Self-consistent clustering analysis: an efficient multi-scale scheme for inelastic heterogeneous materials

Zeliang Liu[a], M.A. Bessa[b], Wing Kam Liu[b,*]

[a] *Theoretical and Applied Mechanics, Northwestern University, Evanston, IL 60208, USA*
[b] *Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA*

Abstract

The discovery of efficient and accurate descriptions for the macroscopic behavior of materials with complex microstructure is an outstanding challenge in mechanics of materials. A mechanistic, data-driven, two-scale approach is developed for predicting the behavior of general heterogeneous materials under irreversible processes such as inelastic deformation. The proposed approach includes two major innovations: 1) the use of a data compression algorithm, k-means clustering, during the offline stage of the method to homogenize the local features of the material microstructure into a group of clusters; and 2) a new method called self-consistent clustering analysis used in the online stage that is valid for any local plasticity laws of each material phase without the need for additional calibration. A particularly important feature of the proposed approach is that the offline stage only uses the linear elastic properties of each material phase, making it efficient. This work is believed to open new avenues in parameter-free multi-scale modeling of complex materials, and perhaps in other fields that require homogenization of irreversible processes.

Keywords: multi-scale, reduced order model, plasticity, data compression, k-means clustering, self-consistent method

---

1. Introduction

Materials are hierarchical in nature, involving an inter-play between simple small-scale constituents that together form elaborate compounds that can span multiple time- and length-scales. This multi-scale nature of heterogeneous materials poses a continuing challenge in computational modeling of macroscopic structures. Ideally, efficient and accurate predictions of the macroscopic behavior of heterogeneous materials should be uniquely obtained from the constitutive behavior of each separate constituent (material phase) and from the information about the material microstructure.

Traditional phenomenological constitutive relations [1, 2, 3] characterize the average behaviors of the material, i.e the contributions from all the material phases are not accounted for as an individual interaction of separate constituents. These laws regard materials as “black boxes,” implying the need for burdensome experimental characterization and tedious calibration. In addition, they are problem-dependent and tend to fail when capturing highly localized microstructure-induced nonlinear material behaviors, such as plasticity, damage and fatigue.

Concurrent multiscale methods [4, 5, 6, 7, 8, 9, 10, 11] avoid this calibration process by directly establishing the connection between the microstructure and the macro-response of materials. These concurrent methods link to *every* macroscopic point of a structure a high-fidelity simulation of a Representative Unit Cell (RUC) [12, 13] of the microstructure. Since each RUC is already associated to a large computational cost, the total computational cost of this approach is tremendous.

---

[*] Corresponding author
Email address: `w-liu@northwestern.edu` (Wing Kam Liu)

A myriad of methods have been developed with the goal of finding an appropriate balance between cost and accuracy; these are generally referred as reduced order models. Analytical micromechanical methods [14, 15, 16, 17], the Voronoi cell finite element method [18], fast Fourier transforms [19], spectral methods [20], the generalized method of cells [21], the transformation field analysis (TFA) [22], the nonuniform transformation field analysis (NTFA) [23, 24], the principal component analysis [25, 26, 27] or proper orthogonal decomposition[1] (POD) [28, 29, 30, 31], and the proper generalized decomposition (PGD) [32, 33, 34, 35, 36] are some of the most successful methods of this kind.

Analytical micromechanical methods [14, 15, 16, 17] are very efficient because they describe the heterogeneous material by several microstructural descriptors, rather than considering the whole representative unit cell explicitly. However, they are based on mean-field assumptions and linear superposition. This limits their applicability when complex microstucture and localized nonlinear material behavior such as plasticity are present.

POD uses linear combinations of all input variables to define the principal components (modes) to make the predictions for general load cases. This is a fundamental issue when describing nonlinear irreversible plastic processes that is only mitigated by using many snapshots in the offline stage which increases the computational cost of the online stage to unreasonable values [29, 37].

The NTFA introduced by Michel and Suquet [23] and further explored by Oskay and Fish [38] among others [39, 40], approaches the problem differently with the aim of drastically reducing computational cost [39]. The idea behind NTFA is to leverage the capabilities of analytical micromechanical methods to define a set of reduced variables (e.g. plastic strain fields and internal variables of the local constitutive laws) that are then subjected to "evolution equations" [23, 41]. Consequently, the computational cost of NTFA is low but the inclusion of empirical laws that require further calibration is a limitation. Furthermore, the representative unit cells need to be subjected to irreversible deformation to obtain the plastic modes (as in POD) which leads to an extensive exploration of the deformation space at the offline stage.

In this article a new approach is proposed where the above-mentioned limitations are addressed simultaneously by meeting three fundamental goals: 1) avoiding to have an extensive exploratory offline stage (as in NTFA and especially POD), by limiting this stage to the characterization of the *elastic behavior* of the representative unit cell; 2) eliminating the need to find reduced macro-constitutive equations (as in NTFA), i.e. only using the local constitutive equations of each phase present in the representative unit cell without calibration of additional laws; and 3) achieving a reduction in computational time of several orders of magnitude without significantly compromising accuracy.

In order to achieve these goals *two fundamental contributions* are introduced in Section 2: a new analysis method applicable to any reduced RUC; and a procedure based on a data compression algorithm to obtain any reduced RUC. Section 3 shows the results and discusses the accuracy and efficiency of the methodology. Concluding remarks are provided in Section 4.

2. Methodology

The starting point is a high-fidelity Representative Unit Cell (RUC) of the material, i.e. a representative domain of the microstructure of a material. This high-fidelity RUC can be analyzed by different methods, e.g. finite element or meshfree methods, such that accurate predictions of the material's mechanical behavior are achieved. However, the computational cost associated to these RUCs is usually significantly large because they involve a very fine numerical discretization to capture the microstructural shapes of the material and its detailed mechanical behavior.

With the aim of lowering the computational expense one may consider to decompose the high-fidelity RUC into a group of large subdomains, obtaining a reduced RUC – Figure 1. These subdomains are designated as **material clusters** henceforth. However, if one uses the same analysis method considered for the high-fidelity RUC, the resolution for the analysis is degraded and the accuracy of the predictions decreases. For example, if using the finite element method the larger subdomains (material clusters) would

---

[1] The mechanics community refers to principal component analysis as proper orthogonal decomposition (POD).

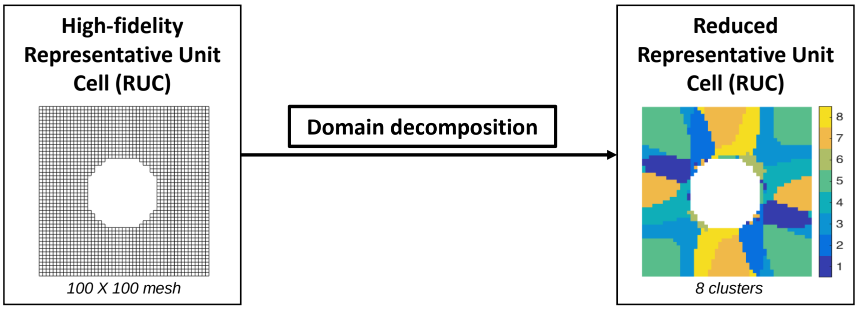


Figure 1: Illustration of arbitrary domain decomposition for a high-fidelity RUC transforming it into a reduced RUC with just 8 subdomains (clusters).

have to correspond to larger finite elements, introducing limitations in the shape of these clusters and the accuracy of the solution.

> This poses an interesting question: for large and irregular material clusters, what is the appropriate analysis method that maintains good accuracy for the predictions of the reduced RUC?

In order to answer this question, assume an arbitrary domain decomposition of the RUC. The procedure used to determine these material clusters will be introduced later in Section 2.2.1. In addition, assume that *every local variable $\boldsymbol{\beta}(\boldsymbol{x})$ within each material cluster is uniform.* This is equivalent to considering a piecewise uniform approximation to the local variables of interest in the RUC[2]:

$$\boldsymbol{\beta}(\mathbf{x}) = \sum_{I=1}^{k} \boldsymbol{\beta}^I \chi^I(\mathbf{x}), \tag{1}$$

where $\boldsymbol{\beta}^I$ is the homogeneous variable in the I-th material cluster and $\chi^I(\mathbf{x})$ is the characteristic function in the domain of the I-th material cluster Ω^I, defined as

$$\chi^I(\mathbf{x}) = \begin{cases} 1 & \text{if } x \in \Omega^I \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

from which the following cluster averaging relationship can be written,

$$\int_{\Omega} \chi^I(\mathbf{x})[\bullet] d\mathbf{x} \equiv \int_{\Omega^I} [\bullet] d\mathbf{x} \tag{3}$$

where $[\bullet]$ signifies any quantity of interest to be averaged in the cluster domain Ω^I.

Under these assumptions, a system of equations can be derived by successively homogenizing each material cluster via the Lippmann-Schwinger equation. The procedure is applicable to any domain decomposition considered and is presented next.

2.1. Lippmann-Schwinger homogenization for multiple clusters

Generally, the equilibrium condition without body force in a material domain Ω can be written as

$$\frac{\partial \sigma_{ij}(\mathbf{x})}{\partial x_i} = 0 \quad \text{in } \Omega. \tag{4}$$

---

[2] For example, if the variable of interest is local strain: $\boldsymbol{\beta}(\mathbf{x}) \equiv \boldsymbol{\varepsilon}(\mathbf{x})$

By introducing a homogeneous isotropic linear elastic reference material with stiffness $\mathbf{C}^0$, the real stress in the heterogeneous material can be divided into two parts,

$$\boldsymbol{\sigma}(\mathbf{x}) = \mathbf{C}^0 : \boldsymbol{\varepsilon}(\mathbf{x}) + \mathbf{p}(\mathbf{x}), \tag{5}$$

where $\mathbf{p}(\mathbf{x})$ is the so-called polarization stress, which denotes the difference between the real stress and the stress in the reference material under the same strain. By substituting equation (5) into (4), we obtain

$$C^0_{ijkl}\frac{\partial \varepsilon_{kl}(\mathbf{x})}{\partial x_i} = -\frac{\partial p_{ij}(\mathbf{x})}{\partial x_i} \quad \text{in } \Omega. \tag{6}$$

With the help of the Green's function $\mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}')$ which represents the strain at $\mathbf{x}$ contributed by a concentrated external stress at $\mathbf{x}'$ in a homogeneous reference material, the original equilibrium condition can be rewritten in an integral form if we regard the polarization stress as an external stress on the reference material [15, 42],

$$\boldsymbol{\varepsilon}(\mathbf{x}) + \int_\Omega \mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}') : \mathbf{p}(\mathbf{x}')d\mathbf{x}' - \boldsymbol{\varepsilon}^0 = 0, \tag{7}$$

where $\boldsymbol{\varepsilon}^0$ is the far-field strain, which is homogeneous in the reference material and controls the evolution of the strain $\boldsymbol{\varepsilon}(\mathbf{x})$. By substituting equation (5) into (7), we obtain the integral equation in terms of the local strain $\boldsymbol{\varepsilon}(\mathbf{x})$,

$$\boldsymbol{\varepsilon}(\mathbf{x}) + \int_\Omega \mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}') : \left[\boldsymbol{\sigma}(\mathbf{x}') - \mathbf{C}^0 : \boldsymbol{\varepsilon}(\mathbf{x}')\right] d\mathbf{x}' - \boldsymbol{\varepsilon}^0 = 0. \tag{8}$$

In order to solve $\boldsymbol{\varepsilon}(\mathbf{x})$ in the integral equation (8), constraints are needed from the macroscopic boundary conditions. In specific, if a macro-strain constraint is used:

$$\frac{1}{\mid \Omega \mid}\int_\Omega \boldsymbol{\varepsilon}(\mathbf{x})d\mathbf{x} = \bar{\boldsymbol{\varepsilon}}, \tag{9}$$

or if a macro-stress constraint is considered:

$$\frac{1}{\mid \Omega \mid}\int_\Omega \boldsymbol{\sigma}(\mathbf{x})d\mathbf{x} = \bar{\boldsymbol{\sigma}}. \tag{10}$$

Similarly, we can also define mixed constraints. For example, the constraints under uniaxial tension loading are

$$\frac{1}{\mid \Omega \mid}\int_\Omega \varepsilon_{11}(\mathbf{x})d\mathbf{x} = \bar{\varepsilon}_{11} \quad \text{if } i=j=1;\ \text{otherwise },\frac{1}{\mid \Omega \mid}\int_\Omega \sigma_{ij}(\mathbf{x})d\mathbf{x} = \bar{\sigma}_{ij} = 0. \tag{11}$$

For convenience, equation (8) can be written in incremental form:

$$\Delta\boldsymbol{\varepsilon}(\mathbf{x}) + \int_\Omega \mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}') : \left[\Delta\boldsymbol{\sigma}(\mathbf{x}') - \mathbf{C}^0 : \Delta\boldsymbol{\varepsilon}(\mathbf{x}')\right] d\mathbf{x}' - \Delta\boldsymbol{\varepsilon}^0 = 0, \tag{12}$$

Integral equation (8) or (12) are also known as Lippmann-Schwinger equation, and solving such an integral equation for every point would be very time-consuming and even slower than the actual simulation of the high-fidelity RUC using finite element analysis, for example. However, recall that the high-fidelity RUC was decomposed into finite material clusters to form a reduced RUC (see Figure 2). With this in mind, the Lippmann-Schwinger equation is averaged for **each** cluster:

$$\frac{1}{c^I \mid \Omega \mid}\int_\Omega \chi^I(\mathbf{x})\Delta\boldsymbol{\varepsilon}(\mathbf{x})d\mathbf{x} + \frac{1}{c^I \mid \Omega \mid}\int_\Omega\int_\Omega \chi^I(\mathbf{x})\mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}') : \left[\Delta\boldsymbol{\sigma}(\mathbf{x}') - \mathbf{C}^0 : \Delta\boldsymbol{\varepsilon}(\mathbf{x}')\right] d\mathbf{x}'d\mathbf{x} - \Delta\boldsymbol{\varepsilon}^0 = 0, \tag{13}$$

where this equation follows directly from averaging equation (12) within the I-th cluster, and where the first term corresponds to $\Delta\boldsymbol{\varepsilon}^I$. Moreover, recalling the piecewise uniform assumption of all the local variables, equation (1),

$$\Delta\boldsymbol{\varepsilon}(\mathbf{x}) = \sum_{J=1}^{k}\chi^J(\mathbf{x})\Delta\boldsymbol{\varepsilon}^J, \quad \Delta\boldsymbol{\sigma}(\mathbf{x}) = \sum_{J=1}^{k}\chi^J(\mathbf{x})\Delta\boldsymbol{\sigma}^J \tag{14}$$

where $\Delta\boldsymbol{\varepsilon}^J$ and $\Delta\boldsymbol{\sigma}^J$ are the incremental stain and stress in the J-th cluster. Hence, equation (13) can be simplified according to these relations as,

$$\Delta\boldsymbol{\varepsilon}^I + \sum_{J=1}^{k} \left[\frac{1}{c^I \mid \Omega \mid} \int_{\Omega} \int_{\Omega} \chi^I(\mathbf{x}) \chi^J(\mathbf{x}') \boldsymbol{\Phi}^0(\mathbf{x}, \mathbf{x}') d\mathbf{x}' d\mathbf{x} \right] : \left[\Delta\boldsymbol{\sigma}^J - \mathbf{C}^0 : \Delta\boldsymbol{\varepsilon}^J \right] - \Delta\boldsymbol{\varepsilon}^0 = 0. \tag{15}$$

where the term within brackets is a well-defined quantity in micromechanical analysis of materials: the interaction tensor $\mathbf{D}^{IJ}$,

$$\mathbf{D}^{IJ} = \frac{1}{c^I \mid \Omega \mid} \int_{\Omega} \int_{\Omega} \chi^I(\mathbf{x}) \chi^J(\mathbf{x}') \boldsymbol{\Phi}^0(\mathbf{x}, \mathbf{x}') d\mathbf{x}' d\mathbf{x}, \tag{16}$$

written as an integral of Green's function $\boldsymbol{\Phi}^0(\mathbf{x}, \mathbf{x}')$ in a RUC domain Ω with periodic boundary conditions, and where c^I is the volume fraction of the I-th cluster and $\mid \Omega \mid$ is the volume of domain Ω.

Equation (13) can then be re-written considering equation (16) as

$$\Delta\boldsymbol{\varepsilon}^I + \sum_{J=1}^{k} \mathbf{D}^{IJ} : \left[\Delta\boldsymbol{\sigma}^J - \mathbf{C}^0 : \Delta\boldsymbol{\varepsilon}^J \right] - \Delta\boldsymbol{\varepsilon}^0 = 0. \tag{17}$$

This is an integral equation for each I-th cluster, which can be used together with the previously defined macro-strain or macro-stress constraints in incremental form:

$$\sum_{I=1}^{k} c^I \Delta\boldsymbol{\varepsilon}^I = \Delta\bar{\boldsymbol{\varepsilon}} \quad \text{or} \quad \sum_{I=1}^{k} c^I \Delta\boldsymbol{\sigma}^I = \Delta\bar{\boldsymbol{\sigma}}, \tag{18}$$

Remark 1. *The system of equations composed by k equations (17) and the respective constraint equations (18) is the main result of this section. This closed system of equations arises from the averaging of the Lippmann-Schwinger equation for each cluster.*

Note that equations (17) and (18) do not depend on the strain concentration tensor or on any additional parameters or constitutive laws. This contrasts with approaches such as TFA [22] and NTFA [23, 39]. Also note that the separate averaging of the Lippmann-Schwinger equation introduced in this section implies that stresses in adjacent clusters are not enforced to be continuous, but the global average of the local strains and stresses are constrained to coincide with the macroscopic applied strain or stress. In a way, this is similar to what happens in the finite element discretization of the conservation of linear momentum equation where the continuity of stresses is also not enforced between adjacent elements, but global equilibrium is satisfied.

The system of equations (17-18) requires the previous determination of the material clusters (domain decomposition), followed by the computation of the interaction tensors $\mathbf{D}^{IJ}$ between each pair of clusters. This *a priori* analysis is conveniently called the "offline" or training stage. The "online" or predictive stage comprises the actual solution of the system of equations (17-18) for any loading conditions applied to the RUC. Section 2.2 describes the offline stage, where a strategy to find the reduced RUC previously mentioned is proposed. Section 2.3 presents the details of the online stage. Figure 2 shows a schematic of the developed framework summarizing the methodology.

2.2. Offline stage

How can one perform the domain decomposition in the "offline stage"? Different methodologies can be adopted. The goal is to group points that have similar mechanical behavior under any applied loading condition. An optimal strategy to find the material clusters is proposed next by using a data compression algorithm called k-means clustering.

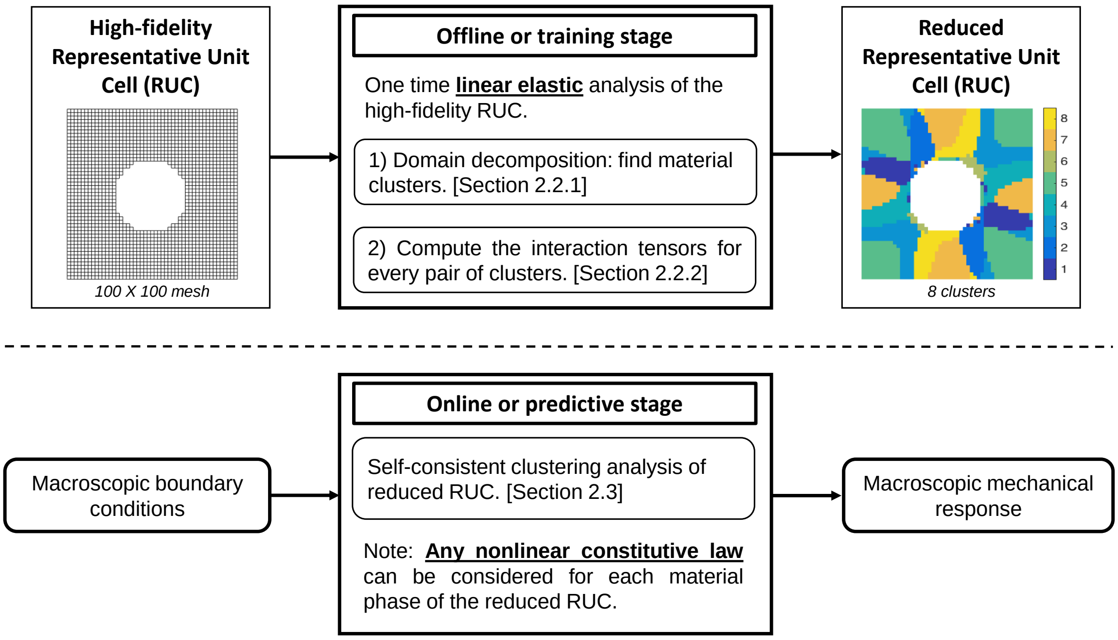


Figure 2: Flowchart summarizing the self-consistent clustering analysis developed in this work. Compared to other reduced order models the offline stage is trivial because it only involves a linear elastic analysis of the representative unit cell (RUC).

2.2.1. Domain decomposition via k-means clustering

A simple yet effective way of characterizing points in the material with similar mechanical behavior is by evaluating the **elastic** response of the high-fidelity RUC and computing at every point the strain concentration tensor $\mathbf{A}(\mathbf{x})$. This tensor relates the elastic microscopic strain $\boldsymbol{\varepsilon}^{\text{micro}}(\mathbf{x})$ to the homogeneous elastic macroscopic strain $\boldsymbol{\varepsilon}^{\text{macro}}$,

$$\boldsymbol{\varepsilon}^{\text{micro}}(\mathbf{x}) = \mathbf{A}(\mathbf{x}) : \boldsymbol{\varepsilon}^{\text{macro}}, \tag{19}$$

where the macroscopic strain $\boldsymbol{\varepsilon}^{\text{macro}}$ is directly related to the periodic boundary conditions of the RUC [12].

By definition, the similarity between two data points is characterized by the difference between their strain concentration tensors. In other words, if two data points have equivalent strain concentration tensors, they have exactly the same mechanical behavior under any loading condition within the elastic regime. Moreover, these points should also have similar nonlinear plastic response, since the localization of plasticity occurs at points with high strain concentrations.

For a 2-dimensional (2D) material, the strain concentration tensor $\mathbf{A}(\mathbf{x})$ in each material point has 9 independent components which are determined by direct numerical simulation (DNS) of the high-fidelity RUC under 3 orthogonal loading conditions. While for a 3-dimensional (3D) material, $\mathbf{A}(\mathbf{x})$ has 36 independent components, needing a set of DNS under 6 orthogonal loading conditions. The format of the raw data for a 2D material is shown below,

DataIndex	A_{11}	A_{22}	A_{33}	A_{12}	A_{21}	A_{23}	A_{32}	A_{13}	A_{31}
1	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	–	–
⋮			⋮			⋮			⋮
N	–	–	–	–	–	–	–	–	–

where the dimension of the data is 9 in the 2D case, and N is the total number of discretization points in the

DNS. For example, if the high-fidelity RUC is discretized by a 600 × 600 finite element mesh with reduced integration elements, then N is equal to 3.6×10^5.

Remark 2. *Understanding the above selection for measuring mechanical similarity is crucial. Other metrics for similarity could be used but that would have direct effects on the reduction of the number of degrees of freedom by forming material clusters. For example, a simpler metric would be to group data points according to their spatial proximity (disregarding their mechanical similarity), as was done later in this article for comparison (Figure 6).*

The next step is to perform the domain decomposition by grouping similar data points using a clustering algorithm called k-means clustering [43]. Note that the data points grouped by the k-means clustering do not need to be adjacent to each other. These points belong to the same cluster because their strain concentration tensor is approximately the same, but they can be in disconnected parts of the RUC – see in Figure 3 irregular and disconnected clusters represented by the same color for two different material microstructures.

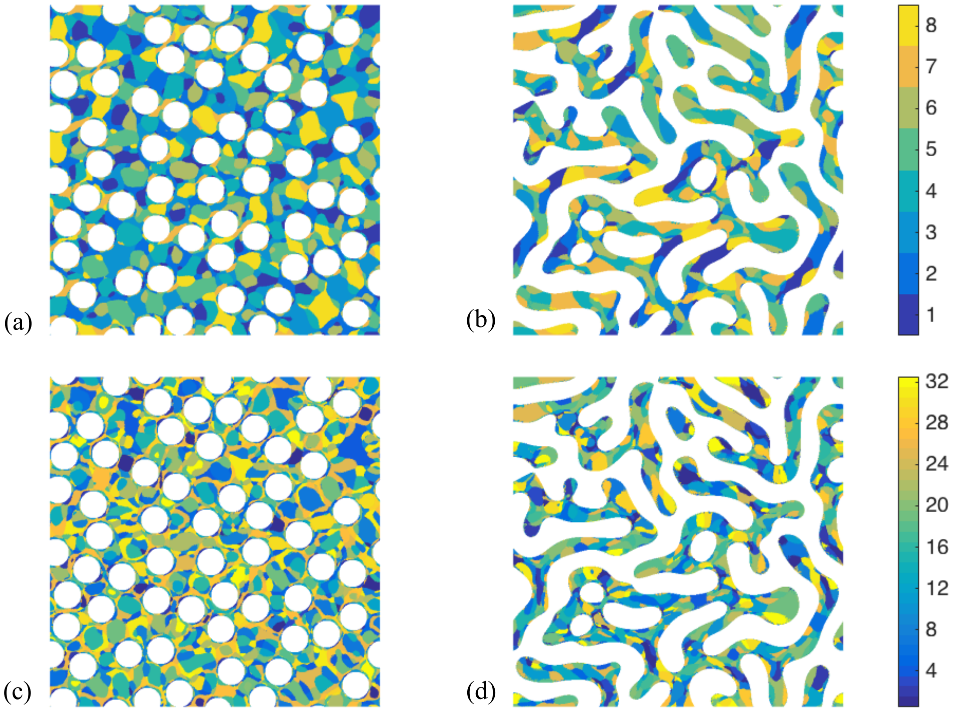


Figure 3: Two-dimensional k-means clustering results of (a,c) the cross-section of a fiber-reinforced composite with identical circular inclusions embedded in a matrix material and (b,d) a two-phase amorphous structure from a two-dimensional phase-field simulation based on the Cahn-Hilliard equation [44]. For both microstructures, phase 1 (matrix) was discretized into **8 clusters** in (a,b) and **32 clusters** in (c,d), while the discretization of phase 2 is not shown here for clarity. Note that **one** cluster is a group of **separate** subdomains indicated by the same color.

Figure 3a) and b) show the reduced RUC domain decomposition considering only 8 clusters for the matrix phase, while figures c) and d) were obtained considering 32 clusters for the same phase. Regions with the same color represent the same cluster. Recall that the methodology proposed in this work assumes that the local variables in each cluster are uniform. Therefore, in this example the finite element mesh of the high-fidelity RUC that had three hundred and sixty thousand elements was reduced to a few clusters (8 or 32) that can be used to solve the system of equations (17-18). This represents an extensive data compression that is expected to lead to tremendous computational savings in the posterior online stage.

Clustering methods have been widely applied to different fields, from pattern recognition and image segmentation to bioinformatics [45, 46]. Recently, a large spectrum of clustering algorithms has been developed, such as hierarchical clustering, k-means clustering and distribution-based clustering [47]. In this manuscript, a particular kind of centroid-based clustering named k-means clustering [43] is applied to compress the data. In k-means clustering all the data points are partitioned into k clusters in the following manner:

1. Evaluate the quantity of interest (e.g, strain-concentration tensor) at each point in the domain;
2. Group all the points into k clusters such that the difference between the quantity of interest at each point within the cluster is minimal when compared to the average value of that quantity inside that cluster.

For the case of the RUC, the suggested quantity of interest is the strain concentration tensor. Hence, this means that a cluster J contains the points whose strain concentration tensor $\mathbf{A}(\mathbf{x})$ is closer to the average strain concentration tensor of that cluster $\bar{\mathbf{A}}_J$ when compared to the average strain concentration tensor $\bar{\mathbf{A}}_I$ of any other cluster I (for $I \neq J$). Once again it is noted that the clusters are not necessarily of the same size and do not need to be continuous, see Figure 3.

Mathematically, given a set of strain concentration tensors the objective of k-means clustering is to minimize the within-cluster least squares sum for the k sets $\mathbf{S} = \{S^1, S^2, ..., S^k\}$ to obtain the shape of the clusters:

$$\mathbf{S} = \underset{\mathbf{S}'}{\operatorname{argmin}} \sum_{J=1}^{k} \sum_{n \in S^J} ||\mathbf{A}_n - \bar{\mathbf{A}}_J||^2, \tag{20}$$

where $\mathbf{A}_n$ is the strain concentration tensor of the n-th data point, and $\bar{\mathbf{A}}_J$ is the mean of all the strain concentration tensors at the points within the cluster S_J. The above norm is defined as usual, e.g. for a general second-order matrix $\mathbf{Z}$ of dimension $m \times m$

$$||\mathbf{Z}|| = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{m} z_{ij}^2} = \sqrt{\operatorname{trace}(\mathbf{Z}^T \mathbf{Z})}, \tag{21}$$

and is called the Frobenius norm of matrix $\mathbf{Z}$.

The algorithms for solving the k-means clustering problem are well developed, such as the standard algorithm (Lloyd's algorithm) [48] or fast k-means clustering algorithm proposed by Kanungo *et al.* [49]. In this article the standard algorithm was used, with more details provided in Appendix A. Note that the choice of the number of clusters k defines the degree of data compression achieved.

Mathematically, the "Silhouette Index" [50] can be used to measure the quality of clustering; however, this may not be sufficient when considering the predictions of the mechanical behavior. Typically, a larger number of clusters (larger k) used in the offline stage leads to more accurate predictions in the online stage since more information was stored, but the number of degrees of freedom increases correspondingly. The computational cost associated with this increase is discussed in later sections. As a result, the choice for the number of clusters should keep a balance between the compression ratio of the data set and the accuracy of the prediction.

Remark 3. *Clustering depends on specific choices of data type (in this case, strain concentration tensor), distance definition (Frobenius norm) and clustering algorithm (k-means clustering). Other selections could be made with direct influence on the mechanical behavior predicted. This exploratory effort is expected to be conducted in future studies.*

2.2.2. Computing the interaction tensors

After the domain decomposition that determines the material clusters, one needs to compute the interaction tensors between clusters. The interaction tensor $\mathbf{D}^{IJ}$ represents the influence of the stress in the J-th cluster on the strain in the I-th cluster. As mentioned previously, the interaction tensor can be written as an integral of Green's function in a RUC domain Ω with periodic boundary conditions,

$$\mathbf{D}^{IJ} = \frac{1}{c^I \mid \Omega \mid} \int_\Omega \int_\Omega \chi^I(\mathbf{x}) \chi^J(\mathbf{x}') \mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}') d\mathbf{x}' d\mathbf{x}, \tag{22}$$

where c^I is the volume fraction of the I-th cluster and $\mid \Omega \mid$ is the volume of domain Ω. $\mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}')$ is the fourth-order periodic Green's function associated with an isotropic linear elastic reference material with stiffness tensor $\mathbf{C}^0$. The definition and function of the reference material is discussed in Section 2.3. Also note that the interaction tensor $\mathbf{D}^{IJ}$ has to be a fourth-order tensor. In addition, by inspection of equation (16), it can be seen that

$$c^I \mathbf{D}^{IJ} = c^J \mathbf{D}^{JI}. \tag{23}$$

When the reference material is linear elastic (including isotropic and anisotropic), the Green's function is explicitly known in Fourier (frequency) space. For simplicity, we restrict the reference material to be isotropic linear elastic in this article, so that its Green's function takes the following form in Fourier space,

$$\hat{\Phi}^0_{ijkl}(\boldsymbol{\xi}) = \frac{1}{4\mu^0 \mid \boldsymbol{\xi} \mid^2} \left(\delta_{ik}\xi_j\xi_l + \delta_{il}\xi_j\xi_k + \delta_{jl}\xi_i\xi_k + \delta_{jk}\xi_i\xi_l \right) - \frac{\lambda^0 + \mu^0}{\mu^0(\lambda^0 + 2\mu^0)} \frac{\xi_i\xi_j\xi_k\xi_l}{\mid \boldsymbol{\xi} \mid^4}, \tag{24}$$

where $\boldsymbol{\xi}$ is the coordinate in Fourier space corresponding to $\mathbf{x}$ in real space, λ^0 and μ^0 are Lamé constants of the reference material. The formulation in equation (24) works for both 2D plane strain and 3D materials. For convenience, the Green's function $\hat{\Phi}^0_{ijkl}(\boldsymbol{\xi})$ can be rewritten in two parts:

$$\hat{\mathbf{\Phi}}^0(\boldsymbol{\xi}) = \frac{1}{4\mu^0} \hat{\mathbf{\Phi}}^1(\boldsymbol{\xi}) + \frac{\lambda^0 + \mu^0}{\mu^0(\lambda^0 + 2\mu^0)} \hat{\mathbf{\Phi}}^2(\boldsymbol{\xi}) \tag{25}$$

where each part can be written as

$$\hat{\Phi}^1_{ijkl}(\boldsymbol{\xi}) = \frac{1}{\mid \boldsymbol{\xi} \mid^2} \left(\delta_{ik}\xi_j\xi_l + \delta_{il}\xi_j\xi_k + \delta_{jl}\xi_i\xi_k + \delta_{jk}\xi_i\xi_l \right) \quad \text{and} \quad \hat{\Phi}^2_{ijkl}(\boldsymbol{\xi}) = -\frac{\xi_i\xi_j\xi_k\xi_l}{\mid \boldsymbol{\xi} \mid^4}, \tag{26}$$

Since $\hat{\mathbf{\Phi}}^1(\boldsymbol{\xi})$ and $\hat{\mathbf{\Phi}}^2(\boldsymbol{\xi})$ do not depend on the material properties, only the coefficients before them need to be updated if the elastic constants λ^0 and μ^0 of the reference material change. This feature of $\hat{\mathbf{\Phi}}^0(\boldsymbol{\xi})$ is essential to the self-consistent scheme as introduced in Section 2.3. Moreover, the convolution term in equation (16) can be translated into a direct multiplication based on a Fourier transformation,

$$\int_\Omega \chi^I(\mathbf{x}') \mathbf{\Phi}^0(\mathbf{x}, \mathbf{x}') d\mathbf{x}' = \mathscr{F}^{-1} \left(\hat{\chi}^I(\boldsymbol{\xi}) \hat{\mathbf{\Phi}}^0(\boldsymbol{\xi}) \right) \tag{27}$$

where $\hat{\chi}^I(\boldsymbol{\xi})$ is the characteristic function in the I-th cluster in the Fourier space. Only a one-time calculation is needed for computing all the interaction tensors, which can then be used for considering complex material behavior.

2.3. Online stage: Self-consistent clustering analysis

If an actual prediction of the local strains in the RUC for any constitutive behavior of the phases is to be obtained, then two points still need to be discussed: 1) the choice of the reference material stiffness $\mathbf{C}^0$; and 2) the algorithmic details to find the iterative solution for equations (17) and (18).

Concerning the first point, it is highlighted that the local strain $\boldsymbol{\varepsilon}(\mathbf{x})$ solved from the **continuous** Lippmann-Schwinger equation, i.e. equation (8) or (12), does not depend on the choice of $\mathbf{C}^0$. On the other hand, the choice of $\mathbf{C}^0$ affects the convergence rate of the iterative solution scheme for the discrete Lippmann-Schwinger equation as pointed out in [19, 51, 52, 53]. Despite the influence on the convergence rate, the converged solution reached by the discretized scheme should be the same and independent of the choice of $\mathbf{C}^0$, which could be regarded as a preconditioner of the iterative scheme. This can be explained by the fact that the physical problem is completely determined by the equilibrium condition and macroscopic constraints. The far-field strain $\boldsymbol{\varepsilon}^0$ in those equations just assumes a different value according to the choice of $\mathbf{C}^0$ such that the same solution for the local strain in the RUC is obtained. Therefore, *in theory* one can choose any reference stiffness $\mathbf{C}^0$ for the homogeneous linear elastic material. In practice, this is not the case – see Remark 4.

Remark 4. *The solution of the* ***discrete*** *Lippmann-Schwinger equations when finding the local strain in each cluster by solving equations (17) and (18) is actually affected by the choice of the reference stiffness* $\boldsymbol{C}^0$.

Section 2.3.1 presents the algorithmic details to solve the discrete equations (17) and (18) considering a constant $\mathbf{C}^0$ first. Then, in Section 2.3.2, a discussion of the importance of a self-consistent scheme for updating $\mathbf{C}^0$ upon deformation is presented.

2.3.1. Algorithm for constant $\boldsymbol{C}^0$

In general cases such as plasticity and damage, the incremental stress in the I-th cluster $\Delta\boldsymbol{\sigma}^I$ is a nonlinear function of its incremental strain $\Delta\boldsymbol{\varepsilon}^I$, which means that the system of integral equations needs to be solved iteratively at every load increment. An implicit scheme is adopted herein, and both macro-strain and macro-stress constraints are considered.

As mentioned before, the unknown variables of the system are the incremental strain in each cluster $\Delta\boldsymbol{\varepsilon}^I$ and the far field strain $\Delta\boldsymbol{\varepsilon}^0$, so $\{\Delta\boldsymbol{\varepsilon}\} = \{\Delta\boldsymbol{\varepsilon}^1; ...; \Delta\boldsymbol{\varepsilon}^k; \Delta\boldsymbol{\varepsilon}^0\}$. To begin, the residual of the integral equation in the I-th cluster at iteration step n is defined as

$$\mathbf{r}^I\left(\{\Delta\boldsymbol{\varepsilon}\}_n\right) = \Delta\boldsymbol{\varepsilon}^I_n + \sum_{J=1}^{k} \mathbf{D}^{IJ} : \left[\Delta\boldsymbol{\sigma}^J_n - \mathbf{C}^0 : \Delta\boldsymbol{\varepsilon}^J_n\right] - \Delta\boldsymbol{\varepsilon}^0_n \quad \text{with } I = 1, 2, ..., k, \tag{28}$$

and the residuals of the macro-strain or macro-stress constraints are

$$\mathbf{r}^{k+1}\left(\{\Delta\boldsymbol{\varepsilon}\}_n\right) = \sum_{I=1}^{k} c^I \Delta\boldsymbol{\varepsilon}^I_n - \Delta\bar{\boldsymbol{\varepsilon}} \quad \text{or} \quad \mathbf{r}^{k+1}\left(\{\Delta\boldsymbol{\varepsilon}\}_n\right) = \sum_{I=1}^{k} c^I \Delta\boldsymbol{\sigma}^I_n - \Delta\bar{\boldsymbol{\sigma}}. \tag{29}$$

In the implicit scheme, the residual $\{\mathbf{r}\} = \{\mathbf{r}^1; ...; \mathbf{r}^k; \mathbf{r}^{k+1}\}$ is linearized with respect to $\{\Delta\boldsymbol{\varepsilon}\}$. After dropping terms of higher order than linear, it gives

$$\{\mathbf{r}\} + \{\mathbf{M}\}\{\delta\boldsymbol{\varepsilon}\} = \mathbf{0} \quad \text{with } \{\mathbf{M}\} = \frac{\partial\{\mathbf{r}\}}{\partial\{\Delta\boldsymbol{\varepsilon}\}}, \tag{30}$$

where $\{\mathbf{M}\}$ is called the system Jacobian matrix. For $I, J = 1, 2, ..., k$, we have

$$\mathbf{M}^{IJ} = \delta_{IJ}\mathbf{I} + \mathbf{D}^{IJ} : \left(\mathbf{C}^J_{alg} - \mathbf{C}^0\right) \quad \text{and} \quad \mathbf{M}^{I(k+1)} = -\mathbf{I}, \tag{31}$$

where δ_{IJ} is the Kronecker delta in terms of indices I and J, and $\mathbf{I}$ is the fourth-order identity tensor. $\mathbf{C}^J_{alg}$ is the so-called algorithmic modulus (or tangent modulus) of the material in the J-th cluster and is an output of the local constitutive law for the current strain increment in that cluster $\Delta\boldsymbol{\varepsilon}^J_n$,

$$\mathbf{C}^J_{alg} = \frac{\partial\Delta\boldsymbol{\sigma}^J}{\partial\Delta\boldsymbol{\varepsilon}^J}. \tag{32}$$

Under the macro-strain constraint, the remaining components in the system Jacobian matrix are

$$\mathbf{M}^{(k+1)I} = c^I\mathbf{I} \quad \text{and} \quad \mathbf{M}^{(k+1)(k+1)} = \mathbf{0}. \tag{33}$$

For macro-stress constraint, we have

$$\mathbf{M}^{(k+1)I} = c^I\mathbf{C}^I_{alg} \quad \text{and} \quad \mathbf{M}^{(k+1)(k+1)} = \mathbf{0}. \tag{34}$$

Finally, the correction of the incremental strain can be expressed as

$$\{\delta\boldsymbol{\varepsilon}\} = -\{\mathbf{M}\}^{-1}\{\mathbf{r}\}. \tag{35}$$

Based on the updated incremental strain, we can then use the constitutive relationship in each cluster to compute the new incremental stress $\{\Delta\boldsymbol{\sigma}\} = \{\Delta\boldsymbol{\sigma}^1; ...; \Delta\boldsymbol{\sigma}^k\}$. The algorithm for the implicit scheme is given in the box below, where it is again highlighted that the constant $\mathbf{C}^0 = \mathbf{C}^{\text{input}}$ is chosen by the analyst as an input $\mathbf{C}^{\text{input}}$. The most logic choice is to consider as reference the effective stiffness of the RUC that can be determined in the offline stage $\mathbf{C}^{\text{input}} = \mathbf{C}^{\text{macro}}$. However, as previously mentioned, other reference stiffnesses could be considered, e.g. the stiffness of the matrix phase $\mathbf{C}^{\text{input}} = \mathbf{C}^{\text{matrix}}$.

Box 2.1 Algorithm for the implicit scheme with constant $\mathbf{C}^0$

```
1. Initial conditions and initialization: set $\{\Delta\boldsymbol{\varepsilon}\}_0 = \mathbf{0}; n = 0; \{\Delta\boldsymbol{\varepsilon}\}_{\text{new}} = \{\Delta\boldsymbol{\varepsilon}\}_0; \mathbf{C}^0 = \mathbf{C}^{\text{input}}$
2. Newton iterations for load increment $n+1$:
   a. compute the incremental stress $\{\Delta\boldsymbol{\sigma}\}_{\text{new}}$ based on material constitutive laws
   b. use 28 and 29 to compute the residual $\{\mathbf{r}\} = \mathbf{f}(\{\Delta\boldsymbol{\varepsilon}\}_{\text{new}}, \{\Delta\boldsymbol{\sigma}\}_{\text{new}})$
   c. compute the system Jacobian $\{\mathbf{M}\}$
   d. solve the linear equation $\{\delta\boldsymbol{\varepsilon}\} = -\{\mathbf{M}\}^{-1}\{\mathbf{r}\}$
   e. $\{\Delta\boldsymbol{\varepsilon}\}_{\text{new}} \leftarrow \{\Delta\boldsymbol{\varepsilon}\}_{\text{new}} + \{\delta\boldsymbol{\varepsilon}\}$
   f. check error criterion; if not met, go to 2a.
3. Update the incremental strain and stress: $\{\Delta\boldsymbol{\varepsilon}\}_{n+1} = \{\Delta\boldsymbol{\varepsilon}\}_{\text{new}}, \{\Delta\boldsymbol{\sigma}\}_{n+1} = \{\Delta\boldsymbol{\sigma}\}_{\text{new}}; n \leftarrow n+1$
4. If simulation not complete, go to 2.
```

2.3.2. Algorithm with self-consistent $\mathbf{C}^0$

Remark 4 highlights that due to the discrete character of the approach the actual choice of the stiffness of the reference material $\mathbf{C}^0$ can affect the solution, contrary to what happens for the continuous form. This raises a question: Is there an optimal choice for $\mathbf{C}^0$?

Here a self-consistent scheme is proposed to set the homogeneous tangent modulus of the reference material $\mathbf{C}^0$ approximately the same as the macroscopic tangent modulus (i.e. the effective tangent modulus of the RUC) even **after** plastic deformation occurs,

$$\mathbf{C}^0 \to \mathbf{C}^{\text{macro}}. \tag{36}$$

Let's start by considering the elastic case first, where the effective tangent modulus of the RUC does not change and coincides with the effective elastic stiffness of the RUC. By averaging the incremental integral equation 12 in the RUC domain Ω, we have

$$\Delta\boldsymbol{\varepsilon}^{macro} + \frac{1}{|\Omega|}\int_\Omega \left(\int_\Omega \boldsymbol{\Phi}^0(\mathbf{x}, \mathbf{x}')d\mathbf{x}\right) : \left[\Delta\boldsymbol{\sigma}(\mathbf{x}') - \mathbf{C}^{\text{macro}} : \Delta\boldsymbol{\varepsilon}(\mathbf{x}')\right] d\mathbf{x}' - \Delta\boldsymbol{\varepsilon}^0 = 0. \tag{37}$$

According to the definition of periodic Green's function $\boldsymbol{\Phi}^0(\mathbf{x}, \mathbf{x}')$, the following integral results in a constant tensor which does not depend on $\mathbf{x}'$,

$$\int_\Omega \boldsymbol{\Phi}^0(\mathbf{x}, \mathbf{x}')d\mathbf{x} = \mathbf{S} : \left(\mathbf{C}^0\right)^{-1}, \tag{38}$$

where $\mathbf{S}$ is the Eshelby's tensor. By substituting equation (38) into (37), the following statement can be proved,

$$\Delta\boldsymbol{\varepsilon}^0 = \Delta\boldsymbol{\varepsilon}^{macro}, \quad \text{if } \mathbf{C}^0 = \mathbf{C}^{\text{macro}}. \tag{39}$$

This result can be extrapolated to the case of irreversible processes such as plasticity. For a given increment n, when $\mathbf{C}^0_n$ tends to the tangent modulus of the RUC $\mathbf{C}^{\text{macro}}_n$, then the far-field strain $\Delta\boldsymbol{\varepsilon}^0_n$ tends to the macroscopic strain $\Delta\boldsymbol{\varepsilon}^{\text{macro}}_n$. Hence, the name **self-consistent** reference tangent modulus.

For convenience, however, the tangent modulus $\mathbf{C}_n^{\mathrm{macro}}$ is considered as isotropic even during plasticity in order to still use the same form of Green's functions, see equation (24). This is not true in general, since the plastic flow may introduce anisotropic behavior, especially under large deformations.

Considering the reference tangent modulus as isotropic at every macroscopic load increment, the self-consistent scheme then needs to find the two independent Lamé parameters λ_n^0 and μ_n^0 for every load increment n that define $\mathbf{C}_n^0$,

$$\mathbf{C}_n^0 = \mathbf{f}\left(\lambda_n^0, \mu_n^0\right), \tag{40}$$

such that the reference tangent modulus is closest to the macroscopic tangent modulus. The objective of the self-consistent scheme is then to find λ_n^0 and μ_n^0 from the following minimization,

$$\left\{\lambda_n^0, \mu_n^0\right\} = \underset{\{\lambda', \mu'\}}{\mathrm{argmin}} \; ||\Delta\boldsymbol{\sigma}_n^{macro} - \mathbf{f}(\lambda', \mu') : \Delta\boldsymbol{\varepsilon}_n^{macro}||^2, \tag{41}$$

where $||\mathbf{Z}||^2 = \mathbf{Z} : \mathbf{Z}$ for an arbitrary second-order tensor $\mathbf{Z}$. The macroscopic strain $\boldsymbol{\varepsilon}_n^{macro}$ and stress $\boldsymbol{\sigma}_n^{macro}$ at increment n are given by

$$\Delta\boldsymbol{\varepsilon}_n^{macro} = \sum_{I=1}^{k} c^I \Delta\boldsymbol{\varepsilon}_n^I \quad \text{and} \quad \Delta\boldsymbol{\sigma}_n^{macro} = \sum_{I=1}^{k} c^I \Delta\boldsymbol{\sigma}_n^I. \tag{42}$$

Then the optimum point $\left(\lambda_{\mathrm{opt}}^0, \mu_{\mathrm{opt}}^0\right)$ is found by computing the minimum of $g(\lambda_n^0, \mu_n^0)$,

$$g(\lambda_n^0, \mu_n^0) = ||\Delta\boldsymbol{\sigma}_n^{macro} - \mathbf{f}\left(\lambda_n^0, \mu_n^0\right) : \Delta\boldsymbol{\varepsilon}_n^{macro}||^2 \tag{43}$$

via the respective partial derivatives,

$$\left.\frac{\partial g}{\partial \lambda_n^0}\right|_{\lambda_{\mathrm{opt}}^0, \mu_{\mathrm{opt}}^0} = 0 \quad \text{and} \quad \left.\frac{\partial g}{\partial \mu_n^0}\right|_{\lambda_{\mathrm{opt}}^0, \mu_{\mathrm{opt}}^0} = 0, \tag{44}$$

which basically forms a system of two linear equations in terms of the Lamé constants. The system always has a unique solution except under a pure-shear loading condition, where λ_{opt}^0 is underdetermined. In this case, the value of λ_n^0 is not updated. Additionally, $g\left(\lambda_{\mathrm{opt}}^0, \mu_{\mathrm{opt}}^0\right)$ vanishes when the effective macroscopic homogeneous material is also isotropic linear elastic.

This self-consistent scheme is also iterative, since the macroscopic incremental stress $\Delta\boldsymbol{\sigma}_{\mathrm{new}}^{macro}$ and strain $\Delta\boldsymbol{\varepsilon}_{\mathrm{new}}^{macro}$ at the beginning of the increment are computed based on previous values of λ_n^0 and μ_n^0. Note that in the algorithm that uses a constant $\mathbf{C}^0$, see **Box 2.1**, since the tangent modulus $\mathbf{C}^0$ is not updated that implies that the interaction tensors $\mathbf{D}^{IJ}$ also do not need to be updated. This is no longer the case for the self-consistent algorithm, as presented in **Box 2.2**, since the interaction tensors given by equation (16) also depend on the updated reference Lamé parameters following equation (24). It is also noted that for the simulations considered in this work, the convergence of λ^0 and μ^0 can be reached within a small number of iterations (e.g. less than 5 for a tolerance 0.1%).

Box 2.2 Algorithm for the updated self-consistent implicit scheme

```
1. Initial conditions and initialization: set $\left(\lambda^0, \mu^0\right)$; $\{\Delta\boldsymbol{\varepsilon}\}_0 = \mathbf{0}; n = 0; \{\Delta\boldsymbol{\varepsilon}\}_{\text{new}} = \{\Delta\boldsymbol{\varepsilon}\}_0$
2. For load increment $n+1$, update the interaction tensor $\mathbf{D}^{IJ}$ and the stiffness tensor $\mathbf{C}^0$
3. Newton iterations:
   a. compute the incremental stress $\{\Delta\boldsymbol{\sigma}\}_{\text{new}}$
   b. use 28 and 29 to compute the residual $\{\mathbf{r}\} = \mathbf{f}\left(\{\Delta\boldsymbol{\varepsilon}\}_{\text{new}}, \{\Delta\boldsymbol{\sigma}\}_{\text{new}}\right)$
   c. compute the system Jacobian $\{\mathbf{M}\}$
   d. solve the linear equation $\{\delta\boldsymbol{\varepsilon}\} = -\{\mathbf{M}\}^{-1}\{\mathbf{r}\}$
   e. $\{\Delta\boldsymbol{\varepsilon}\}_{\text{new}} \leftarrow \{\Delta\boldsymbol{\varepsilon}\}_{\text{new}} + \{\delta\boldsymbol{\varepsilon}\}$
   f. check error criterion; if not met, go to 3a.
4. Solve 44, and $\left(\lambda^0, \mu^0\right) \leftarrow \left(\lambda^0_{\text{opt}}, \mu^0_{\text{opt}}\right)$
5. Check error criterion; if not met, go to 2
6. Update the incremental strain and stress: $\{\Delta\boldsymbol{\varepsilon}\}_{n+1} = \{\Delta\boldsymbol{\varepsilon}\}_{\text{new}}, \{\Delta\boldsymbol{\sigma}\}_{n+1} = \{\Delta\boldsymbol{\sigma}\}_{\text{new}}; n \leftarrow n+1$
7. If simulation not complete, go to 2.
```

With this last update we obtained a "self-consistent Lippmann-Schwinger multiple-cluster homogenization scheme" made computationally efficient by the data compression treatment proposed in Section 2.2.1 where clusters (subdomains) of material with similar mechanical properties were determined. The self-consistency of the proposed scheme is optional, i.e. it is possible to use a constant reference stiffness tensor $\mathbf{C}^0$. However, the convergence of the scheme is improved as shown next. We note that the scheme is standalone: there are no additional parameters nor evolution laws, unlike the schemes reviewed in Section 1.

2.4. Microstructure and material modeling

The proposed reduced order homogenization method was applied to 2D plane strain and 3D heterogeneous materials under different load cases. As detailed in the previous sections, the whole modeling framework can be divided into two stages: 1) offline stage; and 2) online stage.

In the 2D plane strain examples, two types of materials with different microstructures as shown in Figure 3 are analyzed. The first one is a composite material with identical circular fibers (phase 2) embedded in the matrix (phase 1), where the fiber volume fraction was considered as $V_{f_2} = 30\%$. The second 2D plane strain example is a two-phase material with microstructure obtained from a phase-field simulation based on the Cahn-Hilliard equation, where the volume fraction of either phase is equal to 50%. For convenience, the two types of materials are abbreviated as fiber-reinforced composite and amorphous material, respectively.

Two examples in 3D are also presented: one was chosen as a two-phase composite material with identical spherical inclusions embedded in the matrix (volume fraction of 20%), and the other as an amorphous material with a microstructure obtained from a 3D phase-field simulation (volume fraction of 50%). All the microstructures considered in the examples are periodic.

The same material properties are used in the 2D and 3D problems. For the linear elastic direct numerical simulations in the offline stage, the Young's moduli and Poisson's ratios of phase 1 and 2 are:

$$E_1 = 100 \text{ MPa}, \nu_1 = 0.3; \quad E_2 = 500 \text{ MPa}, \nu_2 = 0.19. \tag{45}$$

Figure 3 showed the data compression achieved for the two RUCs used in the 2D examples considering 8 or 32 clusters for phase 1 of the materials. Each reduced RUC obtained for a chosen number of clusters is then analyzed in the online stage. Note that the same reduced RUC can be used for predicting plastic behavior of the material with any local plasticity laws for each phase (no additional calibration needed). Therefore, the following examples considered two different plasticity laws in order to assess the predictive capabilities of the proposed approach.

Two plasticity laws were considered in the online stage for phase 1 (the matrix material). Both laws considered a von Mises yield surface,

$$f = \bar{\sigma} - \sigma_Y(\bar{\varepsilon}) \leqslant 0, \tag{46}$$

where $\bar{\sigma}$ is the von Mises equivalent stress, and the yield stress σ_Y is given by the chosen hardening law that depends on the equivalent plastic strain $\bar{\varepsilon}$.

The first hardening law considered was piecewise linear and isotropic as follows,

$$\sigma_Y(\bar{\varepsilon}) = \begin{cases} 0.5 + 5\bar{\varepsilon} & \bar{\varepsilon} \in [0, 0.04) \\ 0.7 + 2\bar{\varepsilon} & \bar{\varepsilon} \in [0.04, \infty) \end{cases} \text{ MPa.} \tag{47}$$

The other hardening law considered was the following power-law,

$$\sigma_Y(\bar{\varepsilon}) = 0.1 + 0.3\bar{\varepsilon}^{0.4} \text{ MPa.} \tag{48}$$

The plasticity simulations were conducted under different loading conditions. For example, under uniaxial tension the macroscopic strain ε_{11} is prescribed, and the macroscopic stresses vanish in all other directions: $\sigma_{ij} = 0$ for $i, j \neq 1$. For pure shear loading condition, γ_{12} is imposed and the remaining stress-free conditions are applied similarly.

The predictions given by the reduced order modeling, including macroscopic behavior and microscopic stress/strain fields, are validated against high-fidelity direct numerical simulations (DNS) using the finite element method with a fine mesh and imposing periodic boundary conditions. The computational efficiency of the proposed method is discussed as well.

3. Results and discussion

3.1. Two-dimensional linear elastic materials

Before evaluating the plastic predictions of the reduced RUC, the elastic description is considered first. The importance of applying the self-consistent (SC) scheme for updating $\mathbf{C}^0$ is discussed herein as well.

Figure 4 presents the RUC's transverse Young's moduli and Poisson's ratios predicted by the proposed method with and without the self-consistent (SC) scheme as a function of the number of clusters chosen in the offline stage. For the approach without the self-consistent scheme, i.e. considering a constant reference stiffness $\mathbf{C}^0$, phase 1 (matrix) was considered as the reference material: $\mathbf{C}^0 = \mathbf{C}^{\text{matrix}}$. The number of clusters in phase 1 is denoted as k_1, while k_2 denotes the number of clusters in phase 2. In these examples k_2 was defined according to the volume fractions of the phases: for the fiber-reinforced composite the number of clusters in the fibers was chosen as approximately half the number of clusters in the matrix $k_2 = \lceil k_1/2 \rceil$ (recall that $V_{f_2} = 30\%$); and for the amorphous material both phases were discretized with the same number of clusters $k_2 = k_1$ (recall that $V_{f_1} = 50\%$). All the predictions are normalized with respect to the results of direct numerical simulations (DNS) obtained from a converged finite element analysis considering a fine mesh of elements:

1. Fiber-reinforced composite: $E_{\text{DNS}} \approx 156.4$ MPa, $\nu_{\text{DNS}} \approx 0.39$;
2. Amorphous material: $E_{\text{DNS}} \approx 220.8$ MPa, $\nu_{\text{DNS}} \approx 0.35$.

Observing the figure there is an immediate characteristic of the presented solutions: the accuracy of the prediction improves by increasing the number of clusters in the system, as expected. Moreover, the predictions for the effective E and ν of the RUC for the fiber-reinforced composite are similar with or without the self-consistent (SC) scheme: both approaches converge within an error tolerance of 1% after assigning 8 clusters in phase 1 (matrix). However, convergence is significantly faster with the self-consistent scheme[3] for the amorphous material (both for E and ν). These results confirm Remark 4: since the Lippmann-Schwinger equation is being numerically discretized instead of being solved exactly, the choice of the reference material affects the convergence of the solution. This was also pointed out by Chaboche et al. [39] in their complete analysis of the capabilities of mean-field approaches.

---

[3] For the linear-elastic examples, the self-consistent scheme finds $\mathbf{C}^0$ to be the same as the effective stiffness of the RUC $\mathbf{C}^{\text{macro}}$. So, considering $\mathbf{C}^{\text{input}} = \mathbf{C}^{\text{macro}}$ in the non self-consistent scheme of **Box 2.1** would lead to the same convergence as the self-consistent scheme (for linear elasticity only).

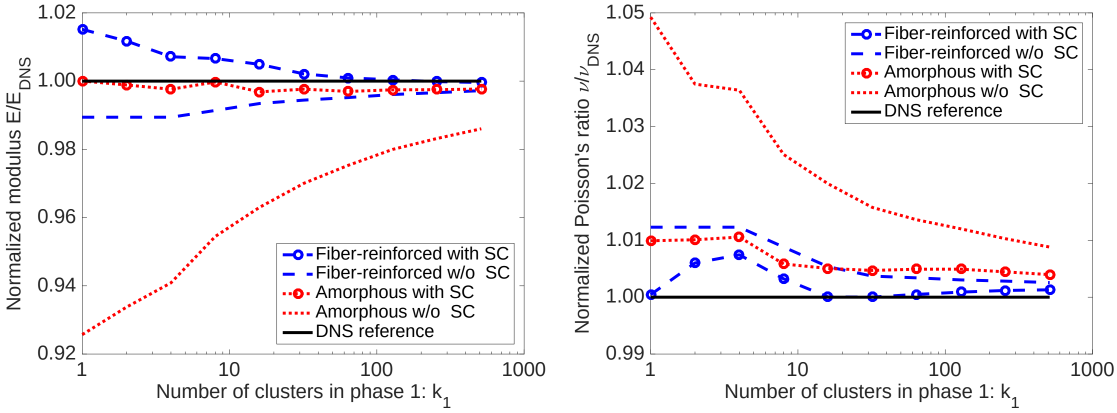


Figure 4: Normalized transverse Young modulus (left) and Poisson's ratio (right) as a function of the number of cluster in phase 1 (matrix). The DNS results for the fiber-reinforced composite and amorphous material were obtained from finite element analyses.

The predictive capabilities of the proposed method can be further assessed by considering the variation of the elastic properties at the online stage without redoing the offline stage. The interested reader is referred to Appendix B and Figure B.14, where new effective elastic properties of the RUC were reasonably well predicted even when using new combinations of elastic properties for the two phases of the material that differ from the ones used in the offline stage, equation (45). This shows the potential of applying the proposed method to microstructure-based property design of heterogeneous material system, such as nanostructured polymers [54].

3.2. Two-dimensional nonlinear elasto-plastic materials

Particularly relevant is the assessment of the predictive capabilities of the proposed model when capturing the nonlinear plastic behavior of the RUC under various loading conditions, as well as the computational cost (CPU time and memory) when compared to the DNS. First the influence of the self-consistent scheme on the plastic predictions is investigated considering the two hardening laws introduced previously: the piece-wise linear hardening, equation (47); and the power-law hardening, equation (48). Figure 5 presents the stress-strain results for the fiber-reinforced composite subjected to uniaxial tension for these two hardening laws of the matrix.

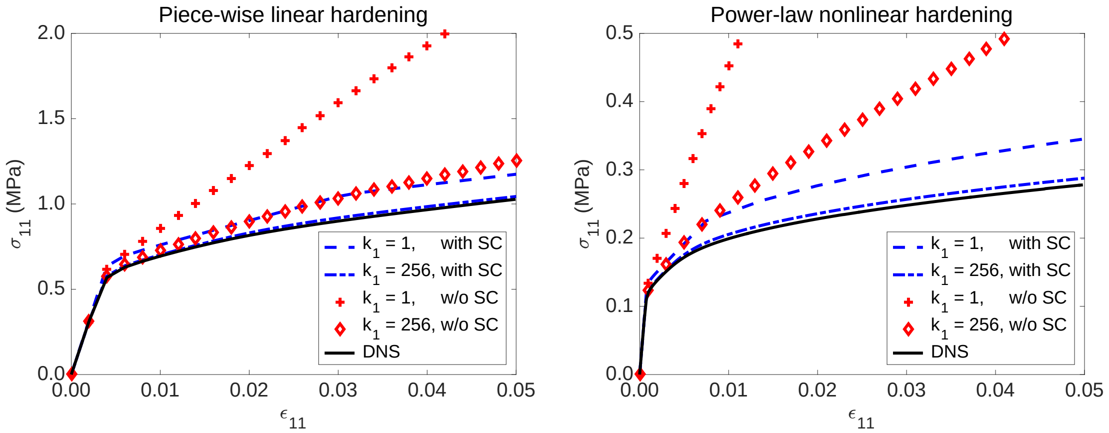


Figure 5: The effect of the self-consistent scheme on the plastic behavior of the fiber-reinforced material under piece-wise linear hardening (left) and power-law nonlinear hardening (right) loading conditions.

As can be observed in the figure, although the error of the method when compared to DNS is reduced in each case by increasing the number of clusters, the self-consistent scheme converges significantly faster for a desired level of accuracy. This is especially visible for the power-law nonlinear hardening. When plastic yielding happens in the heterogeneous material, the macroscopic tangent stiffness tensor starts decreasing when compared to the elastic case due to localized plasticity in the matrix, so that $\mathbf{C}^{\text{matrix}}$ is no longer a good choice for the tangent stiffness $\mathbf{C}^0$ of the reference material. By consistently updating $\mathbf{C}^0$ with an approximate value of the effective stiffness of the RUC, see Box 2.2, the evolution of the local internal variables in each cluster is captured more accurately.

Another point to be considered is the effect of the **A**-based clustering used during the offline stage, introduced in Section 2.2.1. As discussed in that section, the k-means clustering was done based on the strain concentration tensor $\mathbf{A}(\mathbf{x})$ at each data point, with the purpose of grouping the data points with similar mechanical behavior into one cluster. A more naive compression scheme would be to do the clustering purely based on the spatial coordinates $\mathbf{x}$ of the data points, i.e. grouping points by their spatial proximity instead of considering the similarity of mechanical behavior. This position-based clustering leads to a cluster map resembling a Voronoi diagram, as illustrated in Figure 6 (left) for 32 clusters in phase 1 for the fiber-reinforced composite. The reader is encouraged to compare this figure with Figure 3c) to see the difference in the clustering scheme for the **same** number of clusters but using the **A**-based clustering.

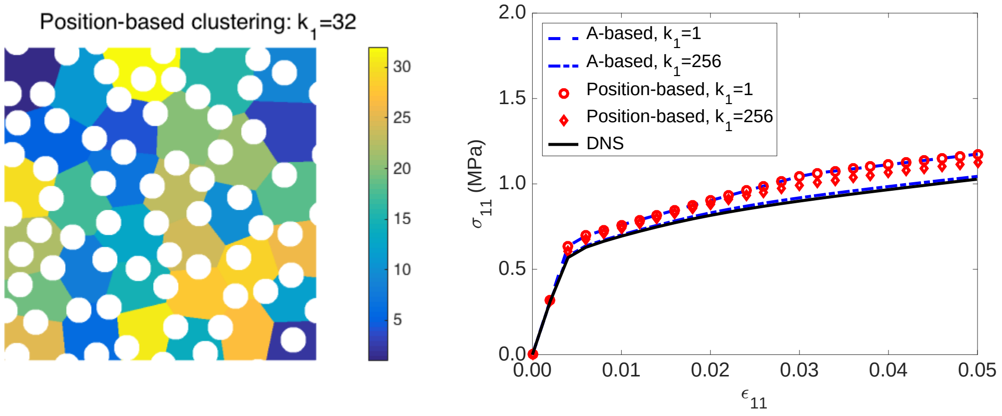


Figure 6: The effect of the raw data type on the clustering results (left), as well as the corresponding predictions on macroscopic response (right), for the fiber-reinforced composite under uniaxial tension. The position-based clustering results with 32 clusters are shown on the left.

On the right of Figure 6 the results of position-based clustering are compared to the **A**-based clustering. In this figure it is seen that the convergence using the position-based clusters is poor and the accuracy of the prediction does not change significantly even considering 256 clusters for phase 1. On the contrary, with the same number of clusters the predictions from **A**-based clustering reproduce the DNS almost exactly.

Figure 7 includes the stress-strain curves for the fiber-reinforced composite and the amorphous material considering two different load cases, uniaxial tension and pure shear, using the self-consistent scheme. The number of clusters in phase 1 ranges from 1 to 256, and the results from DNS are also provided for comparison. By looking at the plots, we can conclude that the proposed method is capable of capturing the nonlinear plastic behavior with significantly fewer degrees of freedom but with a small loss in accuracy. Furthermore, the refinement of the discretization by including more clusters clearly improves the predictions: a desirable characteristic allowing the analyst to choose between accuracy and efficiency.

Figure 8 provides a comparison of the local fields obtained from DNS of the high-fidelity RUC and from the proposed self-consistent clustering analysis of a reduced RUC with 256 clusters for phase 1. The figures show that the new method is able to reproduce the equivalent plastic strain fields, although the localized regions are reproduced in a more diffuse manner leading to regions with lower strain concentrations. This is expected since by definition the goal of the proposed approach is to predict as accurately as possible the global (homogenized) mechanical response with the minimum amount of information possible. In order to improve the resolution at these regions, the total number of clusters could be increased or, more wisely, a

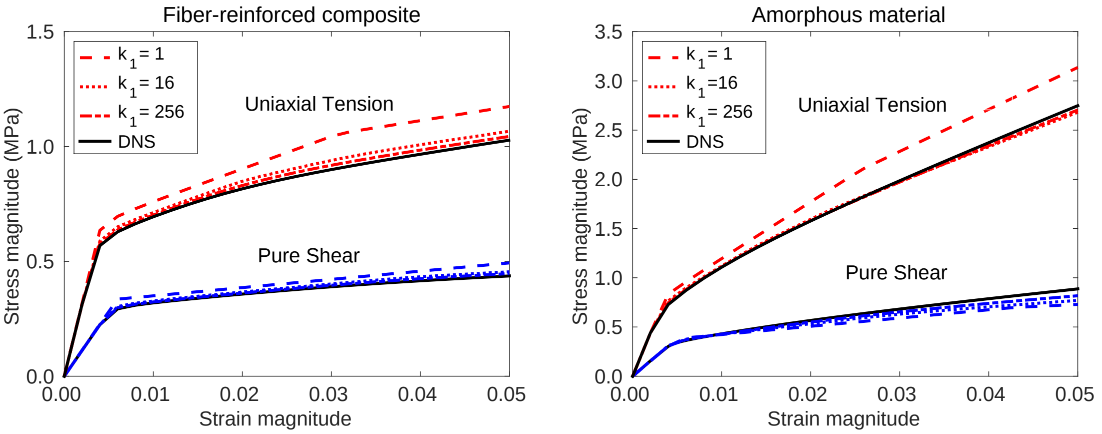
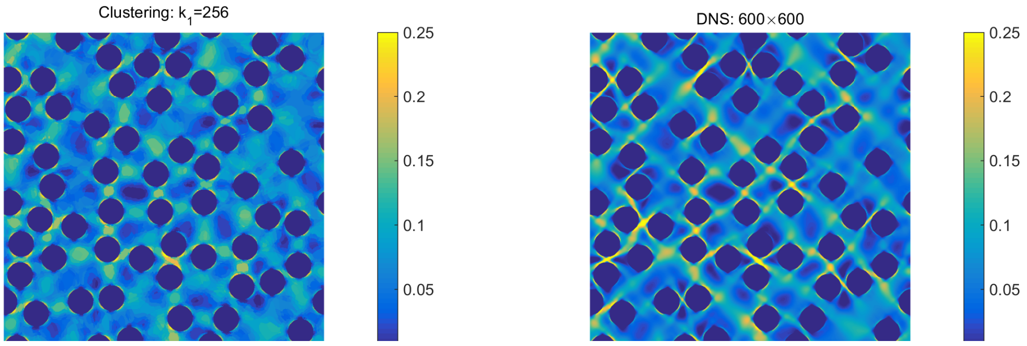
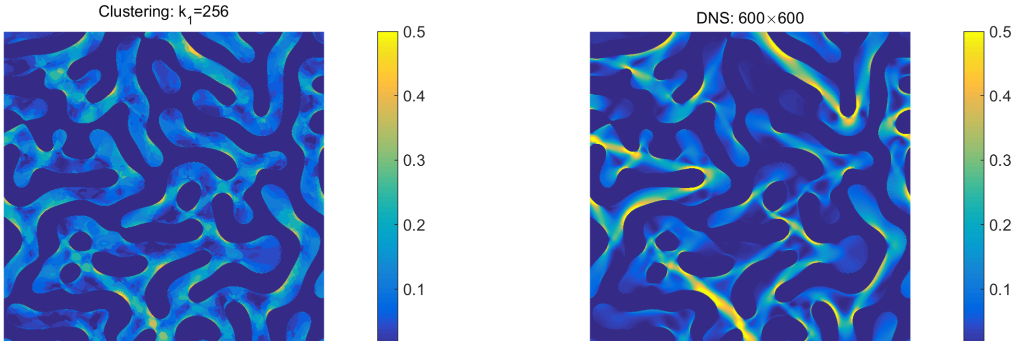


Figure 7: The predictions given by the proposed reduced order method with the self-consistent scheme for the fiber-reinforced composite (left) and amorphous material (right) under uniaxial tension and pure shear conditions. The solid lines represent the DNS results for comparison. Three different numbers of clusters in phase 1 were considered: $k_1 = 1$ (dashes), $k_1 = 16$ (dots) and $k_1 = 256$ (dash-dots).

nonlinear mapping of the k-means clustering could be applied so that more clusters are assigned to regions needing higher resolution (the issue is finding these regions, since different loading conditions lead to different localization areas).



(a) Reduced RUC for composite material (b) High-fidelity RUC for composite material



(c) Reduced RUC for amorphous material (d) High-fidelity RUC for amorphous material

Figure 8: Equivalent plastic strain fields obtained from the direct numerical simulation of the high-fidelity RUC and from the self-consistent clustering analysis of the reduced RUC. Top row shows results for fiber-reinforced composite material, while bottom row presents the results for the amorphous material

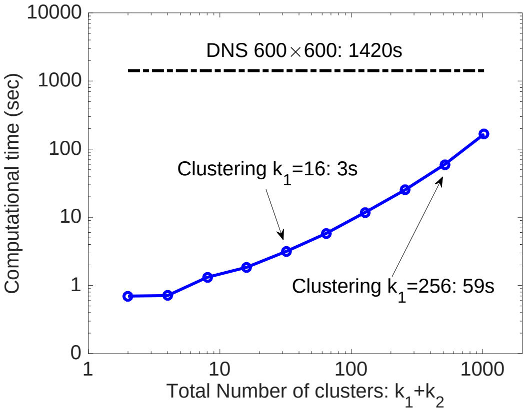


Figure 9: The computational time of the reduced order method with the self-consistent scheme vs. the total number of clusters $k_1 + k_2$. 25 incremental loading steps are considered. A typical DNS with mesh size 600 × 600 takes 1420s.

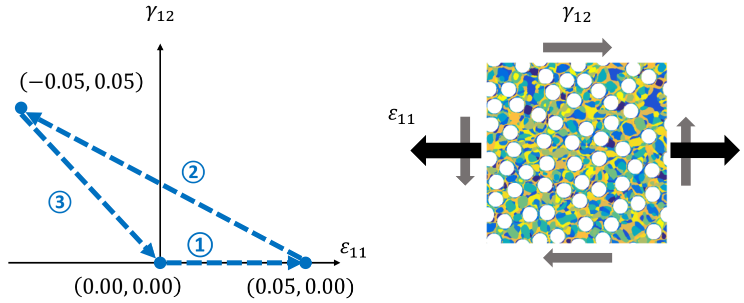
Up to this point the computational cost reduction of the method was not quantified. Even though the reduction achieved in the number of degrees of freedom is obvious, the system of equations that is being solved uses an iterative approach and each equation involves the summation over every cluster. Therefore, there is a clear trade-off between the compression of information achieved by the clustering procedure and the actual solution of the Lippmann-Schwinger equations. Figure 9 shows a comparison of computation time between the proposed approach for different numbers of clusters and the DNS. The results are encouraging, especially considering that the DNS finite element analyses were performed using a commercial code (significantly optimized), while the proposed method was implemented in MATLAB. A typical two-dimensional DNS took about 1420s ($\approx$ 24min) on one intel i7-3632 processor, while the online stage of the reduced order method with the self-consistent scheme (in MATLAB) took 4s for $k_1 = 16$ and 75s for $k_1 = 256$. For $k_1 \in [1, 512]$, the computational time is almost proportional to the number of clusters in the system, which indicates that the most time-consuming part is to update the internal variables locally at each cluster.

Finally, the reduced order model was also validated against a complex loading path. Macro-strain constraints are applied on ε_{11} and γ_{12}, while $\sigma_{22} = 0$. Thus, the loading state can be represented by $(\varepsilon_{11}, \gamma_{12})$. As shown in Figure 10a), there are three steps in the loading path and the material finally returns to the initial state ($\varepsilon_{11} = 0, \gamma_{12} = 0$). Due to plasticity, ε_{22} does not necessarily vanish at the final state, which means that the initial state with $\varepsilon_{22} = 0$ is not reproduced. The stress-strain curves σ_{11}-vs-ε_{11} and σ_{12}-vs-γ_{12} given by the reduced-order method and the DNS results are shown in Figure 10 b) and c). Once again good agreement is observed.

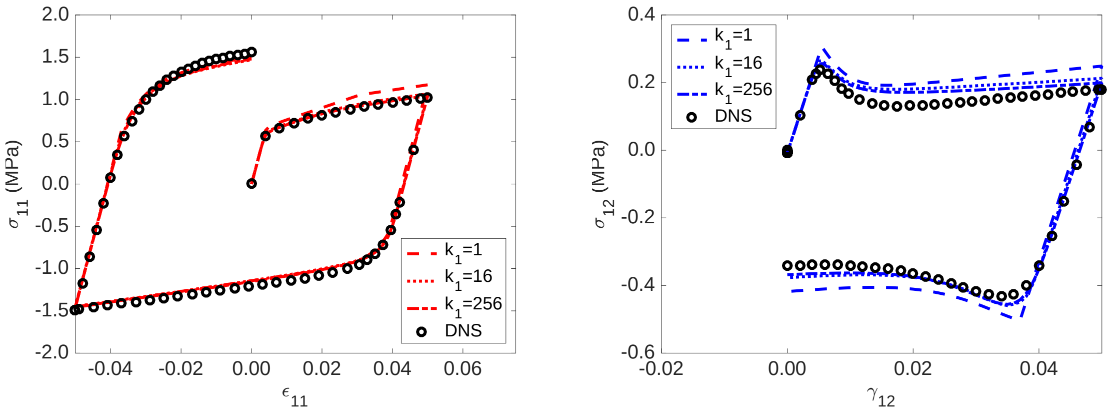
3.3. Three-dimensional nonlinear elasto-plastic material

The proposed method was also applied to 3D nonlinear elasto-plastic materials similar to the previously presented examples in 2D. First, a spherical-particle composite is considered with the phase 1 being the matrix (as before) and the number of matrix clusters being k_1. The volume fraction of the particle phase (phase 2) was considered to be 20% and $k_2 = \lceil k_1/4 \rceil$. All the material properties are the same as before (see equations (45), (46) and (47)). The mesh size of the finite element model considered as the DNS was 64 × 64 × 64, and the illustration of the mesh in the inclusion phase (phase 2) is demonstrated in Figure 11. The reader is reminded that when generating the raw data, the strain concentration tensor $\mathbf{A}(\mathbf{x})$ has 36 independent components since this is a 3D RUC. Thus, the dimension of data becomes 36, and a linear increase in the running time of the k-means clustering occurs.

The stress-strain curves given by the self-consistent clustering analysis of the reduced RUC and the DNS of the high-fidelity RUC results are plotted in Figure 11. Although the convergence rates of macroscopic mechanical properties with respect to the number of clusters in 3D is not as high as the ones in 2D, good accuracy can be achieved by assigning 256 clusters in phase 1 (320 in total) under both loading conditions. In terms of computational cost, a typical DNS with 64 × 64 × 64 mesh took 7.3 hours on 24 cores (in a



(a) Illustrations of the complex loading path $(\varepsilon_{11}, \gamma_{12})$, where ε_{11} and γ_{12} are imposed as shown and $\sigma_{22} = 0$



(b) Normal stress versus normal strain

(c) Shear stress versus shear strain

Figure 10: Response of the reduced RUC for the fiber-reinforced composite material under the three step loading path shown in (a): $(0.00, 0.00) \to (0.05, 0.00) \to (-0.05, 0.05) \to (0.00, 0.00)$. The σ_{11}-vs-ε_{11} response is shown in (b), and the σ_{12}-vs-ε_{12} response is presented in (c).

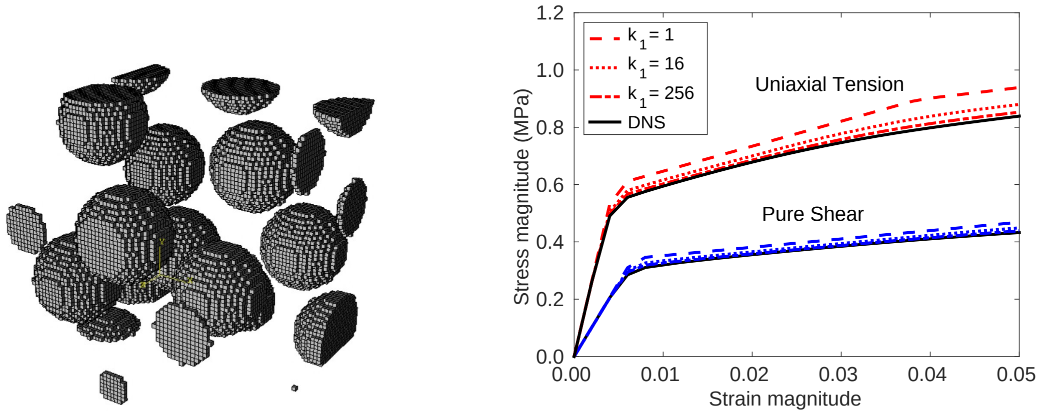


Figure 11: The 64 × 64 × 64 FE mesh of the 3D material displayed in the inclusion phase (phase 2) and stress-strain curves under uniaxial tension and pure shear loading conditions. The solid lines represent the DNS results for comparison. Three different numbers of clusters in phase 1 were considered: $k_1 = 1$ (dashes), $k_1 = 16$ (dots) and $k_1 = 256$ (dash-dots).

state-of-the art high performance computing cluster with the following compute nodes: Intel Haswell E5-2680v3 2.5GHz 12-cores). The reduced order method (in MATLAB) took 5s and 214s on one intel i7-3632

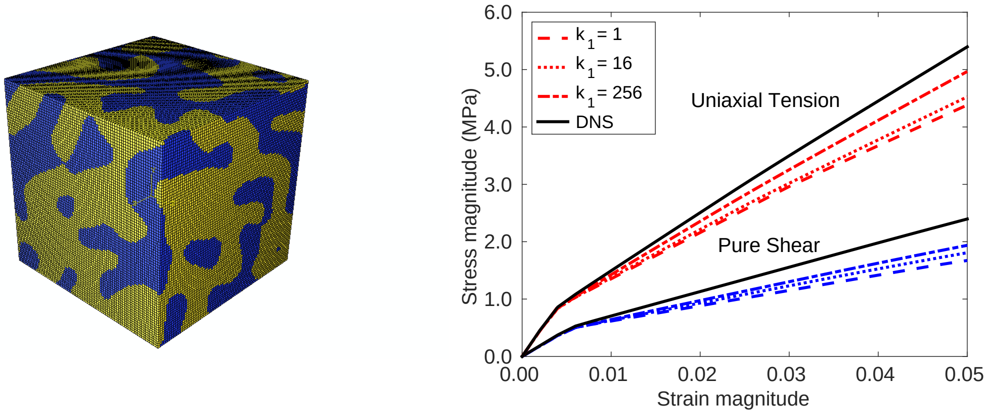


Figure 12: The $80 \times 80 \times 80$ FE mesh of the 3D amorphous material and stress-strain curves under uniaxial tension and pure shear loading conditions. The solid lines represent the DNS results for comparison. Three different numbers of clusters in phase 1 were considered: $k_1 = 1$ (dashes), $k_1 = 16$ (dots) and $k_1 = 256$ (dash-dots).

processor for $k_1 = 16$ and $k_1 = 256$, respectively. A more refined finite element mesh was used for the same microstructure, $80 \times 80 \times 80$, and the differences in the stress-strain curves were less than 1% for both DNS and reduced order methods, indicating good convergence. For this case the DNS takes approximately 25.7 hours on the same 24 cores, while the reduced model computation time is the same (it only depends on the number of clusters).

The final example chosen to present in this work is a 3D two-phase amorphous material obtained by solving the Cahn-Hilliard equation in 3D. The illustration of the 3D microstructure is shown in Figure 12 and the finite element mesh considered for the DNS was $80 \times 80 \times 80$. The volume fraction of each phase was defined as 50%, so $k_2 = k_1$ in the clustering step. The stress-strain curves in Figure 12 show that the predictions given by the reduced order method converge slower than all the previous examples considered: this is attributed to the complexity of microstructure in the 3D amorphous material that leads to fewer similarities in the local stress/strain field, implying the need to use more clusters in order to characterize the nonlinear behavior appropriately.

4. Conclusion

A new computational approach was introduced for modeling the behavior of heterogeneous materials with complex microstructure. The proposed method includes two major contributions: 1) a data-compression procedure to improve the offline stage based on k-means clustering to group material subdomains with similar mechanical behavior; and 2) a new analysis method termed "self-consistent clustering analysis" derived via the homogeniation of each material subdomain (material cluster) through the Lippmann-Schwinger equation. The combination of the data compression achieved at the offline stage and the new scheme for solving the Lippmann-Schwinger equation *without* additional parameters or constitutive laws was shown to lead to a methodology with several attractive features: accurate, good convergence under refinement, computationally efficient, and involving a minimum amount of effort for both the offline and online stages. The offline stage is trivial since it only involves linear elastic simulations under 3 loading conditions for two-dimensional cases, or 6 loading conditions for three-dimensional ones.

As a final note, the proposed method is believed to have a wide range of applications to various material systems. Multifunctional and multiphysics material systems such as piezoelectric and thermomechanical materials may be interesting future applications. The particular case of nonlinear materials with strain softening or damage requires special attention, since the effective modulus becomes negative in the softening region and the current fix-point method for solving the self-consistent reference material needs to be modified to improve convergence. A possible way is to work with the total form of the Lippmann-Schwinger equation instead of the incremental form, i.e. with $\boldsymbol{\sigma}(\mathbf{x})$ and $\boldsymbol{\varepsilon}(\mathbf{x})$ as variables.

Acknowledgement

Z.L., M.A.B and W.K.L. warmly thank the support from AFOSR grant No. FA9550-14-1-0032. Z.L. would like to thank Prof. Wei Chen and Stephen Lin for their parts in helpful discussions. M.A.B. would like to acknowledge support from the Portuguese National Science Foundation and the Fulbright Program, and to thank Prof. Jacob Fish and Dr. Brendan Abberton for their parts in helpful discussions.

Appendix A. Calculation and clustering of A

In order to generate the database with the strain concentration tensors $\mathbf{A}(\mathbf{x})$, we implemented the DNS method based on fast Frourier transforms [19], which iteratively solves the full Lippmann-Schwinger equation with periodic boundary conditions. The finite element method could also be used to calculate $\mathbf{A}(\mathbf{x})$. As mentioned in Section 2.2.1, three orthogonal loading conditions are needed for a 2D material. In Voigt notation $\{\boldsymbol{\varepsilon}^{\text{macro}}\} = [\varepsilon_{11}^{\text{macro}}, \varepsilon_{22}^{\text{macro}}, \gamma_{12}^{\text{macro}}]$, the loading conditions in terms of the macroscopic strain are listed as $[1,0,0]$, $[0,1,0]$ and $[0,0,1]$. For instance, component A_{11}, A_{21} and A_{31} are obtained under condition $\{\boldsymbol{\varepsilon}^{\text{macro}}\} = [1,0,0]$.

In this paper, we use the standard algorithm (Lloyd's algorithm) [48] to solve the k-means clustering problem. The standard algorithm is essentially an optimization process of minimizing the within-cluster least squares sum,

$$\mathbf{S} = \underset{\mathbf{S}'}{\operatorname{argmin}} \sum_{J=1}^{k} \sum_{n \in S^J} ||\mathbf{A}_n - \bar{\mathbf{A}}_J||^2, \tag{A.1}$$

where $\mathbf{A}_n$ is the strain concentration tensor of the n-th data point in Voigt notation, and $\bar{\mathbf{A}}_J$ is the mean of all the strain concentration tensors at the points within the cluster S_J.

At the initialization step, k data points are randomly selected from the data set and served as the initial means (Forgy method [55]). Then the algorithm iterates between the following two steps,

1. Assignment step: Each data point is assigned to the cluster whose mean is nearest to the data point. In other words, within the t-th iteration, $\forall\{\mathbf{A}_n\} \in S_I^{(t)}$, we have

$$||\{\mathbf{A}_n\} - \bar{\mathbf{A}}_I^{(t)}||^2 \leqslant ||\{\mathbf{A}_n\} - \bar{\mathbf{A}}_J^{(t)}||^2 \quad \forall J,\ J \neq I \tag{A.2}$$

2. Update step: The mean values of the data points in the new cluster are recalculated for iteration $t+1$,

$$\bar{\mathbf{A}}_I^{(t)} = \frac{1}{N_I^{(t)}} \sum_{\{\mathbf{A}_n\} \in S_I^{(t)}} \{\mathbf{A}_n\} \tag{A.3}$$

   where $N_i^{(t)}$ is the number of data points in cluster $S_I^{(t)}$.

When the assignment of data points no longer changes, the algorithm is said to converge to a local optimum. However, the global optimum is not guaranteed in the optimization process. Therefore, k-means clustering is always performed with multiple replications in real application in order to overcome the local barrier.

Appendix B. Different elastic properties in the online stage without redoing the offline stage

The proposed approach predicts new effective elastic properties of the RUC when considering new elastic properties for each phase of the material with the same microstructure, even when considering the same selection of elastic properties in the offline stage, equation (45).

Several simulations in the online stage were conducted using different Young's modulus of phase 2 (E_2 was varied from 1 to 5000MPa), while keeping the Young's modulus of phase 1 at 100 MPa. The Young's moduli ratio between phase 2 and phase 1 ranged from 0.01 to 50. The DNS results obtained for this range are provided in Figure B.13, where the microstructural effect on the overall material behavior is clear. Note that since phase 2 in the fiber-reinforced composite is discontinuous (fibers are separated), the effect of varying its modulus E_2 is less pronounced and the overall properties are mainly determined by the matrix material. On the other hand, the overall modulus keeps increasing with E_2 in the amorphous material, where phases 1 and 2 are irregular and similar to each other.

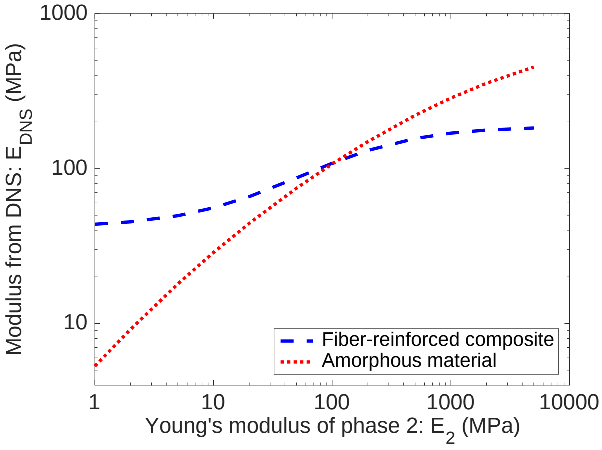


Figure B.13: Equivalent transverse Young modulus E_{DNS} of the RUC obtained from DNS as a function of the Young's modulus of phase 2.

Figure B.14 shows the results from the proposed method with and without the self-consistent scheme. Once again the improvement from the self-consistent scheme is evident, where the transverse Young's modulus is close to the DNS values for a significant range of ratios $(0.01-50)$. Without the self-consistent scheme, especially for the amorphous material, there is a tendency to overestimate or underestimate the modulus at extreme cases even using a large number of clusters, indicating that phase 1 is no longer a good candidate for the reference material.

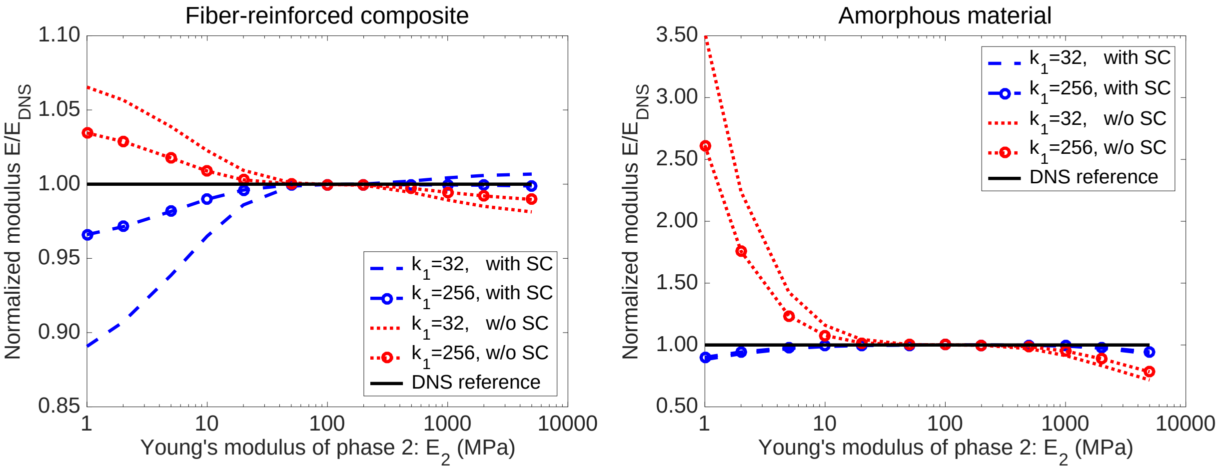


Figure B.14: The effects of phase 2 Young's modulus E_2 on the overall normalized equivalent transverse Young's modulus for the fiber-reinforced material (left) and amorphous material (right). Results from $k_1 = 32$ and 256 are provided with and without the self-consistent scheme (SC).

These results demonstrate that reasonable elastic predictions are possible even without going through the offline process for new combinations of elastic properties. Nevertheless, in practice the offline stage of the proposed method is so simple (three loading conditions in 2D, and six in 3D) that avoiding it is usually not worthwhile.

References

[1] A. L. Gurson, Continuum theory of ductile rupture by void nucleation and growth: Part iyield criteria and flow rules for porous ductile media, Journal of engineering materials and technology 99 (1) (1977) 2–15.

[2] E. de Souza Neto, D. Peric, D. Owen, Computational methods for plasticity: theory and applications, Wiley, 2008.

[3] P. Camanho, M. Bessa, G. Catalanotti, M. Vogler, R. Rolfes, Modeling the inelastic deformation and fracture of polymer composites part ii: Smeared crack model, Mechanics of Materials 59 (0) (2013) 36 – 49.

[4] F. Feyel, J.-L. Chaboche, Fe 2 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.

[5] V. Kouznetsova, M. G. D. Geers, W. A. M. Brekelmans, Multi-scale constitutive modelling of heterogeneous materials with a gradient-enhanced computational homogenization scheme, International Journal for Numerical Methods in Engineering 54 (8) (2002) 1235–1260.

[6] F. Feyel, A multilevel finite element method (fe2) to describe the response of highly non-linear structures using generalized continua, Computer Methods in Applied Mechanics and Engineering 192 (2830) (2003) 3233 – 3244, ¡ce:title¿Multiscale Computational Mechanics for Materials and Structures¡/ce:title¿.

[7] G. J. Wagner, W. K. Liu, Coupling of atomistic and continuum simulations using a bridging scale decomposition, Journal of Computational Physics 190 (1) (2003) 249 – 274.

[8] H. Kadowaki, W. K. Liu, Bridging multi-scale method for localization problems, Computer Methods in Applied Mechanics and Engineering 193 (3032) (2004) 3267 – 3302, computational Failure Mechanics.

[9] H. Kadowaki, W. K. Liu, A multiscale approach for the micropolar continuum model, Computer Modeling in Engineering and Sciences 7 (3) (2005) 269–282.

[10] H. S. Park, E. G. Karpov, P. A. Klein, W. K. Liu, Three-dimensional bridging scale analysis of dynamic fracture, Journal of Computational Physics 207 (2) (2005) 588 – 609.

[11] S. Tang, T. Y. Hou, W. K. Liu, A pseudo-spectral multiscale method: Interfacial conditions and coarse grid equations, Journal of Computational Physics 213 (1) (2006) 57 – 85.

[12] X. Bai, M. Bessa, A. Melro, P. Camanho, L. Guo, W. Liu, High-fidelity micro-scale modeling of the thermo-visco-plastic behavior of carbon fiber polymer matrix composites, Composite Structures (2015) –.

[13] N. Vu-Bac, M. Bessa, T. Rabczuk, W. Liu, A multiscale model for the quasi-static thermo-plastic behavior of highly cross-linked glassy polymers, Submitted to publication.

[14] J. D. Eshelby, The determination of the elastic field of an ellipsoidal inclusion, and related problems, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, Vol. 241, The Royal Society, 1957, pp. 376–396.

[15] Z. Hashin, S. Shtrikman, A variational approach to the theory of the elastic behaviour of multiphase materials, Journal of the Mechanics and Physics of Solids 11 (2) (1963) 127–140.

[16] R. Hill, A self-consistent mechanics of composite materials, Journal of the Mechanics and Physics of Solids 13 (4) (1965) 213–222.

[17] Z. Liu, J. A. Moore, S. M. Aldousari, H. S. Hedia, S. A. Asiri, W. K. Liu, A statistical descriptor based volume-integral micromechanics model of heterogeneous material with arbitrary inclusion shape, Computational Mechanics (2015) 1–19.

[18] S. Ghosh, K. Lee, S. Moorthy, Two scale analysis of heterogeneous elastic-plastic materials with asymptotic homogenization and voronoi cell finite element model, Computer Methods in Applied Mechanics and Engineering 132 (12) (1996) 63 – 116.

[19] H. Moulinec, P. Suquet, A numerical method for computing the overall response of nonlinear composites with complex microstructure, Computer Methods in Applied Mechanics and Engineering 157 (12) (1998) 69 – 94.

[20] J. Aboudi, A continuum theory for fiber-reinforced elastic-viscoplastic composites, International Journal of Engineering Science 20 (5) (1982) 605–621.

[21] M. Paley, J. Aboudi, Micromechanical analysis of composites by the generalized cells model, Mechanics of Materials 14 (2) (1992) 127–139.

[22] G. J. Dvorak, Transformation field analysis of inelastic composite materials 437 (1900) (1992) 311–327.

[23] J. Michel, P. Suquet, Nonuniform transformation field analysis, International Journal of Solids and Structures 40 (25) (2003) 6937 – 6955, special issue in Honor of George J. Dvorak.

[24] S. Roussette, J.-C. Michel, P. Suquet, Nonuniform transformation field analysis of elastic–viscoplastic composites, Composites Science and Technology 69 (1) (2009) 22–27.

[25] K. Karhunen, Zur spektraltheorie stochastischer prozesse, Suomalainen tiedeakatemia, 1946.

[26] M. Loève, Probability Theory; Foundations, Random Sequences, New York: D. Van Nostrand Company, 1955.

[27] I. Jolliffe, Principal component analysis, Wiley Online Library, 2002.

[28] G. Berkooz, P. Holmes, J. L. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, Annual review of fluid mechanics 25 (1) (1993) 539–575.

[29] J. Yvonnet, Q.-C. He, The reduced model multiscale method (r3m) for the non-linear homogenization of hyperelastic media at finite strains, Journal of Computational Physics 223 (1) (2007) 341–368.

[30] O. Goury, D. Amsallem, S. P. A. Bordas, W. K. Liu, P. Kerfriden, Automatised selection of load paths to construct reduced-order models in computational damage micromechanics: from dissipation-driven random selection to bayesian optimisation, Computational Mechanics.

[31] D. Néron, P.-A. Boucard, N. Relun, Time-space pgd for the rapid solution of 3d nonlinear parametrized problems in the many-query context, International Journal for Numerical Methods in Engineering 103 (4) (2015) 275–292.

[32] P. Ladevèze, Nonlinear computational structural mechanics: new approaches and non-incremental methods of calculation, Springer Science & Business Media, 2012.

[33] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids, Journal of Non-Newtonian Fluid Mechanics 139 (3) (2006) 153 – 176.

[34] P. Ladevèze, J.-C. Passieux, D. Néron, The {LATIN} multiscale computational method and the proper generalized decomposition, Computer Methods in Applied Mechanics and Engineering 199 (2122) (2010) 1287 – 1296, multiscale Models and Mathematical Aspects in Solid and Fluid Mechanics.

[35] F. Chinesta, A. Ammar, A. Leygue, R. Keunings, An overview of the proper generalized decomposition with applications in computational rheology, Journal of Non-Newtonian Fluid Mechanics 166 (11) (2011) 578 – 592.

[36] C. Ghnatios, F. Masson, A. Huerta, A. Leygue, E. Cueto, F. Chinesta, Proper generalized decomposition based dynamic data-driven control of thermal processes, Computer Methods in Applied Mechanics and Engineering 213216 (2012) 29 – 41.

[37] P. Kerfriden, O. Goury, T. Rabczuk, S. P.-A. Bordas, A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics, Computer methods in applied mechanics and engineering 256 (2013) 169–188.

[38] C. Oskay, J. Fish, Eigendeformation-based reduced order homogenization for failure analysis of heterogeneous materials, Computer Methods in Applied Mechanics and Engineering 196 (7) (2007) 1216 – 1243.

[39] J. Chaboche, P. Kanout, A. Roos, On the capabilities of mean-field approaches for the description of plasticity in metal matrix composites, International Journal of Plasticity 21 (7) (2005) 1409 – 1434.

[40] J. Fish, Practical multiscaling, John Wiley & Sons, 2013.

[41] F. Fritzen, M. Leuschner, Reduced basis hybrid computational homogenization based on a mixed incremental formulation, Computer Methods in Applied Mechanics and Engineering 260 (2013) 143–154.

[42] Z. Hashin, S. Shtrikman, A variational approach to the theory of the elastic behaviour of polycrystals, Journal of the Mechanics and Physics of Solids 10 (4) (1962) 343–352.

[43] J. MacQueen, et al., Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 1, Oakland, CA, USA., 1967, pp. 281–297.

[44] J. W. Cahn, J. E. Hilliard, Free energy of a nonuniform system. i. interfacial free energy, The Journal of chemical physics 28 (2) (1958) 258–267.

[45] L. Barghout, J. Sheynin, Real-world scene perception and perceptual organization: Lessons from computer vision, Journal of Vision 13 (9) (2013) 709–709.

[46] W. Li, L. Jaroszewski, A. Godzik, Clustering of highly homologous sequences to reduce the size of large protein databases, Bioinformatics 17 (3) (2001) 282–283.

[47] I. H. Witten, E. Frank, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2005.

[48] S. P. Lloyd, Least squares quantization in pcm, Information Theory, IEEE Transactions on 28 (2) (1982) 129–137.

[49] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, A. Y. Wu, An efficient k-means clustering algorithm: Analysis and implementation, Pattern Analysis and Machine Intelligence, IEEE Transactions on 24 (7) (2002) 881–892.

[50] P. J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of computational and applied mathematics 20 (1987) 53–65.

[51] J. C. Michel, H. Moulinec, P. Suquet, A computational scheme for linear and non-linear composites with arbitrary phase contrast, International Journal for Numerical Methods in Engineering 52 (1-2) (2001) 139–160.

[52] V. Monchiet, G. Bonnet, A polarization-based fft iterative scheme for computing the effective properties of elastic composites with arbitrary contrast, International Journal for Numerical Methods in Engineering 89 (11) (2012) 1419–1436.

[53] J. Yvonnet, A fast method for solving microstructural problems defined by digital images: a space lippmann–schwinger scheme, International Journal for Numerical Methods in Engineering 92 (2) (2012) 178–205.

[54] C. M. Breneman, L. C. Brinson, L. S. Schadler, B. Natarajan, M. Krein, K. Wu, L. Morkowchuk, Y. Li, H. Deng, H. Xu, Stalking the materials genome: A data-driven approach to the virtual design of nanostructured polymers, Advanced Functional Materials 23 (46) (2013) 5746–5752.

[55] E. W. Forgy, Cluster analysis of multivariate data: efficiency versus interpretability of classifications, Biometrics 21 (1965) 768–769.