor due to the changes of $\boldsymbol{w}$ inside Ω, i.e.,

$$\langle \boldsymbol{A} \rangle \neq \bar{\boldsymbol{A}}, \tag{17}$$

thus complicating the calculation of the correct effective stiffness [14]. A popular way to compute the stiffness with a perturbation method is suggested in [17], which essentially approximates the stiffness with a finite difference scheme.

3. Reduced order model

The two-scale simulation presented above can be solved using the finite element method (FEM) on both scales, leading to a nested FE scheme, also known as FE². Although effective, FE² method tends to be computationally very expensive. To lower the computational cost, we approximate the microscopic problem using a non-intrusive reduced order model, constructed by a proper orthogonal decomposition (POD) and a Gaussian process regression (GPR), which we will refer to as P-PODGPR later. The proposed method differs from existing works in the literature (see e.g. [54,56,57,64]) in that the non-intrusive reduced order model is constructed directly for the stresses instead of the displacements or strains. As seen in Fig. 1, both displacement and strain are not required by the macroscopic solver and, moreover, the stresses still need to be calculated, which means that the microscopic constitutive law has to be implemented anew, negating the advantages of a non-intrusive method. Moreover, depending on the material model the evaluation might be costly and compromise the speed up significantly. It is also not obvious how one would derive the effective stiffness for that case. On the other hand, reducing the stresses directly circumvents the above-mentioned disadvantages, and both the effective stress and stiffness can be rapidly obtained.

3.1. Proper orthogonal decomposition (POD)

Principal component analysis (PCA) is a powerful tool in data science to find a low-dimensional approximation in Euclidean space. By introducing an appropriate function metric, PCA can be generalized for functions in Hilbert spaces and is referred to as proper orthogonal decomposition (POD). POD can utilize the correlation of solutions of a problem for different parameters to find a low-dimensional orthonormal basis, thus reducing the number of unknowns to a very small fraction of the dimension of the original high-fidelity model. The procedure for computing the low-dimensional basis of the microscopic stress $\boldsymbol{P} \in L^2(\Omega)$ is outlined in the following.

3.1.1. Basis construction

After N_{pod} high-fidelity simulations for different parameters have been carried out, the microscopic stress in all quadrature points $\mathbf{P} \in \mathbb{R}^{N_{\text{qp}} \times 3 \times 3}$, where N_{qp} is the total number of quadrature points in the high-fidelity model, is collected for each solution in the snapshot tensor $\mathbf{S} \in \mathbb{R}^{N_{\text{qp}} \times 3 \times 3 \times N_{\text{pod}}}$,

$$\mathbf{S} = \left[\mathbf{P}^{(1)}, \mathbf{P}^{(2)}, \ldots, \mathbf{P}^{(N_{\text{pod}})} \right]. \tag{18}$$

Next, the correlation matrix $\mathbf{C} \in \mathbb{R}^{N_{\text{pod}} \times N_{\text{pod}}}$ can be formed by computing the L^2-inner product between every pair of two snapshots,

$$\begin{aligned} \mathbf{C}_{ij} &= \left(\boldsymbol{P}^{(i)}, \boldsymbol{P}^{(j)} \right)_{L^2(\Omega)} \\ &= \int_{\Omega} \boldsymbol{P}^{(i)} : \boldsymbol{P}^{(j)} dV \\ &= \sum_{q=1}^{N_{\text{qp}}} w_q \mathbf{P}_q^{(i)} : \mathbf{P}_q^{(j)}, \end{aligned} \tag{19}$$

where $(\bullet, \bullet)_{L^2(\Omega)}$ denotes the L^2-inner product and $i, j = 1, 2, \ldots, N_{\text{pod}}$. The weights w_q are the integration weights corresponding to each quadrature point; they only depend on the mesh and can be easily computed. Note the

subtle difference between italic and nonitalic characters used, corresponding to $L^2(\Omega)$ functions and their discrete counterparts respectively.

After obtaining the correlation matrix, its eigenvalues λ_l and eigenvectors $\mathbf{v}_l$, $l = 1, 2, \ldots, \min(L, N_{\text{pod}})$ are computed, where L is the maximum number of basis functions specified by the user. The number of basis functions is often chosen from the criterion

$$\frac{\sum_{l=1}^{L} \lambda_l}{\sum_{l=1}^{N_{\text{pod}}} \lambda_l} > \mathcal{E}_{\text{pod}}, \tag{20}$$

where $\mathcal{E}_{\text{pod}}$ corresponds to the energy captured by L basis functions and is specified by the user. Moreover, the eigenvalues can reveal if a problem is reducible using POD: if the eigenvalues decay rapidly, the solution space can be accurately captured by a few basis functions. Finally, the lth POD basis function $\mathbf{B}_l \in \mathbb{R}^{N_{\text{qp}} \times 3 \times 3}$ is found with

$$\mathbf{B}_l = \frac{1}{\sqrt{\lambda_l}} \mathbf{S} \mathbf{v}_l. \tag{21}$$

By construction, the POD basis is orthonormal, i.e.

$$(\boldsymbol{B}_i, \boldsymbol{B}_j)_{L^2(\Omega)} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \tag{22}$$

3.1.2. Approximation of stress and stiffness

With the POD basis functions, $\mathbf{P}$ can be approximated with

$$\mathbf{P}^{\text{RB}}(\bar{\boldsymbol{F}}, \boldsymbol{\mu}) = \sum_{l=1}^{L} \alpha_l(\bar{\boldsymbol{F}}, \boldsymbol{\mu}) \mathbf{B}_l, \tag{23}$$

where the loading $\bar{\boldsymbol{F}}$ and material parameters $\boldsymbol{\mu}$ are listed separately, as the partial derivative over the loading parameters yields the stiffness tensor, whereas the partial derivative with respect to $\boldsymbol{\mu}$ could be used for optimization or inverse problems. Since the basis functions $\mathbf{B}_l$ are constant, only the coefficients $\alpha_l \in \mathbb{R}$ depend on the parameters and need to be found for a new parameter value. Note that with this approach the ability to determine the microscopic displacement, energy and stiffness tensor is lost, if the constitutive relation $\boldsymbol{P}(\boldsymbol{F})$ cannot be inverted. However, if knowledge of the deformation is required, a similar procedure using displacement or deformation gradient snapshots could be implemented [54,56]. This should in general work well if the stresses can be approximated accurately, because the stress is a non-linear function of the displacement and therefore should be more difficult to approximate.

By averaging the microscopic stress, the macroscopic stress $\bar{\boldsymbol{P}} \in \mathbb{R}^{3 \times 3}$ is found:

$$\begin{aligned} \bar{\boldsymbol{P}} &= \langle \boldsymbol{P}^{\text{RB}} \rangle \\ &= |\Omega|^{-1} \int_{\Omega} \boldsymbol{P}^{\text{RB}} dV \\ &= |\Omega|^{-1} \sum_{q=1}^{N_{\text{qp}}} w_q \mathbf{P}_q^{\text{RB}} \\ &= |\Omega|^{-1} \sum_{q=1}^{N_{\text{qp}}} w_q \left(\sum_{l=1}^{L} \alpha_l(\bar{\boldsymbol{F}}, \boldsymbol{\mu}) \mathbf{B}_{lq} \right). \end{aligned} \tag{24}$$

Lastly, the effective stiffness $\bar{A} \in \mathbb{R}^{3 \times 3 \times 3 \times 3}$ is determined by differentiating the macroscopic stress over the macroscopic deformation gradient,

$$\begin{aligned} \bar{A} &= \frac{\partial \bar{\boldsymbol{P}}}{\partial \bar{\boldsymbol{F}}} \\ &= |\Omega|^{-1} \frac{\partial \left(\sum_{q=1}^{N_{\text{qp}}} w_q \left(\sum_{l=1}^{L} \alpha_l(\bar{\boldsymbol{F}}, \boldsymbol{\mu}) \mathbf{B}_{lq} \right) \right)}{\partial \bar{\boldsymbol{F}}} \\ &= |\Omega|^{-1} \sum_{q=1}^{N_{\text{qp}}} w_q \left(\sum_{l=1}^{L} \mathbf{B}_{lq} \otimes \frac{\partial \alpha_l(\bar{\boldsymbol{F}}, \boldsymbol{\mu})}{\partial \bar{\boldsymbol{F}}} \right). \end{aligned} \tag{25}$$

As seen above, by introducing the approximation Eq. (23), the macroscopic stiffness tensor can be naturally derived when the derivative of the coefficients α_l with respect to $\bar{F}$ is available. Moreover, the derivatives of stress with respect to the material parameters μ can be obtained analogously and used in optimization or inverse problems.

Remark. As this surrogate model is used to replace the microstructural simulation, it is sufficient to construct the basis for macroscopic stretches $\bar{U}$ instead of $\bar{F}$, since the macroscopic deformation gradient can be split into $\bar{F} = \bar{R}\bar{U}$ by using the polar decomposition, where $\bar{R}$ accounts for rotations, see [65]. Due to the symmetry of $\bar{U}$, the number of parameters reduces from 4 to 3 in 2D or from 9 to 6 in 3D. In the following, $\bar{F}$ is therefore replaced by $\bar{U}$ and all snapshots are assumed to have been obtained for different $\bar{U}$ and μ.

3.2. Regression model

The mapping $\alpha_l(\bar{U}, \mu)$ for all parameters can be constructed by means of a regression, using the training data that have been collected for POD. If the regression quality is insufficient, more data has to be generated.

3.2.1. Data preparation

The stress data $\mathbf{P}^{(i)}$ collected was generated for parameters $(\bar{U}^{(i)}, \mu^{(i)})$ with $i = 1, 2, \ldots, N_{\text{reg}}$, where N_{reg} denotes the total number of snapshots available. The best approximation of the ith snapshot on the POD basis is given by

$$\mathbf{P}^{(i)} \approx \sum_{l=1}^{L} (\boldsymbol{P}^{(i)}, \boldsymbol{B}_l)_{L^2(\Omega)} \mathbf{B}_l, \tag{26}$$

and hence $\alpha_l^{(i)} = (\boldsymbol{P}^{(i)}, \boldsymbol{B}_l)_{L^2(\Omega)}$. All coefficients for all snapshots $\alpha_l^{(i)}$ are collected together with the parameter data; a mapping between the parameters and the POD coefficients is then constructed by regression, i.e.

$$\hat{\alpha}_l : \mathcal{P} \to \mathbb{R}, (\bar{U}, \mu) \mapsto \hat{\alpha}_l(\bar{U}, \mu), \tag{27}$$

where $\mathcal{P}$ denotes the parameter space. Due to the orthonormality of the POD basis functions, see Eq. (22), the coefficients α_l are fully uncorrelated and a separate regression for each POD coefficient can be constructed independently. With this mapping, the stress $\mathbf{P}$ is approximated as

$$\hat{\mathbf{P}}^{\text{RB}}(\bar{U}, \mu) = \sum_{l=1}^{L} \hat{\alpha}_l(\bar{U}, \mu)\mathbf{B}_l. \tag{28}$$

3.2.2. Gaussian process regression

Since the material model considered here is hyperelastic and therefore history independent, each stress corresponds to exactly one set of deformations and therefore, a wide range of regression techniques can be used. Many different regression approaches have been used to approximate the coefficients, e.g. radial basis functions [58], neural networks (NNs) [55], and Gaussian process regression (GPR) [54,56,59]. A systematic investigation on these three methods has been conducted in [66]. A comparison of different machine learning methods for regression was performed in [57]. For this work, GPR is used since it offers some desirable properties:

1. It can reconstruct the training data perfectly, i.e. it reproduces the exact solution at the training points.
2. Depending on the chosen kernel, the regression function has a specified smoothness and its derivatives can be obtained analytically.
3. The trained GPR model returns an uncertainty measure for each evaluation, which can be used to estimate the regression error or to develop an active learning scheme [54,67].

In GPR, a scalar regression function $f(\mathbf{X})$ with $\mathbf{X} \in \mathbb{R}^d$ is assumed to be distributed as a Gaussian process (GP) with a zero mean function and kernel $k_\theta(\mathbf{X}, \mathbf{X}')$,

$$f \sim \mathcal{GP}(0, k_\theta(\mathbf{X}, \mathbf{X}')). \tag{29}$$

The form of the kernel $k_\theta(\mathbf{X}, \mathbf{X}')$ has to be chosen by the user, and each kernel has hyperparameters θ that must be fitted to the data. In this work, we use, as in [54], the *automatic relevance determination* (ARD) squared exponential

kernel,

$$k_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{X}') = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{k=1}^{d}\frac{(X_k - X_k')^2}{l_k^2}\right), \tag{30}$$

where $\boldsymbol{\theta} = [\sigma_f, l_1, l_2, \ldots, l_d]$.

Given a finite number of data points $\{\mathbf{X}^{(i)}, f(\mathbf{X}^{(i)})\}_{i=1}^{N_{\mathrm{reg}}}$, the optimal hyperparameters $\boldsymbol{\theta}$ can be determined with a maximum likelihood estimation [54,68], hence yielding a prior GP f. By Bayesian inference, the posterior GP $f^*(\mathbf{x})$ can then be given as

$$f^*|(\mathbf{X}, \mathbf{y}) \sim \mathcal{GP}(m^*, k^*), \tag{31}$$

$$m^*(\mathbf{x}) = \mathbf{k}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{x})^T \mathbf{k}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{X})^{-1}\mathbf{y}(\mathbf{X}), \tag{32}$$

$$k^*(\mathbf{x}, \mathbf{x}') = k_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') - \mathbf{k}_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x})^T \mathbf{k}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{X})^{-1}\mathbf{k}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{x}'), \tag{33}$$

where $\mathbf{k}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{x}) = [k_{\boldsymbol{\theta}}(\mathbf{X}^{(1)}, \mathbf{x}), k_{\boldsymbol{\theta}}(\mathbf{X}^{(2)}, \mathbf{x}), \ldots, k_{\boldsymbol{\theta}}(\mathbf{X}^{(N_{\mathrm{reg}})}, \mathbf{x})]^T$, $\mathbf{k}_{\boldsymbol{\theta}}(\mathbf{X}, \mathbf{X}) = [k_{\boldsymbol{\theta}}(\mathbf{X}^{(i)}, \mathbf{X}^{(j)})]_{i,j=1}^{N_{\mathrm{reg}}}$, and $\mathbf{y}(\mathbf{X}) = [f(\mathbf{X}^{(1)}), f(\mathbf{X}^{(2)}), \ldots, f(\mathbf{X}^{(N_{\mathrm{reg}})})]^T$. For any arbitrary input $\mathbf{x}$, a distribution $\mathcal{N}(m^*(\mathbf{x}), k^*(\mathbf{x}, \mathbf{x}))$ can thus be obtained.

A GPR is separately performed for each POD basis coefficient, yielding in total L GPR models. The ARD kernel is infinitely smooth, hence the Jacobian of α_l can be easily obtained and Eq. (25) is fully specified. With the trained GPR models and the POD basis, the surrogate model for the microscopic simulation is complete. Given a parameter, the microscopic stress field can be rapidly evaluated, from which the effective stress and its derivative can be derived. Therefore the surrogate model can be considered a learned constitutive model, where additionally the stress field on the microscale is still accessible.

As indicated earlier, the non-intrusive reduced order model will be referred to as P-PODGPR.

3.2.3. Error estimate

For brevity of notation, we use the symbol $\boldsymbol{\pi}$ to denote both the loading and material parameters in this section. Using the Cauchy–Schwarz inequality, it can be shown that the total error between the high fidelity simulation and the approximation by P-PODGPR for a parameter $\boldsymbol{\pi} \in \mathcal{P}$ consists of the projection error and regression error,

$$\begin{aligned}
\epsilon_{\mathrm{P\text{-}PODGPR}}(\boldsymbol{\pi}) &= \|\boldsymbol{P}^{\mathrm{HF}}(\boldsymbol{\pi}) - \hat{\boldsymbol{P}}^{\mathrm{RB}}(\boldsymbol{\pi})\|_{L^2(\Omega)} \\
&= \|\boldsymbol{P}^{\mathrm{HF}}(\boldsymbol{\pi}) - \boldsymbol{P}^{\mathrm{RB}}(\boldsymbol{\pi}) + \boldsymbol{P}^{\mathrm{RB}}(\boldsymbol{\pi}) - \hat{\boldsymbol{P}}^{\mathrm{RB}}(\boldsymbol{\pi})\|_{L^2(\Omega)} \\
&\leq \underbrace{\|\boldsymbol{P}^{\mathrm{HF}}(\boldsymbol{\pi}) - \boldsymbol{P}^{\mathrm{RB}}(\boldsymbol{\pi})\|_{L^2(\Omega)}}_{\text{projection error}} + \underbrace{\|\boldsymbol{P}^{\mathrm{RB}}(\boldsymbol{\pi}) - \hat{\boldsymbol{P}}^{\mathrm{RB}}(\boldsymbol{\pi})\|_{L^2(\Omega)}}_{\text{regression error}}.
\end{aligned} \tag{34}$$

If the parameter space has been appropriately explored by the snapshots used for POD and the eigenvalues exhibit a rapid decay, the projection error can be reasonably expected to quickly approach zero for any parameter value as the number of basis functions L is increased.

The squared regression error for the stresses can be rearranged,

$$\begin{aligned}
\|\boldsymbol{P}^{\mathrm{RB}}(\boldsymbol{\pi}) - \hat{\boldsymbol{P}}^{\mathrm{RB}}(\boldsymbol{\pi})\|_{L^2(\Omega)}^2 &= \left\|\sum_{l=1}^{L}(\alpha_l - \hat{\alpha}_l)\boldsymbol{B}_l\right\|_{L^2(\Omega)}^2 \\
&= \left(\sum_{l=1}^{L}(\alpha_l - \hat{\alpha}_l)\boldsymbol{B}_l, \sum_{l'=1}^{L}(\alpha_{l'} - \hat{\alpha}_{l'})\boldsymbol{B}_{l'}\right)_{L^2(\Omega)} \\
&= \sum_{l=1}^{L}\sum_{l'=1}^{L}(\alpha_l - \hat{\alpha}_l)(\alpha_{l'} - \hat{\alpha}_{l'})(\boldsymbol{B}_l, \boldsymbol{B}_{l'})_{L^2(\Omega)} \\
&= \sum_{l=1}^{L}(\alpha_l - \hat{\alpha}_l)^2 \\
&= \|\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}\|_2^2,
\end{aligned} \tag{35}$$

where the orthonormality of $\boldsymbol{B}_l$ was used in the second last line and $\|\bullet\|_2$ denotes the Euclidean norm. Hence, the regression error is simply equal to the Euclidean norm of the error of the coefficient vectors. This result is analogous to the result presented in [56] for the displacement.

3.3. Comparison to neural networks

After the pioneering works by Ghaboussi et al. [25,26] and recent advances in Deep Learning, many papers have used methods of deep learning to extract a constitutive model from pairs of stress and deformation data by training a deep neural network (DNN), e.g. [27,30,33,69,70]. Frequently their data is collected from RVE simulations by applying boundary conditions that fulfill Eq. (14) and solving the microscopic simulation to generate training data, analogous to what is done in this work. However, one difference is that in this work the stress field is collected while the other works only collect the averaged stress and dispose of the stress field. A brief comparison between both approaches is given below:

1. **Training**: The DNN approach essentially performs a regression on the stress and deformation data. This means that a mapping from $\mathbb{R}^{3\times3} \to \mathbb{R}^{3\times3}$ (or $\mathbb{R}^6 \to \mathbb{R}^6$ in the case where the stress and deformation measures are symmetric) has to be learned. P-PODGPR, on the other hand, uses the correlation of the microscopic stress field solutions and finds a few uncorrelated coefficients which can be independently learned, i.e. a mapping from $\mathbb{R}^{3\times3} \to \mathbb{R}$ or $\mathbb{R}^6 \to \mathbb{R}$ is learned.
2. **Implementation**: Both surrogate models, after they have been trained, can be easily adopted into any simulation software as they are both non-intrusive and therefore entirely independent of the software used to solve the microscale simulation.
3. **Evaluation**: The DNN approach needs to compute one forward pass through the neural network to get the effective stress for a given deformation. The effective stiffness can then be computed with one backward pass with automatic differentiation. However, the microstructural stresses, which are desirable for finding local stress concentrations and designing improved microstructures, cannot be obtained. In the P-PODGPR method the stress field is fully obtained by evaluating L GPR models, from which the effective stress can be computed. The effective stiffness is obtained with Eq. (25) as the GPR model also supports an analytical way of computing the derivative of α_l over the deformation.

4. Results

To demonstrate the performance of P-PODGPR and the influence of the number of basis functions L, number of samples used for the basis construction N_{pod}, and number of samples used for the regression N_{reg}, two single-scale examples with different 2D RVEs are presented. The third example involves a two-scale problem, in which the results obtained with P-PODGPR are compared with the FE^2 solution.

4.1. Single-scale RVE simulation

To measure the accuracy of P-PODGPR on test data, the following relative error measure for the effective stress is defined:

$$\epsilon_{\bar{\mathbf{P}}} := \frac{\left\| \bar{\mathbf{P}}^{\text{HF}} - \hat{\bar{\mathbf{P}}}^{\text{RB}} \right\|_{\text{F}}}{\left\| \bar{\mathbf{P}}^{\text{HF}} \right\|_{\text{F}}}, \tag{36}$$

where $\|\bullet\|_{\text{F}}$ denotes the Frobenius norm, $\bar{\mathbf{P}}^{\text{HF}}$ the effective stress from the high-fidelity simulation and $\hat{\bar{\mathbf{P}}}^{\text{RB}}$ the effective stress resulting from P-PODGPR. The projection error of the basis is attained when $\hat{\bar{\mathbf{P}}}^{\text{RB}} = \bar{\mathbf{P}}^{\text{RB}}$, where $\bar{\mathbf{P}}^{\text{RB}}$ is the effective stress after projecting the high fidelity solution onto the POD basis when the solution is known. As the regression error Eq. (35) goes towards zero, $\hat{\bar{\mathbf{P}}}^{\text{RB}}$ tends towards $\bar{\mathbf{P}}^{\text{RB}}$, so the projection error can be interpreted as a lower bound of the solution from P-PODGPR. All microscopic simulations were conducted with the FE framework MOOSE [71].
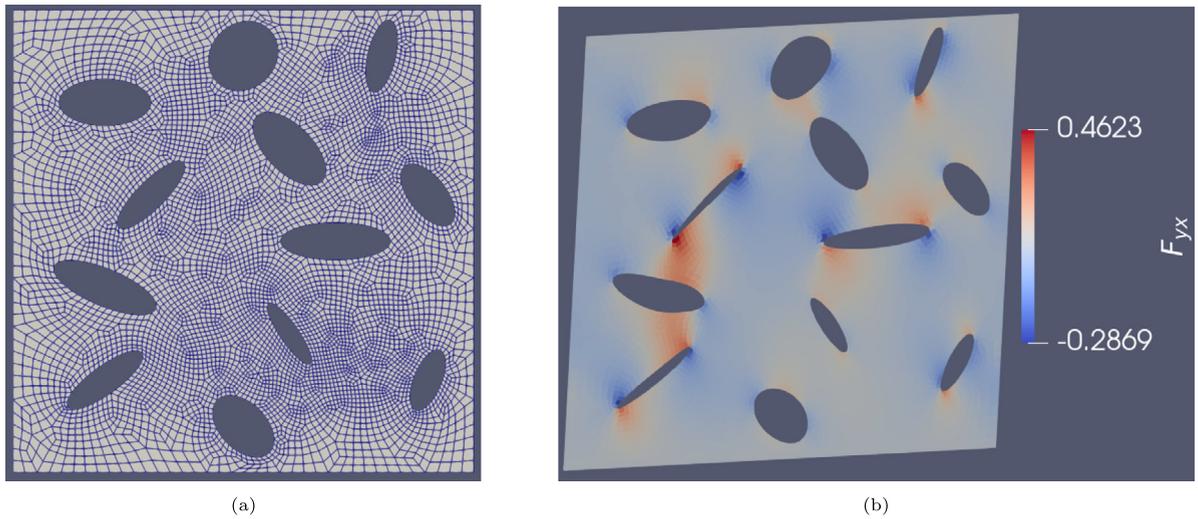
Fig. 2. Porous material: (a) The considered microstructure consists of a Neo-Hookean matrix material with a volume ratio of 86%. The mesh consists of 6000 bilinear elements and 6293 nodes. (b) An exemplary deformation with $[\bar{U}_{xx}, \bar{U}_{yy}, \bar{U}_{xy}] = [0.95, 0.95, 0.05]$. The yx-component of the local deformation gradient ranges from $[-0.2869, 0.4623]$ and is hence much larger than the prescribed boundary deformation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.1.1. Porous material

A porous microstructure as depicted in Fig. 2(a) is considered for the first example, where the pores account for 14% of the total area. Four-node quadrilateral elements with four quadrature points are employed. The matrix material is modeled as a Neo-Hookean material with material parameters $C_1 = 1$ and $D_1 = 1$. Linear displacement boundaries with $\bar{U} - \mathbf{I} \in [-0.05, 0.05]$ are considered. Due to the geometry of the problem, such a deformation leads to much higher local strains, see Fig. 2(b) for an exemplary deformation with $[\bar{U}_{xx}, \bar{U}_{yy}, \bar{U}_{xy}] = [0.95, 0.95, 0.05]$. If larger displacements on the boundaries were considered, some pores might close, requiring contact detection. In this example, the material parameters $\boldsymbol{\mu}$ are fixed, hence this problem has three varying parameters.

Data generation. To investigate the number of precomputations needed for an accurate representation, a set of 500 training snapshots for training P-PODGPR was sampled via a Sobol sequence sampling procedure and another set of 1000 test snapshots, to evaluate the accuracy of the surrogate model, was generated randomly from a uniform distribution. It was observed in [72] that Sobol sequence sampling fills the parameter space more evenly as compared to random sampling and thus provides better results.

Eigenvalues. The eigenvalues of the correlation matrix for different numbers of training snapshots N_{pod} are plotted in Fig. 3. For all cases, an exponential decay is observed, indicating the reducibility of the problem. The von Mises stress of the first three POD basis functions is plotted in Fig. 4. It can be seen that the basis functions specifically capture the stress concentrations around the pores. Subsequently, $L = 20$ basis functions are considered which corresponds to an energy $\mathcal{E}_{pod}$ of 99.9996%.

Influence of N_{pod} and N_{reg}. For $L = 20$ basis functions, combinations of $N_{pod} \in \{20, 50, 100, 200, 500\}$ and $N_{reg} \in \{50, 100, 200, 300, 400, 500\}$ were investigated. All 1000 test data were evaluated and the resulting mean and maximum error of the stresses are plotted in Fig. 5. It can be observed that for all cases the mean error is below 0.1%. From $N_{pod} = 20$ to $N_{pod} = 50$, a significant improvement in the error can be observed. However, for $N_{pod} > 50$, the error only changes marginally. As N_{reg} is increased, more data becomes available for the regression and therefore the mapping in Eq. (27) becomes increasingly more accurate. Nevertheless, even using only 50 snapshots for both basis and regression yields a mean error of 0.1% and a maximum error of roughly 0.65%.

Influence of L. Using $N_{pod} = N_{reg} = 50$, the influence of the number of basis functions L was then investigated. The relative errors of the effective stress are given in Fig. 6. The projection error is also plotted to compare the quality of the regression in Eq. (27). As seen from the figure, both the mean and maximum error of the first 8 basis
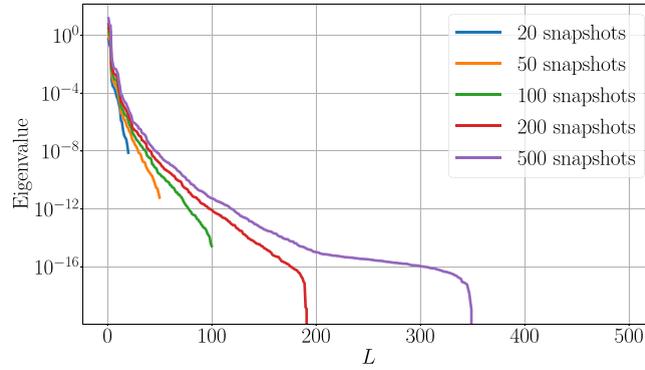
Fig. 3. Porous material: Eigenvalues of the correlation matrix for different numbers of snapshots N_{pod} used for POD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

| (a) 1st basis function | (b) 2nd basis function | (c) 3rd basis function |
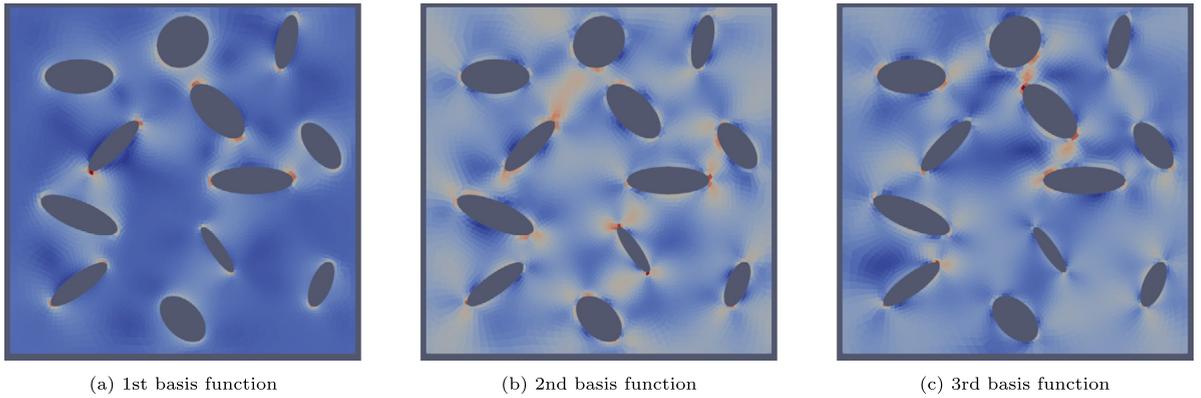
Fig. 4. Porous material: Von Mises stress of the first three POD basis functions of the microscopic stress field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Porous material: Validation and effective stress error for different feedforward neural network architectures. The lowest values in each column have been highlighted.

Architecture	Validation loss	$\epsilon_{\mathbf{P}}^{mean}$	$\epsilon_{\mathbf{P}}^{max}$
$N_h = 1, N_n = 20$	2.33×10^{-7}	**0.0016**	0.0199
$N_h = 1, N_n = 50$	$\mathbf{1.72 \times 10^{-7}}$	0.0017	**0.0136**
$N_h = 2, N_n = 20$	2.4×10^{-7}	0.0019	0.0184
$N_h = 2, N_n = 50$	1.94×10^{-7}	0.0018	0.0166

functions nearly perfectly match. However, for a larger number of basis functions, the discrepancy slowly grows and the error flattens. Generally, the POD coefficients get increasingly more oscillatory with increasing number and hence require more data for an accurate regression. To show that with increasing N_{reg}, the regression error decreases and approaches the projection error, the error over L for $N_{pod} = 50$ with $N_{reg} = 500$ is also plotted in Fig. 7. For this case, the mean error matches the projection error for 13 basis functions, while the maximum error also gets much closer to the projection error. Note that the maximum error of P-PODGPR is sometimes slightly lower than the projection error. This can be explained with errors in the regression that lead to a stress field which after averaging ends up closer to the high fidelity solution than the best projection. This is entirely random and as N_{reg} is further increased, the regression error will tend towards zero and the curves will match eventually.

Comparison with neural networks. For comparison, a deep forward neural network with different architectures was trained with the available deformation and stress data collected from the RVE simulations. The stress field
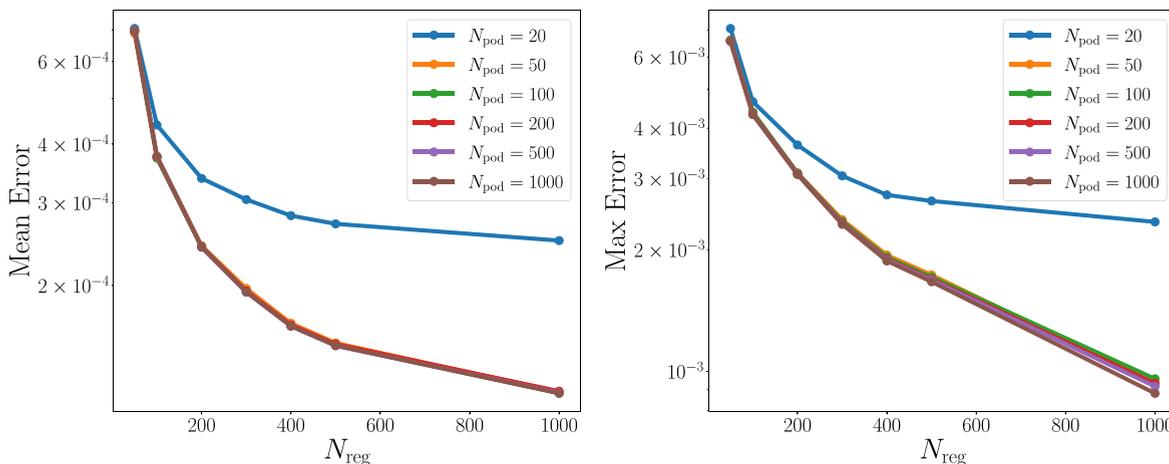
Fig. 5. Porous material: Comparison of stress errors for different combinations of N_{pod} and N_{reg} for $L = 20$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
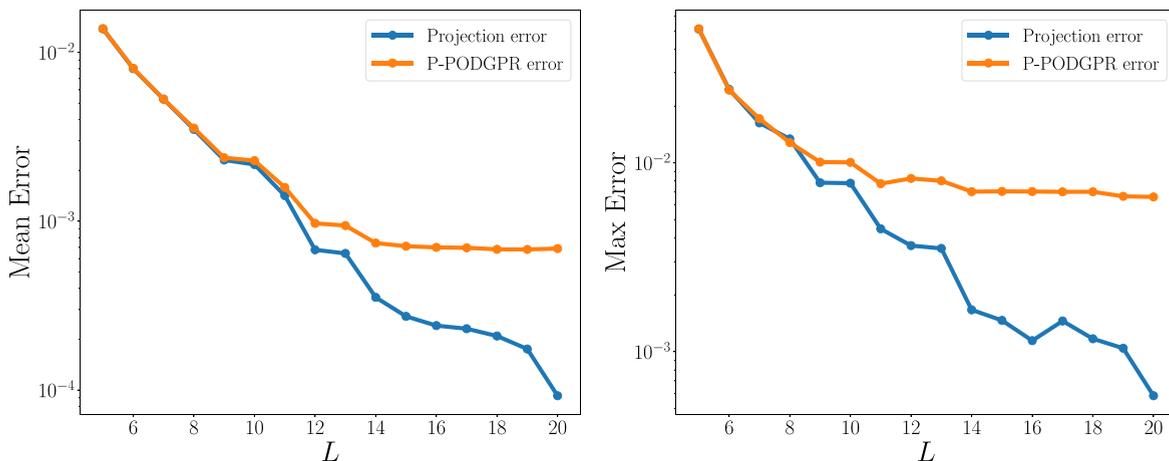
Fig. 6. Porous material: Comparison of stress errors for different L with $N_{pod} = N_{reg} = 50$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
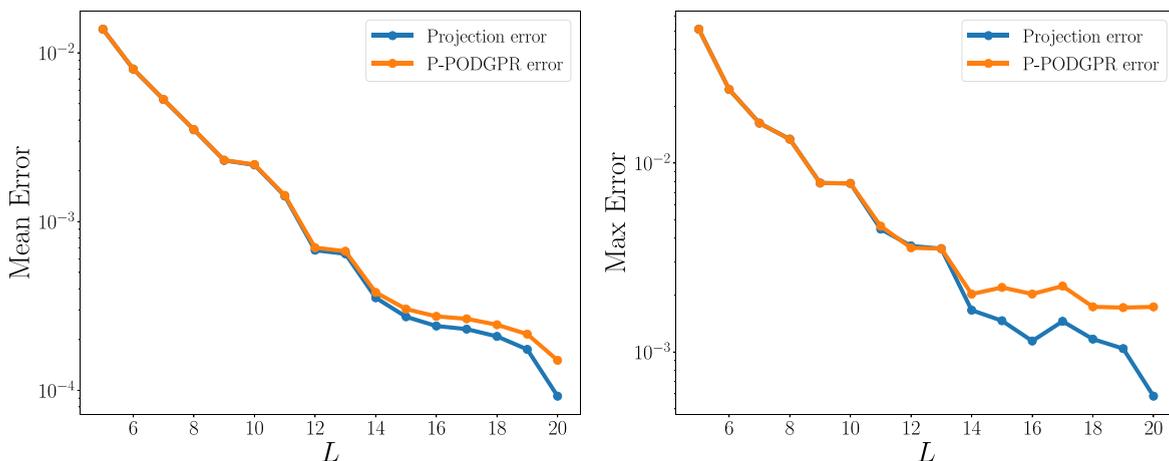
Fig. 7. Porous material: Comparison of stress errors for different L with $N_{pod} = 50$ and $N_{reg} = 500$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
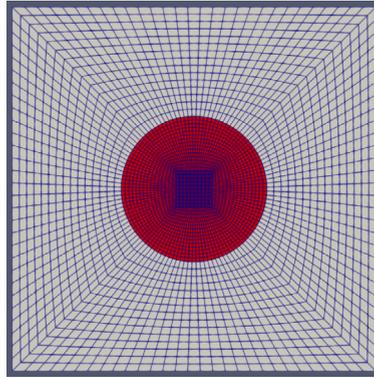
Fig. 8. Fiber reinforced material: The considered microstructure consisted of a Neo-Hookean matrix material (gray) and a Neo-Hookean fiber material (red). The fiber volume fraction was 12.56%. The mesh consisted of 4321 eight-node elements and 13080 nodes.

data was averaged to give effective stresses. Regarding the network architecture, different neural networks with $N_h = \{1, 2\}$ hidden layers with each $N_n = \{20, 50\}$ neurons were tested. The input layer and output layer both had 3 and 4 neurons respectively, one for each component of the stretch tensor and effective stress. Apart from the output layer an *ELU* activation function was applied after each layer. Before training, the deformation and stress data were normalized to [0, 1]. Then, all available $N_{reg} = 500$ training snapshots and all 1000 test snapshots were set as training and validation dataset for the optimization. The network was trained using a mean squared error loss function, the *Adam* optimizer with a learning rate of 1×10^{-4} and a batch size of 32 for 10000 epochs. The validation loss and relative errors $\epsilon_{\bar{P}}$ obtained for each architecture are given in Table 1, where the best architecture is highlighted. All architectures produce similar results with the second architecture $N_h = 1, N_n = 50$ yielding the best results with an average and maximum error of 0.17% and 1.36%. On the other hand, P-PODGPR already yields a lower mean and maximum error of around 0.065% and 0.65% when using only 50 training data, showing that P-PODGPR can utilize the information of each snapshot more efficiently than the neural network. When 500 snapshots are used for the regression in P-PODGPR, a mean and max error of less than 0.02% and 0.2% is achieved, hence outperforming the neural network greatly.

4.1.2. Fiber reinforced material

In the second example, the considered microstructural RVE consists of two different phases: a soft matrix and a stiff fiber material. Both materials are Neo-Hookean, but the matrix has parameters $C_1 = 1, D_1 = 1$, while the fiber has variable parameters $C_1 = D_1 \in [50, 150]$, corresponding to a Young's modulus that is 50–150 times higher than the matrix, with the Poisson ratio remaining the same ($\nu = 0.25$), cf. Eq. (12). The geometry used is depicted in Fig. 8, where the matrix (gray) completely surrounds the fiber (red). Eight node quadrilateral elements with four quadrature points are employed. The volume fraction of the fiber is 12.56%. Linear displacement boundaries with $\bar{U} - \mathbf{I} \in [-0.3, 0.3]$ and fiber parameters of $C_1 = D_1 \in [50, 150]$ are considered. Therefore, the considered problem has 4 varying parameters.

Data generation. A set of 1000 training snapshots and 1000 test snapshots were generated. The first set contained the $2^4 = 16$ corner points of the parameter domain and the remaining points were sampled from a Sobol sequence, while the second set was generated from a random uniform distribution.

Eigenvalues. The eigenvalues of the correlation matrix for different numbers of training snapshots are plotted in Fig. 9. Similar to the last example, for all cases, an exponential decay can be observed, showing the reducibility of the problem. The von Mises stress of the first three POD basis functions are plotted in Fig. 10. Subsequently, $L = 20$ basis functions are considered which correspond to an energy $\mathcal{E}_{pod}$ of 99.9901%.

Influence of N_{pod} and N_{reg}. For $L = 20$ basis functions, combinations of $N_{pod} \in \{20, 50, 100, 200, 500\}$ and $N_{reg} \in \{100, 200, 300, 400, 500, 1000\}$ were tested. The mean and maximum error plots can be seen in Fig. 11. It can be observed that for all cases with $N_{reg} > 50$, the mean error is below 1%. The mean error improves as more
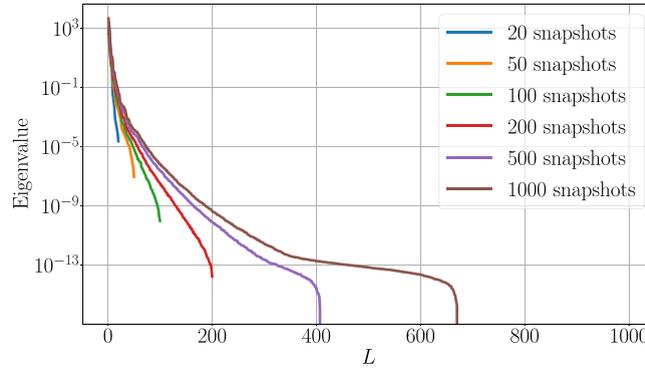
Fig. 9. Fiber reinforced material: Eigenvalues of the correlation matrix for different numbers of snapshots used for POD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(a) 1st basis function   (b) 2nd basis function   (c) 3rd basis function
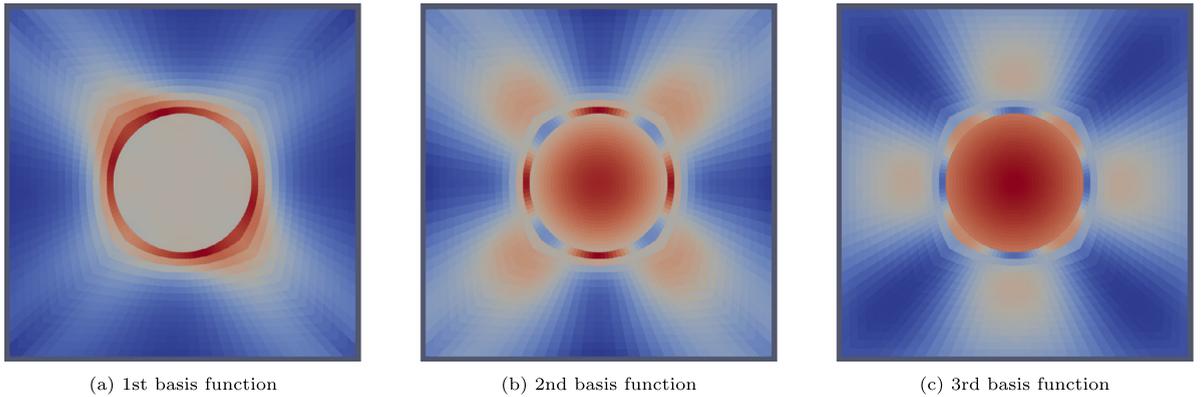
Fig. 10. Fiber reinforced material: Von Mises stress of the first three POD basis functions of the microscopic stress field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

snapshots were considered for the POD basis, suggesting that there was still new information in the snapshots, which fine-tune the optimal basis. The mean error reaches around 0.04% for $N_{\mathrm{reg}} = 500$ snapshots with a maximum error of around 1%. Increasing N_{reg} from 500 to 1000 does not improve the mean error significantly, meaning that the projection error has already been reached with 500 snapshots and $L = 20$ basis functions. Weighing the approximation error over the number of full computations needed, $N_{\mathrm{pod}} = N_{\mathrm{reg}} = 200$ is considered in the subsequent analysis.

Influence of L. Using $N_{\mathrm{pod}} = N_{\mathrm{reg}} = 200$, the influence of L is investigated. The mean and maximum error of the approximation error over the number of basis functions L used are shown in Fig. 12. The projection error is also plotted to reveal the quality of the regression of Eq. (27). As seen from the figure, the mean and maximum error for the first 11 basis functions match the projection error quite well, and afterwards flatten out, similar to the last example. When the number of snapshots N_{reg} is increased to 500, the mean error matches the projection error for 16 basis functions, as seen in Fig. 13, while the maximum error also slightly improves.

Comparison with neural network. Similarly to the last example, deep forward neural networks with different architectures were trained for comparison. The available 500 training data and 1000 testing data were used for training and validation. The same network architecture configurations as before were tested, with one additional combination $N_h = 2, N_n = 100$. The results are given in Table 2. The fourth architecture $N_h = 2, N_n = 50$ performs the best with an average error of 0.39% and a maximum error of 2.06%. Same as for the previous example, P-PODGPR outperforms the neural network. With only 200 training snapshots, P-PODGPR achieves roughly the same accuracy as the NN using 500 snapshots. With 500 employed snapshots, P-PODGPR reaches a mean and
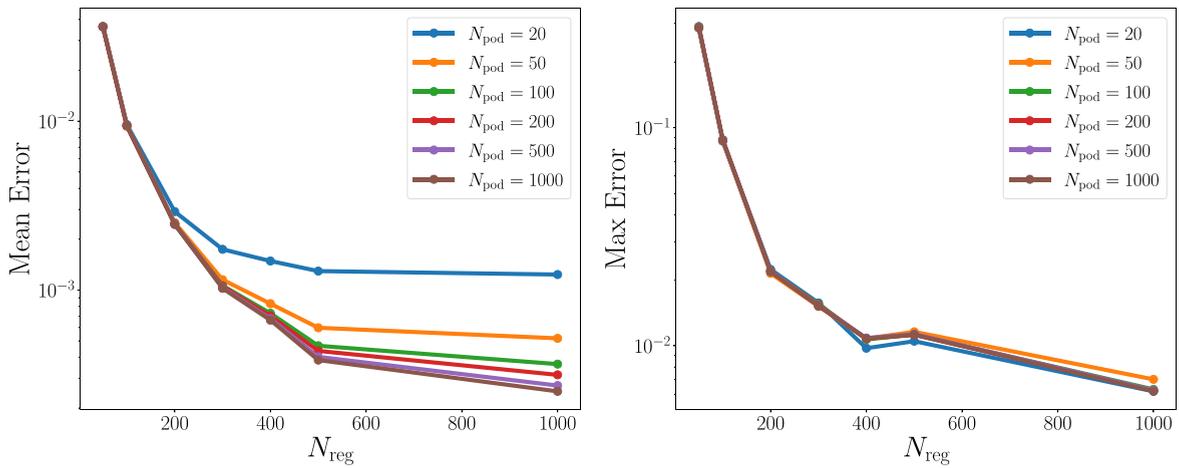
Fig. 11. Fiber reinforced material: Comparison of stress errors for different combinations of N_{pod} and N_{reg} for $L = 20$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
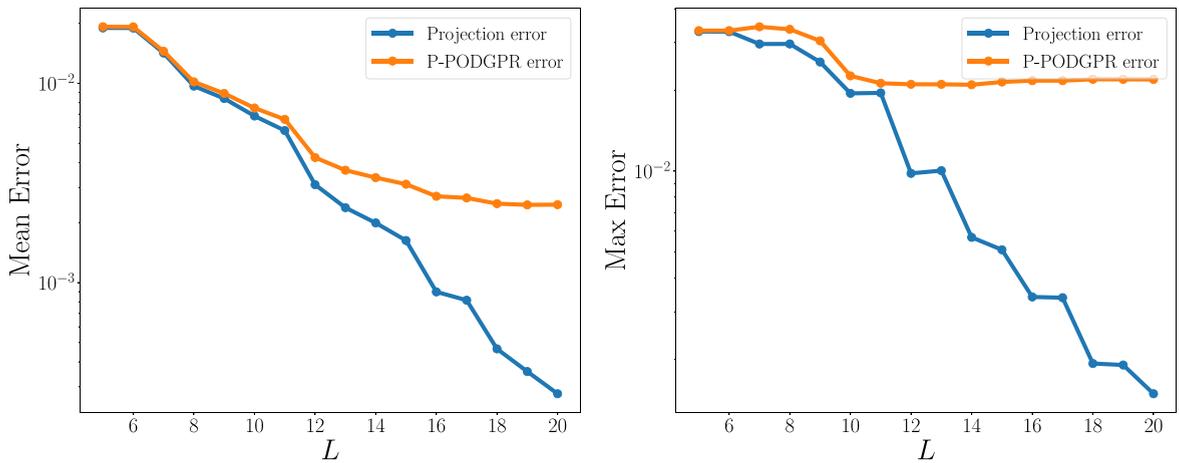
Fig. 12. Fiber reinforced material: Comparison of stress errors for different L with $N_{pod} = 200$ and $N_{reg} = 200$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
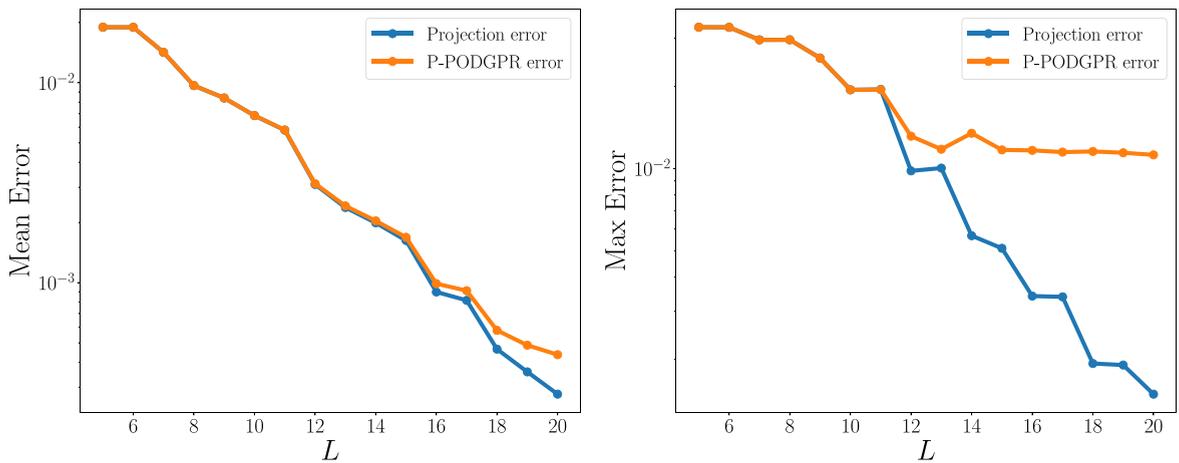
Fig. 13. Fiber reinforced material: Comparison of stress errors for different L with $N_{pod} = 200$ and $N_{reg} = 500$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Fiber reinforced material: Validation loss and effective stress error for different feedforward neural network architectures. The lowest values in each column have been highlighted in bold face.

Architecture	Validation loss	$\epsilon_{\bar{\mathbf{P}}}^{\text{mean}}$	$\epsilon_{\bar{\mathbf{P}}}^{\text{max}}$
$N_h = 1, N_n = 20$	1.7×10^{-6}	0.0077	0.0368
$N_h = 1, N_n = 50$	8.55×10^{-7}	0.0056	0.0289
$N_h = 2, N_n = 20$	5.36×10^{-7}	0.0047	**0.0176**
$N_h = 2, N_n = 50$	$\mathbf{2.97 \times 10^{-7}}$	**0.0039**	0.0206
$N_h = 2, N_n = 100$	7.91×10^{-7}	0.0052	0.0315
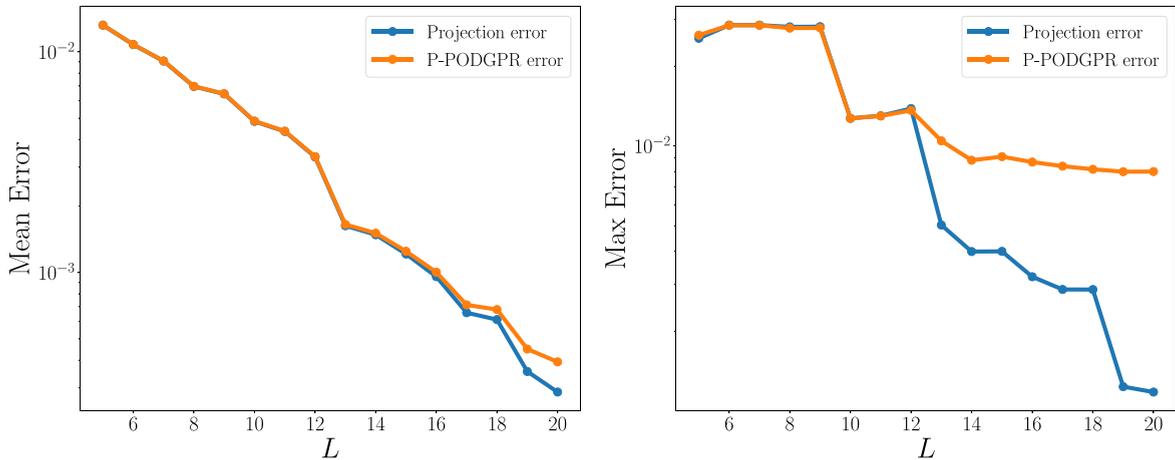
Fig. 14. Fiber reinforced material with periodic boundary conditions: Comparison of stress errors for different L with $N_{\text{pod}} = 200$ and $N_{\text{reg}} = 500$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

maximum error of 0.04% and 1%, hence outperforming the neural network, while also being able to recover the microscopic stress field.

Remark on periodic boundary conditions. For the same microstructure (Fig. 8), an analogous analysis using periodic boundary conditions was considered. In contrast to linear boundary conditions, the fluctuation displacement field $\boldsymbol{w}$ is assumed to be periodic on the RVE boundary. The obtained results regarding the approximation quality are shown in Fig. 14, where $N_{\text{pod}} = 200$ and $N_{\text{reg}} = 500$ was chosen. In comparison to Fig. 13, the results are comparable and hence it can be concluded that P-PODGPR works independently of the chosen boundary conditions.

4.2. Two-scale simulation

To show the performance of P-PODGPR in a two-scale simulation, the learned constitutive model for the fiber reinforced RVE with $N_{\text{pod}} = N_{\text{reg}} = 200$ is embedded inside a FE solver. For the macroscopic problem the Cook's membrane, consisting of the fiber reinforced RVE, is chosen. The geometry and mesh of the membrane are given in Fig. 15. The macroscopic mesh consists of 200 bilinear quadrilateral elements with 4 quadrature points. The quadrature points A and B, as marked on the figure, denote points in which the microscopic stress field is compared against the reference solution. The reference solution is obtained by running a full two-scale FE^2 simulation. The left side of the sample is fixed and the right side is loaded in five time steps with a vertical traction of 0.1. The material parameters of the fiber are taken to be fixed with $C_1 = D_1 = 100$ in this example.

The yx-component of the macroscopic stress $\bar{P}_{yx}$ obtained by the full FE^2 and FE with P-PODGPR simulation are given in Fig. 16, while the microscopic stress P_{yx} at point A and B are shown in Figs. 17 and 18 respectively. The relative error defined as $\epsilon_{P_{yx}} := |P_{yx}^{\text{FE2}} - P_{yx}^{\text{ROM}}| / \langle |P_{yx}^{\text{FE2}}| \rangle$, with $\langle |P_{yx}^{\text{FE2}}| \rangle$ the averaged absolute stress, is also shown (Figs. 16(c), 17(c), 18(c)). As can be seen, the shape of the stress field of both solutions is indistinguishable.
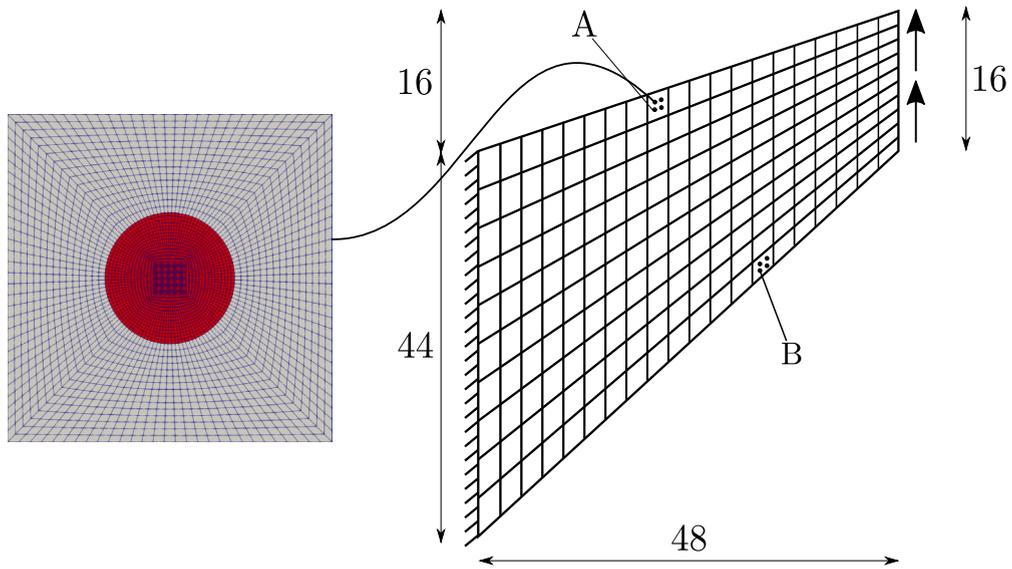
Fig. 15. Cook's membrane: The left side was fixed while a traction in the vertical direction of 0.1 was applied along the right side. The mesh consisted of 200 bilinear elements and 231 nodes. The microstructural stress field at the marked quadrature points A and B will be shown later.

(a) FE^2 (b) FE with P-PODGPR   (c) Relative error
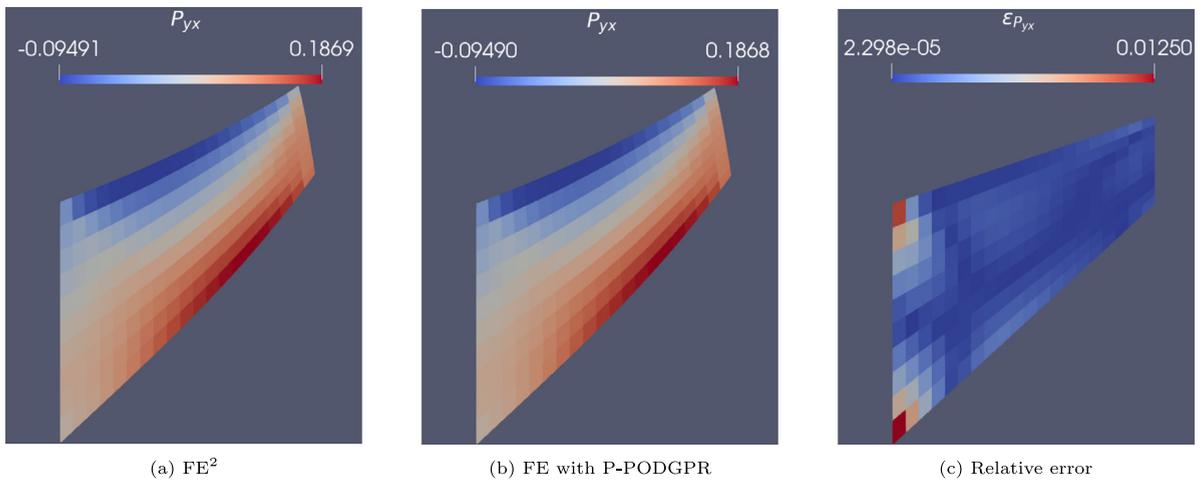
Fig. 16. Cook's membrane: Macroscopic stress field. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Even though the relative errors for the microscopic problem reaches a maximum of 7% near the interface of matrix and fiber, after homogenization the highest error reduces to merely 1%. This discrepancy is due to the fact that the method tries to reduce the L_2-norm of the error in the stress field and therefore allows locally high errors.

Using 48 cores,[1] the FE^2 simulation takes around 100 min while the simulation with P-PODGPR is completed within 1 min on a single core,[2] resulting in a speedup of about three orders of magnitude. For P-PODGPR, 200 RVE simulations have to be pre-computed, which takes less than one hour on a single core. Performing the POD and GPR to construct P-PODGPR takes around one minute.

---

[1] Intel Xeon Platinum 8260.
[2] Intel Core i7-8750H.

(a) FE^2 (b) FE with P-PODGPR   (c) Relative error
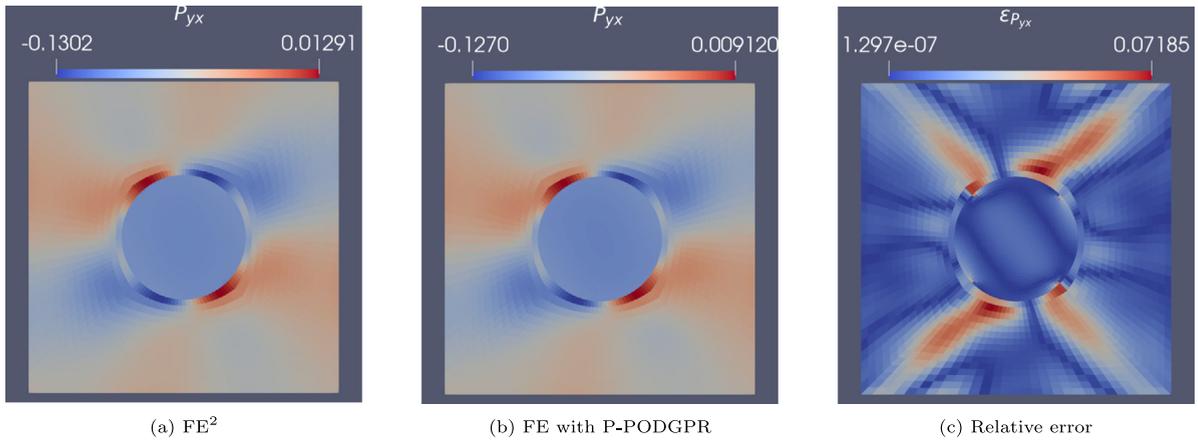
Fig. 17. Cook's membrane: Microscopic stress field at Point A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(a) FE^2 (b) FE with P-PODGPR   (c) Relative error
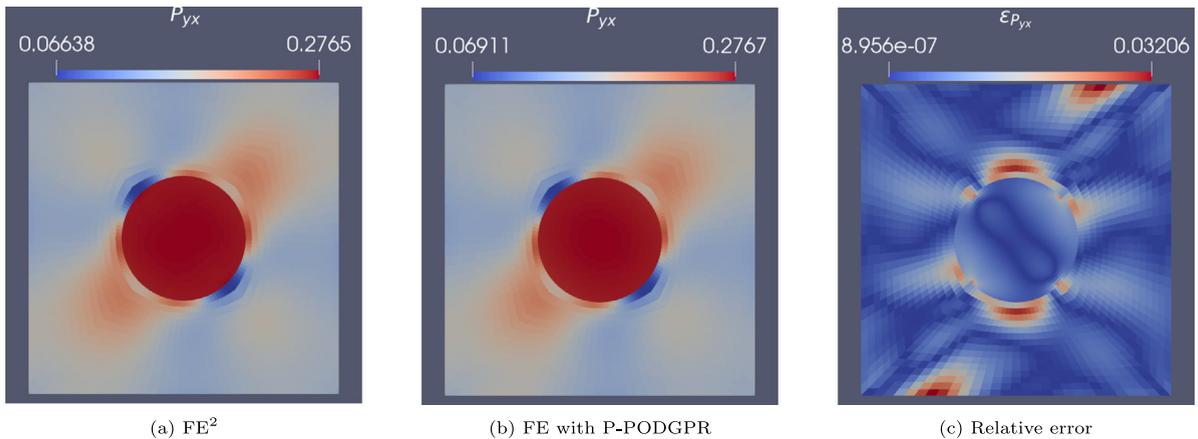
Fig. 18. Cook's membrane: Microscopic stress field at Point B. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

In this work, the proper orthogonal decomposition (POD) has been applied on the microscopic stress field to find a reduced basis. A direct mapping from the loading and material parameters to the POD basis coefficients was built with Gaussian process regression (GPR). For the two considered microstructures, involving a porous and a fiber-reinforced material, the proposed method captured the stress field accurately with a mean error of 0.1%, even in the case of varying material parameters. Finally, the learned constitutive model was employed inside a macroscopic simulation and compared to a full scale FE^2 simulation, reaching high accuracy on both macro- and microscale, while gaining a speedup of the order of 10^3.

Although the examples presented here are two dimensional, the theory has been derived for three dimensions and can readily be applied to 3D problems. Due to the non-intrusive nature of the method, it can be easily implemented into any existing finite element solver. It is noted that the framework is not limited to POD and GPR, but in fact any method for discovering the reduced basis and any regression method can be employed.

This novel method has the potential to open up new ways of material design, as the construction of the surrogate model only requires a relatively small dataset and the surrogate can cover a large range of microstructural parameters. Since the derivatives with respect to the parameters are at hand, it is furthermore possible to define and solve macroscopic optimization problems. Moreover, the microscopic stress field can be fully recovered and visualized in

order to make informed modifications of the microstructure to relieve local stress concentrations and remove flaws. Currently, works on more complicated microstructures are being conducted.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

Acknowledgments

References

[1] M.G.D. Geers, V.G. Kouznetsova, W.A.M. Brekelmans, Multi-scale computational homogenization: Trends and challenges, J. Comput. Appl. Math. 234 (7) (2010) 2175–2182, http://dx.doi.org/10.1016/j.cam.2009.08.077, URL https://linkinghub.elsevier.com/retrieve/pii/S0377042709005536.

[2] M.G.D. Geers, J. Yvonnet, Multiscale modeling of microstructure–property relations, MRS Bull. 41 (08) (2016) 610–616, http://dx.doi.org/10.1557/mrs.2016.165, URL http://link.springer.com/10.1557/mrs.2016.165.

[3] K. Matouš, M.G.D. Geers, V.G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, J. Comput. Phys. (2017) http://dx.doi.org/10.1016/j.jcp.2016.10.070.

[4] F. Feyel, Multiscale FE2 elastoviscoplastic analysis of composite structures, Comput. Mater. Sci. 16 (1–4) (1999) 344–354, http://dx.doi.org/10.1016/s0927-0256(99)00077-4.

[5] H. Moulinec, P. Suquet, A numerical method for computing the overall response of nonlinear composites with complex microstructure, Comput. Methods Appl. Mech. Engrg. 157 (1–2) (1998) 69–94, http://dx.doi.org/10.1016/S0045-7825(97)00218-1.

[6] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707, http://dx.doi.org/10.1016/j.jcp.2018.10.045.

[7] E. Haghighat, M. Raissi, A. Moure, H. Gomez, R. Juanes, A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics, Comput. Methods Appl. Mech. Engrg. 379 (2021) 113741, http://dx.doi.org/10.1016/j.cma.2021.113741, URL https://linkinghub.elsevier.com/retrieve/pii/S0045782521000773.

[8] E. Weinan, B. Yu, The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems, Commun. Math. Stat. 6 (1) (2018) 1–12, http://dx.doi.org/10.1007/s40304-018-0127-z, URL http://link.springer.com/10.1007/s40304-018-0127-z.

[9] E. Samaniego, C. Anitescu, S. Goswami, V.M. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, T. Rabczuk, An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications, Comput. Methods Appl. Mech. Engrg. 362 (2020) 112790, http://dx.doi.org/10.1016/j.cma.2019.112790.

[10] V.G. Kouznetsova, W.A.M. Brekelmans, F.P.T. Baaijens, Approach to micro-macro modeling of heterogeneous materials, Comput. Mech. 27 (1) (2001) 37–48, http://dx.doi.org/10.1007/s004660000212.

[11] O. Rokoš, M.M. Ameen, R.H.J. Peerlings, M.G.D. Geers, Micromorphic computational homogenization for mechanical metamaterials with patterning fluctuation fields, J. Mech. Phys. Solids 123 (2019) 119–137, http://dx.doi.org/10.1016/j.jmps.2018.08.019.

[12] V.P. Nguyen, O. Lloberas-Valls, M. Stroeven, L.J. Sluys, Computational homogenization for multiscale crack modeling. Implementational and computational aspects, Internat. J. Numer. Methods Engrg. 89 (2) (2012) 192–226, http://dx.doi.org/10.1002/nme.3237.

[13] L. Xia, P. Breitkopf, Recent advances on topology optimization of multiscale nonlinear structures, Arch. Comput. Methods Eng. 24 (2) (2017) 227–249, http://dx.doi.org/10.1007/s11831-016-9170-7.

[14] S. Saeb, P. Steinmann, A. Javili, Aspects of computational homogenization at finite deformations: A unifying review from Reuss' to Voigt's Bound, Appl. Mech. Rev. 68 (5) (2016) http://dx.doi.org/10.1115/1.4034024.

[15] C. Miehe, A. Koch, Computational micro-to-macro transitions of discretized microstructures undergoing small strains, Arch. Appl. Mech. 72 (4–5) (2002) 300–317, http://dx.doi.org/10.1007/s00419-002-0212-2.

[16] 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, Comput. Methods Appl. Mech. Engrg. 192 (5–6) (2003) 559–591, http://dx.doi.org/10.1016/S0045-7825(02)00564-9.

[17] I. Temizer, P. Wriggers, On the computation of the macroscopic tangent for multiscale volumetric homogenization problems, Comput. Methods Appl. Mech. Engrg. 198 (3–4) (2008) 495–510, http://dx.doi.org/10.1016/j.cma.2008.08.018.

[18] J.I. Okada, T. Washio, T. Hisada, Study of efficient homogenization algorithms for nonlinear problems: Approximation of a homogenized tangent stiffness to reduce computational cost, Comput. Mech. 46 (2) (2010) 247–258, http://dx.doi.org/10.1007/s00466-009-0432-1.

[19] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, Comput. Methods Appl. Mech. Engrg. 304 (2016) 81–101, http://dx.doi.org/10.1016/j.cma.2016.02.001.

[20] T. Kirchdoerfer, M. Ortiz, Data driven computing with noisy material data sets, Comput. Methods Appl. Mech. Engrg. 326 (2017) 622–641, http://dx.doi.org/10.1016/j.cma.2017.07.039.

[21] R. Eggersmann, T. Kirchdoerfer, S. Reese, L. Stainier, M. Ortiz, Model-free data-driven inelasticity, Comput. Methods Appl. Mech. Engrg. 350 (2019) 81–99, http://dx.doi.org/10.1016/j.cma.2019.02.016.

[22] K. Karapiperis, L. Stainier, M. Ortiz, J.E. Andrade, Data-driven multiscale modeling in mechanics, J. Mech. Phys. Solids 147 (July 2020) (2021) 104239, http://dx.doi.org/10.1016/j.jmps.2020.104239.

[23] D.Z. Huang, K. Xu, C. Farhat, E. Darve, Learning constitutive relations from indirect observations using deep neural networks, J. Comput. Phys. 416 (2020) 109491, http://dx.doi.org/10.1016/j.jcp.2020.109491, URL https://linkinghub.elsevier.com/retrieve/pii/S0021999120302655.

[24] K. Xu, D.Z. Huang, E. Darve, Learning constitutive relations using symmetric positive definite neural networks, J. Comput. Phys. 428 (2021) 110072, http://dx.doi.org/10.1016/j.jcp.2020.110072, URL https://linkinghub.elsevier.com/retrieve/pii/S0021999120308469.

[25] J. Ghaboussi, J. Garrett Jr., X. Wu, Knowledge-based modeling of material behavior with neural networks, J. Eng. Mech. 117 (1991) 132–153.

[26] J. Ghaboussi, D.A. Pecknold, M. Zhang, R.M. Haj-Ali, Autoprogressive training of neural network constitutive models, Internat. J. Numer. Methods Engrg. 42 (1) (1998) 105–126, http://dx.doi.org/10.1002/(SICI)1097-0207(19980515)42:1<105::AID-NME356>3.0.CO;2-V.

[27] F. Ghavamian, A. Simone, Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network, Comput. Methods Appl. Mech. Engrg. 357 (2019) 112594, http://dx.doi.org/10.1016/j.cma.2019.112594.

[28] 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, Comput. Methods Appl. Mech. Engrg. 369 (2020) 113234, http://dx.doi.org/10.1016/j.cma.2020.113234.

[29] H.J. Logarzo, G. Capuano, J.J. Rimoli, Smart constitutive laws: Inelastic homogenization through machine learning, Comput. Methods Appl. Mech. Engrg. 373 (2021) 113482, http://dx.doi.org/10.1016/j.cma.2020.113482.

[30] M. Mozaffar, R. Bostanabad, W. Chen, K. Ehmann, J. Cao, M.A. Bessa, Deep learning predicts path-dependent plasticity, Proc. Natl. Acad. Sci. 116 (52) (2019) 26414–26420, http://dx.doi.org/10.1073/pnas.1911815116.

[31] K. Linka, M. Hillgärtner, K.P. Abdolazizi, R.C. Aydin, M. Itskov, C.J. Cyron, Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning, J. Comput. Phys. 429 (2021) 110010, http://dx.doi.org/10.1016/j.jcp.2020.110010.

[32] F. Masi, I. Stefanou, P. Vannucci, V. Maffi-Berthier, Thermodynamics-based artificial neural networks for constitutive modeling, J. Mech. Phys. Solids 147 (October 2020) (2021) 104277, http://dx.doi.org/10.1016/j.jmps.2020.104277, URL https://linkinghub.elsevier.com/retrieve/pii/S0022509620304841.

[33] B.A. Le, J. Yvonnet, Q.C. He, Computational homogenization of nonlinear elastic materials using neural networks, Internat. J. Numer. Methods Engrg. 104 (12) (2015) 1061–1084, http://dx.doi.org/10.1002/nme.4953.

[34] W. Noll, B.D. Coleman, W. Noll, The thermodynamics of elastic materials with heat conduction and viscosity, The Foundations of Mechanics and Thermodynamics, 1974, pp. 145–156, http://dx.doi.org/10.1007/978-3-642-65817-4{_}9.

[35] G.J. Dvorak, Transformation field analysis of inelastic composite materials, Proc. R. Soc. A 437 (1900) (1992) 311–327, http://dx.doi.org/10.1098/rspa.1992.0063.

[36] J.C. Michel, P. Suquet, Nonuniform transformation field analysis, Int. J. Solids Struct. 40 (25) (2003) 6937–6955, http://dx.doi.org/10.1016/S0020-7683(03)00346-9, URL https://linkinghub.elsevier.com/retrieve/pii/S0020768303003469.

[37] F. Fritzen, M. Leuschner, Reduced basis hybrid computational homogenization based on a mixed incremental formulation, Comput. Methods Appl. Mech. Engrg. 260 (2013) 143–154, http://dx.doi.org/10.1016/j.cma.2013.03.007.

[38] F. Fritzen, S. Marfia, V. Sepe, Reduced order modeling in nonlinear homogenization: A comparative study, Comput. Struct. 157 (2015) 114–131, http://dx.doi.org/10.1016/j.compstruc.2015.05.012.

[39] Z. Liu, M.A. Bessa, W.K. Liu, Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials, Comput. Methods Appl. Mech. Engrg. 306 (2016) 319–341, http://dx.doi.org/10.1016/j.cma.2016.04.004.

[40] C. Yu, O.L. Kafka, W.K. Liu, Self-consistent clustering analysis for multiscale modeling at finite strains, Comput. Methods Appl. Mech. Engrg. 349 (2019) 339–359, http://dx.doi.org/10.1016/j.cma.2019.02.027.

[41] S. Wulfinghoff, F. Cavaliere, S. Reese, Model order reduction of nonlinear homogenization problems using a Hashin–Shtrikman type finite element method, Comput. Methods Appl. Mech. Engrg. 330 (2018) 149–179, http://dx.doi.org/10.1016/j.cma.2017.10.019.

[42] F. Chinesta, R. Keunings, A. Leygue, The Proper Generalized Decomposition for Advanced Numerical Simulations, in: SpringerBriefs in Applied Sciences and Technology, Springer International Publishing, Cham, 2014, pp. 25–46, http://dx.doi.org/10.1007/978-3-319-02865-1.

[43] P. Ladevèze, J.C. Passieux, D. Néron, The LATIN multiscale computational method and the Proper Generalized Decomposition, Comput. Methods Appl. Mech. Engrg. 199 (21–22) (2010) 1287–1296, http://dx.doi.org/10.1016/j.cma.2009.06.023.

[44] S. Niroomandi, D. González, I. Alfaro, F. Bordeu, A. Leygue, E. Cueto, F. Chinesta, Real-time simulation of biological soft tissues: A PGD approach, Int. J. Numer. Methods Biomed. Eng. 29 (5) (2013) 586–600, http://dx.doi.org/10.1002/cnm.2544.

[45] C. Prud'homme, D.V. Rovas, K. Veroy, L. Machiels, Y. Maday, A.T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods, Trans. ASME, J. Fluids Eng. 124 (1) (2002) 70–80, http://dx.doi.org/10.1115/1.1448332.

[46] A. Quarteroni, A. Manzoni, F. Negri, Reduced Basis Methods for Partial Differential Equations, in: UNITEXT, vol. 92, Springer International Publishing, Cham, 2016, pp. 1–263, http://dx.doi.org/10.1007/978-3-319-15431-2.

[47] J.S. Hesthaven, G. Rozza, B. Stamm, Certified Reduced Basis Methods for Parametrized Partial Differential Equations, in: SpringerBriefs in Mathematics, Springer International Publishing, Cham, 2016, pp. 1–131, http://dx.doi.org/10.1007/978-3-319-22470-1.

[48] K.T. Carlberg, C. Bou-Mosleh, C. Farhat, Efficient non-linear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations, Internat. J. Numer. Methods Engrg. 86 (2) (2011) 155–181, http://dx.doi.org/10.1002/nme.3050.

[49] A. Radermacher, S. Reese, A comparison of projection-based model reduction concepts in the context of nonlinear biomechanics, Arch. Appl. Mech. 83 (8) (2013) 1193–1213, http://dx.doi.org/10.1007/s00419-013-0742-9.

[50] D. Soldner, B. Brands, R. Zabihyan, P. Steinmann, J. Mergheim, A numerical study of different projection-based model reduction techniques applied to computational homogenisation, Comput. Mech. 60 (4) (2017) 613–625, http://dx.doi.org/10.1007/s00466-017-1428-x.

[51] F. Ghavamian, P. Tiso, A. Simone, POD–DEIM model order reduction for strain softening viscoplasticity, Comput. Methods Appl. Mech. Engrg. 317 (2017) 458–479, http://dx.doi.org/10.1016/j.cma.2016.11.025.

[52] D. Ryckelynck, Hyper-reduction of mechanical models involving internal variables, Internat. J. Numer. Methods Engrg. 77 (1) (2009) 75–89, http://dx.doi.org/10.1002/nme.2406.

[53] J.A. Hernández, J. Oliver, A.E. Huespe, M.A. Caicedo, J.C. Cante, High-performance model reduction techniques in computational multiscale homogenization, Comput. Methods Appl. Mech. Engrg. 276 (2014) 149–189, http://dx.doi.org/10.1016/j.cma.2014.03.011.

[54] M. Guo, J.S. Hesthaven, Reduced order modeling for nonlinear structural analysis using Gaussian process regression, Comput. Methods Appl. Mech. Engrg. 341 (2018) 807–826, http://dx.doi.org/10.1016/j.cma.2018.07.017.

[55] J.S. Hesthaven, S. Ubbiali, Non-intrusive reduced order modeling of nonlinear problems using neural networks, J. Comput. Phys. 363 (2018) 55–78, http://dx.doi.org/10.1016/j.jcp.2018.02.037.

[56] M. Kast, M. Guo, J.S. Hesthaven, A non-intrusive multifidelity method for the reduced order modeling of nonlinear problems, Comput. Methods Appl. Mech. Engrg. 364 (2020) 112947, http://dx.doi.org/10.1016/j.cma.2020.112947.

[57] R. Swischuk, L. Mainini, B. Peherstorfer, K. Willcox, Projection-based model reduction: Formulations for physics-based machine learning, Comput. & Fluids 179 (2019) 704–717, http://dx.doi.org/10.1016/j.compfluid.2018.07.021.

[58] M. Guénot, I. Lepot, C. Sainvitu, J. Goblet, R.F. Coelho, Adaptive sampling strategies for non-intrusive POD-based surrogates, Eng. Comput. 30 (4) (2013) 521–547, http://dx.doi.org/10.1108/02644401311329352.

[59] M. Yang, Z. Xiao, POD-based surrogate modeling of transitional flows using an adaptive sampling in Gaussian process, Int. J. Heat Fluid Flow 84 (December 2019) (2020) 108596, http://dx.doi.org/10.1016/j.ijheatfluidflow.2020.108596.

[60] M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An 'empirical interpolation' method: Application to efficient reduced-basis discretization of partial differential equations, C. R. Math. 339 (9) (2004) 667–672, http://dx.doi.org/10.1016/j.crma.2004.08.006.

[61] S. Chaturantabut, D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, SIAM J. Sci. Comput. 32 (5) (2010) 2737–2764, http://dx.doi.org/10.1137/090766498.

[62] J. Yvonnet, Q.C. He, The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains, J. Comput. Phys. 223 (1) (2007) 341–368, http://dx.doi.org/10.1016/j.jcp.2006.09.019.

[63] D.W. Abueidda, S. Koric, N.A. Sobh, H. Sehitoglu, Deep learning for plasticity and thermo-viscoplasticity, Int. J. Plast. 136 (August 2020) (2021) 102852, http://dx.doi.org/10.1016/j.ijplas.2020.102852.

[64] R. Chakir, J.K. Hammond, A non-intrusive reduced basis method for elastoplasticity problems in geotechnics, J. Comput. Appl. Math. 337 (2018) 1–17, http://dx.doi.org/10.1016/j.cam.2017.12.044.

[65] O. Kunc, F. Fritzen, Finite strain homogenization using a reduced basis and efficient sampling, Math. Comput. Appl. 24 (2) (2019) 56, http://dx.doi.org/10.3390/mca24020056.

[66] A. Bērziņš, J. Helmig, F. Key, S. Elgeti, Standardized non-intrusive reduced order modeling using different regression models with application to complex flow problems, arXiv, 2020.

[67] H. Liu, Y.S. Ong, J. Cai, A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design, Struct. Multidiscip. Optim. 57 (1) (2018) 393–416, http://dx.doi.org/10.1007/s00158-017-1739-8.

[68] C.E. Rasmussen, Gaussian Processes in machine learning, in: O. Bousquet, U. von Luxburg, G. Rätsch (Eds.), Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, pp. 63–71, http://dx.doi.org/10.1007/978-3-540-28650-9{_}4.

[69] D. Huang, J.N. Fuhg, C. Weißenfels, P. Wriggers, A machine learning based plasticity model using proper orthogonal decomposition, Comput. Methods Appl. Mech. Engrg. 365 (2020) 113008, http://dx.doi.org/10.1016/j.cma.2020.113008.

[70] P. Wu, J. Sun, X. Chang, W. Zhang, R. Arcucci, Y. Guo, C.C. Pain, Data-driven reduced order model with temporal convolutional neural network, Comput. Methods Appl. Mech. Engrg. 360 (2020) 112766, http://dx.doi.org/10.1016/j.cma.2019.112766.

[71] C.J. Permann, D.R. Gaston, D. Andrš, R.W. Carlsen, F. Kong, A.D. Lindsay, J.M. Miller, J.W. Peterson, A.E. Slaughter, R.H. Stogner, R.C. Martineau, MOOSE: Enabling massively parallel multiphysics simulation, SoftwareX 11 (2020) 100430, http://dx.doi.org/10.1016/j.softx.2020.100430.

[72] M.A. Bessa, R. Bostanabad, Z. Liu, A. Hu, D.W. Apley, C. Brinson, W. Chen, W.K. Liu, A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality, Comput. Methods Appl. Mech. Engrg. 320 (2017) 633–667, http://dx.doi.org/10.1016/j.cma.2017.03.037.