A k-means clustering machine learning-based multiscale method for anelastic heterogeneous structures with internal variables

Mohamed Amine Benaimeche, Julien Yvonnet, Benoît Bary, Qi-Chang He

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

A k-means clustering machine learning-based multiscale method for anelastic heterogeneous structures with internal variables

M.A. Benaimeche[1,2] | J. Yvonnet*[2] | B. Bary[1] | Q.-C. He[2]

[1]Université Paris-Saclay, CEA, Service d'Étude du Comportement des Radionucléides, 91191, Gif-sur-Yvette, France

[2]MSME, Univ Gustave Eiffel, CNRS UMR 8208, F-77454, France

Correspondence
*Corresponding Julien Yvonnet, Univ Gustave Eiffel, 5 BD Descartes 77454 Marne-la-Vallée Cedex 2, France. Email: julien.yvonnet@univ-eiffel.fr

Abstract

A new machine-learning based multiscale method, called k-means FE^2, is introduced to solve general nonlinear multiscale problems with internal variables and loading history-dependent behaviors, without use of surrogate models. The macro scale problem is reduced by constructing clusters of Gauss points in a structure which are estimated to be in the same mechanical state. A k-means clustering - machine learning technique is employed to select the Gauss points based on their strain state and sets of internal variables. Then, for all Gauss points in a cluster, only one micro nonlinear problem is solved, and its response is transferred to all integration points of the cluster in terms of mechanical properties. The solution converges with respect to the number of clusters, which is weakly depends on the number of macro mesh elements. Accelerations of FE^2 calculations up to a factor 50 are observed in typical applications. Arbitrary nonlinear behaviors including internal variables can be considered at the micro level. The method is applied to heterogeneous structures with local quasi-brittle and elastoplastic behaviors and, in particular, to a nuclear waste package structure subject to internal expansions.

KEYWORDS:
Multiscale; FE2; k-means clustering; Nonlinear; Homogenization; Machine learning

1 | INTRODUCTION

Machine learning, Artificial Intelligence and Data Sciences have recently attracted a lot of attention in the computational mechanics community. These techniques may offer new possibilities like reducing computational times, automating construction of models from data, or automating mappings between models and material characteristics or microstructural patterns.

One first example of use of machine learning in mechanics is the construction of models from data. For example, Artificial Neural Networks (ANN) have shown promising results to reproduce stress-strain curves of behaviors in complex configurations like high temperatures loadings[1], design of steel structures[2], vibrational behavior of structures[3,4], or structural stability problems[5].

Another application importing techniques from machine learning is the recognition of patterns and the construction of a relationship between material topological features and some models. For examples in[6,7], machine learning was used to relate microstructures to their effective thermal or mechanical effective properties. In[8] Ryckelynck et al. used image recognition and non supervised machine learning techniques for classification issues and the construction of digital twins of material defects in nonlinear mechanical simulations.

More recently, Ortiz et al.[9] developed a fully data-driven approach to replace constitutive law in finite elements simulations by direct use of data, by searching in a data base the closest data point to be used at the local Gauss points, then removing the use of empirical constitutive laws. In[10], Carrara et al. constructed fracture models where modeling assumptions were removed and the governing equations stemming from variational principle were combined with discrete data points, leading to a model-free data-driven method for solving fracture problems.

Strategies using machine learning have also recently been developed to describe complex phenomena like dynamics trajectories or fluid sloshing without solving the partial differential equations[11,12].

Another exciting field of application of these techniques is multiscale modeling. One major difficulty in multiscale numerical simulation occurs when dealing with nonlinear materials. One direct approach to simulate the behavior of a nonlinear heterogeneous structure resorts to the so-called FE^2 technique[13,14], where two levels of Finite Element simulations are conducted, one at the scale of the whole structure, and the other at the microscale where nonlinear simulations are performed over Representative Volume Elements. The major drawback of such an approach is its computational cost, related to the unacceptable number of nonlinear calculations to be performed. Among the different strategies to reduce these calculations, one method, initiated in[15,16], makes use of so-called *data-driven* approaches in which micro scale calculations are performed during an off-line stage, and which are then used as data in an *on-line* stage to reconstruct the macroscopic (effective) behavior. For this purpose, several techniques have been proposed, including interpolation methods[15,17], Neural Networks[16,18,19,20,21,22,23], Bayesian inference[24], Fourier series expansions[25], Gaussian process[26], or self-clustering analysis[27,28]. In the related techniques, the off-line data collection is used in a regression process to construct an accurate surrogate model whose evaluation is of several orders of magnitude lower than performing one micro scale nonlinear calculation. A critical comparison of several regression techniques used in data-driven multiscale approaches can be found in[29]. In[30], Avery et al. investigated and discussed several regression methods with ANN in homogenization problems of hyperelastic woven composites, and demonstrated its use in advanced dynamic/fluid structure applications. Recent advances of data driven techniques, including handling loading history-dependent behaviors like plasticity can be found in[31,11,23]. On-the fly construction of the surrogate model by probabilistic machine learning was proposed in[26]. Developments of neural networks techniques in FE^2, including Feed-Forward and recurrent Neural Networks can be found in[19,31]. In[32,33], both homogenization and localization of heterogeneous hyperelastic materials were performed using a digital database and the manifold-based nonlinear reduced order model (MNROM).

However, the main challenge in the aforementioned techniques is that they are adapted only to some classes of constitutive behaviors and lack generality. More specifically, handling internal variables and loading-dependent behaviors still remains delicate. In addition, neural networks and related surrogate model techniques lack error estimation, especially when the parameters are out of the training parameter data set[34].

The objective of this paper is then to propose a new machine-learning based multiscale method, called k-means FE^2, to solve general nonlinear multiscale problems with internal variables and loading-dependent behaviors, without use of surrogate models. For this purpose, we employ a new paradigm: instead of reducing the microscale problem calculations, we reduce the macro scale problem by constructing clusters of Gauss points in a structure which are assumed to be in the same mechanical state. A k-means clustering - machine learning technique is employed to select the Gauss points based on their strain state and sets of internal variable. Then, for all Gauss points in a cluster, only one micro nonlinear problem is solved, and its response is transferred to all points of the cluster. The operation is repeated until convergence is reached at both macro and micro scales. Drastic computational reductions can be achieved, and error can be controlled by gradually increasing the number of clusters at the macro scale. Note that the present method is totally different from the so-called *Self-Clustering analysis* proposed by Liu et al.[27,28], which involves clustering algorithms at the micro scale, and which uses a Lippmann-Schwinger integral equation to accelerate micro equations. We summarize the main features of the proposed technique as follows:

- Convergence is obtained with respect to the number of clusters.
- Validation is made with respect to direct FE^2 solutions.
- Convergence is given with respect to the macro mesh size.
- Computational gains between 1 and 2 orders of magnitude can be achieved for similar accuracy when compared to full FE^2 solutions. Larger gains can be obtained if larger errors are accepted.
- A simple error analysis can be achieved by gradually increasing the number of clusters.
- Nonlinear behaviors can be arbitrary (plasticity, damage, etc.) with internal variables.

- The number of clusters weakly depends on the one of macro mesh elements.

The organization of the paper is as follows. In section 2, the basics of $k-means$ clustering analysis are recalled. In section 3, the k-means FE^2 method is presented. In section 4, numerical examples are provided to illustrate the aforementioned features of the method and its potential use for industrial applications.

2 | K-MEANS CLUSTERING

In this section, we briefly recall the basics of the k-means clustering method[35]. This algorithm is used to classify data into groups, or clusters, with similar features. Clustering methods are widely used in data mining[36], data compression[37], and pattern classification[38]. Given a set of N vectors

$$\mathcal{S} = \left\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_N\right\}, \tag{1}$$

the k-means algorithms aims at defining K clusters $\mathcal{S}^k$, $k = 1, 2, ..., K$:

$$\mathcal{S}^k = \left\{\mathbf{v}_1^k, \mathbf{v}_2^k, ..., \mathbf{v}_{N_k}^k\right\}, \tag{2}$$

each containing N_k vectors $\mathbf{v}_i^k$, where N_k can be different for each cluster k. We define the centroid of a cluster according to

$$\overline{\mathbf{v}}^k = \frac{1}{N_k}\sum_{i\in\mathcal{S}^k}\mathbf{v}_i^k \tag{3}$$

and the Euclidean distance between one data point (vector) $\mathbf{v}_j$ and the centroid of a cluster k by

$$d_j^k = \|\mathbf{v}_j - \overline{\mathbf{v}}^k\|. \tag{4}$$

where $\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}}$. Mathematically, the objective of the k-means algorithm is to find the sets $\mathcal{S}^k$ such that the following objective function

$$J = \sum_k\sum_{i\in\mathcal{S}^k}\|\mathbf{v}_i - \overline{\mathbf{v}}^k\|^2 \tag{5}$$

is minimized. The number K is chosen by the user. The centroids are initialized by choosing K arbitrary vectors $\mathbf{v}_m^k \in \mathcal{S}$, and assigning $\overline{\mathbf{v}}^k = \mathbf{v}_m^k$ at the first iteration. Then, the algorithm runs as follows[39]. At each iteration r, for each data point $\mathbf{v}_i \in \mathcal{S}$, the Euclidean distance to the centroid of each cluster is computed according to (4). A data point (vector) $\mathbf{v}_i$ is simply assigned to a cluster k if its distance to the cluster is minimal among all clusters

$$\mathcal{S}^k = \left\{\mathbf{v}_i |\ \ \|\mathbf{v}_i - \overline{\mathbf{v}}^k\| \leq \|\mathbf{v}_i - \overline{\mathbf{v}}^m\|, \forall m \text{ with } m \neq k\right\}. \tag{6}$$

The process goes on until a criterion is reached. Here, we choose the simple stopping criterion:

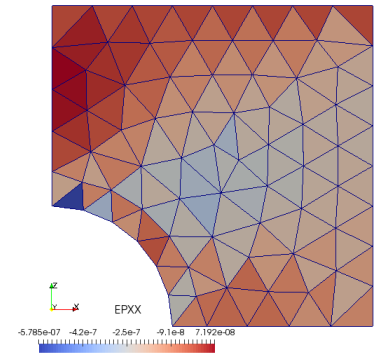
$$\Delta J = |J^{(r)} - J^{(r-1)}| = 0 \tag{7}$$

where $J^{(r)}$ is given by (5) at iteration r of the algorithm. The k-means algorithm is summarized in Algorithm 1.

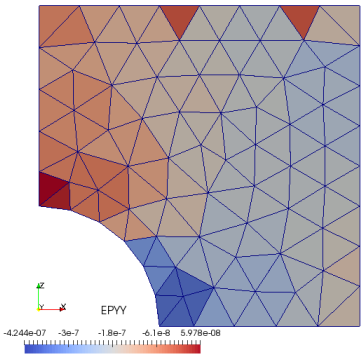
An illustration in the context of a 3D linear finite element calculation is provided in Fig. 1. Here, the algorithm is applied to classify the Gauss points in the finite elements mesh according to their strain components values. A perforated plate under uniaxial tension is considered, whose dimensions, boundary conditions, and loading are described in section 4.2. Figs. 1a, 1b, 1c show the distributions of the strain components ε_{11}, ε_{22} and ε_{12} in the elements of the mesh. Here, as one Gauss point per element is used, the elements are associated to the Gauss points. After solving the FEM problem, the strain components are obtained in each element in a vector form as:

$$\mathbf{v}_i^T = \left[\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, 2\varepsilon_{13}, 2\varepsilon_{23}, 2\varepsilon_{12}\right]. \tag{8}$$

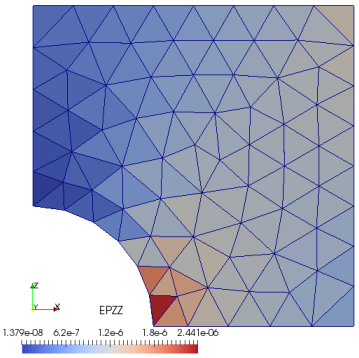
Here the number of clusters is chosen as $k = 5$. Fig. 1d depicts the initial data points (elements) selected to initialize the clusters. Figs. 1e-1l show the positions of the clusters at different iterations. In this example, the convergence is achieved in 25 iterations. Note that the convergence rate of the algorithm depends on the number of clusters k and on the number of vectors N in $\mathcal{S}$ (here the number of elements in the FEM mesh).



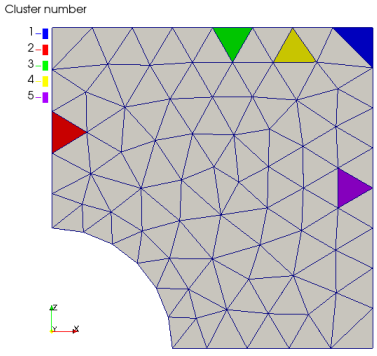
(a) ϵ_{11}



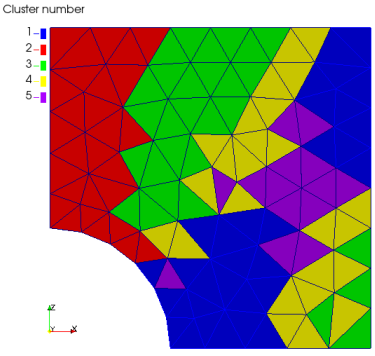
(b) ϵ_{22}



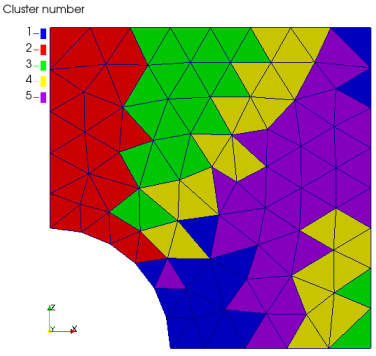
(c) ϵ_{33}



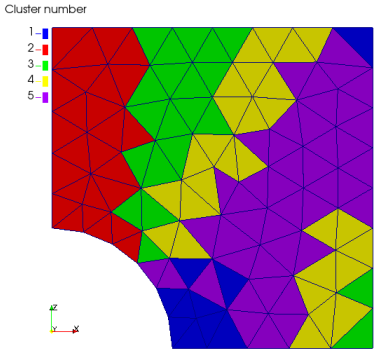
(d) iteration 0



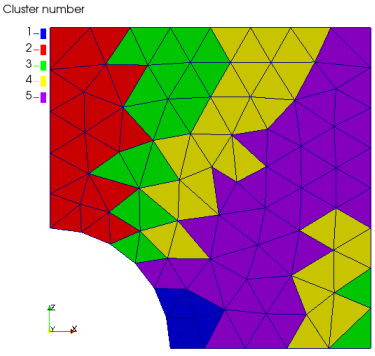
(e) iteration 1



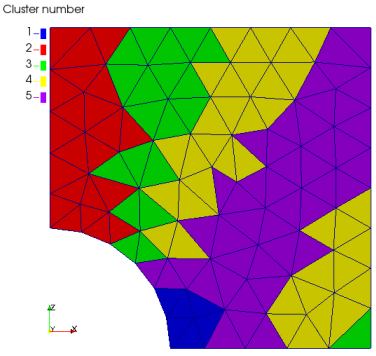
(f) iteration 2



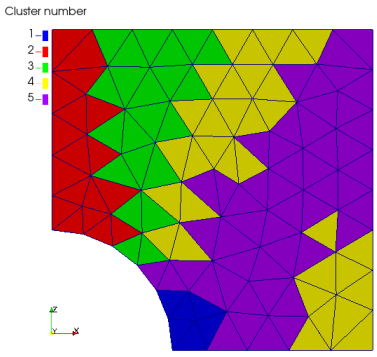
(g) iteration 3



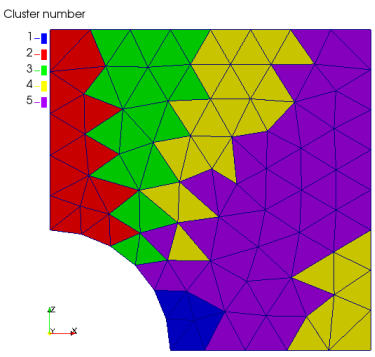
(h) iteration 5



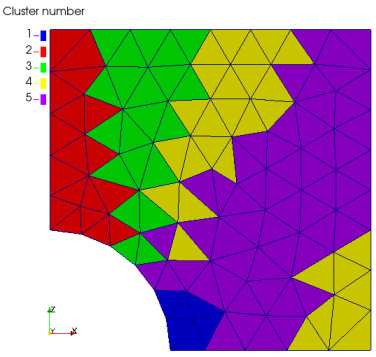
(i) iteration 9



(j) iteration 15



(k) iteration 20



(l) iteration 25

FIGURE 1 Application of the k-means algorithm for classifying the elements in a FEM linear calculation according to their strain components: (a), (b), (c): components of the strain tensor; (d): initialization of the clusters by choosing arbitrary elements in the structure; (e)-(l): positions of the clusters at iterations 1, 2,3,5,9,15,20 and 25 respectively.

Algorithm 1 k-means clustering

1: Initialize $\overline{\mathbf{v}}^k = \mathbf{v}_m^k$, where $\mathbf{v}_m^k$ are randomly chosen vectors in $\mathcal{S}$.
2: Set the number of clusters K.
3: WHILE $\Delta J^{(R)} \neq 0$
4: Assignment phase: assign each data point to the cluster $\mathcal{S}^k$ according to the nearest cluster centroid criterion.
$$\mathcal{S}^k = \left\{ \mathbf{v}_i |\ \|\mathbf{v}_i - \overline{\mathbf{v}}^k\| \leq \|\mathbf{v}_i - \overline{\mathbf{v}}^m\|, \forall m,\ m \neq k \right\}.$$
5: Update centroids:
$$\overline{\mathbf{v}}^k = \frac{1}{N_k} \sum_{i \in \mathcal{S}^k} \mathbf{v}_i^k$$
6: Test convergence:
$$J^{(r)} = \sum_k \sum_{i \in \mathcal{S}^{k,(r)}} \|\mathbf{v}_i - \overline{\mathbf{v}}^{k,(r)}\|^2$$
$$\Delta J = |J^{(r)} - J^{(r-1)}|$$
7: END

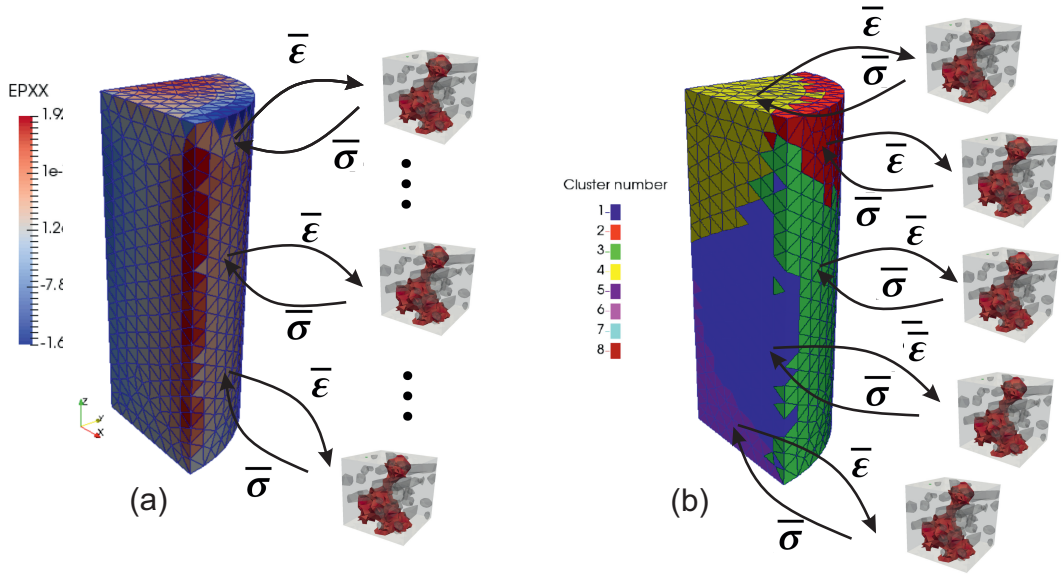


FIGURE 2 (a) Classical FE^2 method: one nonlinear RVE calculation is required at each Gauss point of the macro mesh; (b) k-means FE^2: only one RVE calculation is required for each cluster (depicted in different colors).

3 | A K-MEANS FE^2

3.1 | Brief review of the FE^2 method

The multilevel finite-element method [13,14], also called FE^2 in the literature, as it involves two levels of finite-element simulations, and independently proposed by several other authors and groups [40,41,42,43,44,45], was introduced as a general multiscale method for solving nonlinear heterogeneous structural problems. The basic underlying idea is that two levels of finite element simulations are concurrently solved, one for each scale. At the macroscale, each integration point of the finite-element mesh is associated with a representative volume element (RVE). Boundary conditions depending on the macroscopic strain are prescribed on the boundary of each RVE. After solving each nonlinear RVE problem at each integration point, the macroscopic stress is averaged over the RVE and provided at the macro integration point. Then, the macroscopic constitutive law is available only through solving a nonlinear problem. These operations are repeated until convergence is reached at both scales (see Figure 2 (a)).

We consider a macroscopic structure associated with a domain $\overline{\Omega} \subset \mathbb{R}^3$, with a boundary $\partial\overline{\Omega}$. The assumption of scale separation is adopted (an extension of FE^2 to second-order homogenization can be found in [43]). The microstructure is assumed to be characterized by an RVE associated with a domain $\Omega \subset \mathbb{R}^3$, with boundary $\partial\Omega$. For a general nonlinear constitutive law, the strain $\boldsymbol{\epsilon}(\mathbf{x})$ is related to the stress $\boldsymbol{\sigma}(\mathbf{x})$ by a nonlinear local constitutive relationship, where $\mathbf{x}$ is a material point within Ω.

In the following, $\overline{(.)}$ notations will denote macroscale quantities. In the absence of body forces, for a given macroscopic strain $\overline{\boldsymbol{\varepsilon}}$ and a set of microscopic internal variables defined by a vector $\boldsymbol{\alpha}(\mathbf{x})$, the RVE problem is to find $\varepsilon(\mathbf{x})$, such that

$$\nabla \cdot \boldsymbol{\sigma}(\mathbf{x}) = 0 \; \forall \mathbf{x} \in \Omega, \tag{9}$$

where $\nabla \cdot (.)$ is the divergence operator. The constitutive law is given by

$$\boldsymbol{\sigma}(\mathbf{x}) = \mathcal{F}^{nl}\left(\boldsymbol{\varepsilon}(\mathbf{x}), \boldsymbol{\alpha}(\mathbf{x})\right). \tag{10}$$

where $\mathcal{F}^{nl}$ is a local nonlinear operator. Under the assumption of sufficient displacement smoothness, the strain macroscopic strain is defined by

$$\overline{\boldsymbol{\varepsilon}} = \frac{1}{V}\int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{x}) d\Omega, \tag{11}$$

where V is the volume of Ω. Equation (11) holds, e.g. for the following boundary conditions (see e.g.[46]):

$$\mathbf{u}(\mathbf{x}) = \overline{\boldsymbol{\varepsilon}}\mathbf{x} + \tilde{\mathbf{u}}(\mathbf{x}) \;\text{ on } \partial\Omega, \tag{12}$$

where $\tilde{\mathbf{u}}(\mathbf{x})$ is a periodic function over Ω. Again upon a sufficient stress regularity, the macroscopic stress is defined by

$$\overline{\boldsymbol{\sigma}} = \frac{1}{V}\int_{\Omega} \boldsymbol{\sigma}(\mathbf{x}) d\Omega. \tag{13}$$

In the so-called FE^2 method, the constitutive law $\overline{\boldsymbol{\sigma}}$ - $\overline{\boldsymbol{\varepsilon}}$ is unknown, but can be obtained numerically by solving a nonlinear problem over the RVE, as follows (see Figure 2 (a)):

Given $\overline{\boldsymbol{\varepsilon}}, \boldsymbol{\alpha}(\mathbf{x})$:

1. Prescribe boundary conditions (12) on $\partial\Omega$.
2. Use a numerical method such as FEM with an iterative solver such as the Newton method to solve the nonlinear problems (9), (10), and (12) (see details in the following).
3. Compute the spatial average of the stress over the RVE to obtain $\overline{\boldsymbol{\sigma}}$.

In what follows, a detailed numerical implementation of the FE^2 method is presented to better understand where problems (9), (10), and (12) have to be solved within finite-element calculation at the macroscopic scale. In the absence of body forces, the problem at the macroscale is given by

$$\nabla \cdot \overline{\boldsymbol{\sigma}} = 0 \;\text{ in } \Omega, \tag{14}$$

and completed with boundary conditions such as prescribed tractions $\overline{\mathbf{F}}^{*}$ over $\partial\overline{\Omega}_F$ and prescribed displacemnts $\overline{\mathbf{u}} = \overline{\mathbf{u}}^{*}$ over $\partial\overline{\Omega}_u$. The weak form associated with (14) is given by:

$$\int_{\overline{\Omega}} \overline{\boldsymbol{\sigma}}\left(\overline{\boldsymbol{\varepsilon}}, \boldsymbol{\alpha}\right) : \overline{\boldsymbol{\varepsilon}}(\delta\overline{\mathbf{u}}) d\Omega - \int_{\partial\overline{\Omega}_F} \overline{\mathbf{F}}^{*} \cdot \delta\overline{\mathbf{u}} d\Gamma = \overline{R}(\overline{\mathbf{u}}) = \mathbf{0}. \tag{15}$$

In a nonlinear context, Eq. (15) must be solved using an iterative method, e.g. a Newton method, as described in the following. The linearized problem related to (15) is given by:

$$\int_{\overline{\Omega}} \overline{\mathbb{C}}_{tan}\left(\overline{\boldsymbol{\varepsilon}}\left(\overline{\mathbf{u}}^{p}\right), \boldsymbol{\alpha}^{p}\right) : \overline{\boldsymbol{\varepsilon}}\left(\Delta\overline{\mathbf{u}}\right) : \overline{\boldsymbol{\varepsilon}}\left(\delta\overline{\mathbf{u}}\right) d\Omega = -\int_{\overline{\Omega}} \overline{\boldsymbol{\sigma}}\left(\overline{\boldsymbol{\varepsilon}}\left(\overline{\mathbf{u}}^{p}\right), \boldsymbol{\alpha}^{p}\right) : \overline{\boldsymbol{\varepsilon}}(\delta\overline{\mathbf{u}}) d\Omega + \int_{\partial\overline{\Omega}_F} \overline{\mathbf{F}}^{*} \cdot \delta\overline{\mathbf{u}} d\Gamma, \tag{16}$$

where $\overline{\mathbf{u}}^{p}$ and $\boldsymbol{\alpha}^{p}$ denote the displacement solution and the set of internal variables at the previous $p-$th iteration, respectively, and $\overline{\mathbb{C}}_{tan} = \frac{\partial\overline{\sigma}}{\partial\overline{\varepsilon}}$ is the tangent elastic tensor. Solving (16), we obtain the correction to be carried out on the macroscopic displacements:

$$\overline{\mathbf{u}}^{p+1} = \overline{\mathbf{u}}^{p} + \Delta\overline{\mathbf{u}}. \tag{17}$$

Problem (16) is solved until a convergence criterion is reached, i.e. when the norm of the residual $\overline{\mathbf{R}}$ is lower than a prescribed tolerance. Introducing an FEM discretization into (16), we obtain a linear system of equations in the form:

$$\overline{\mathbf{K}}_{tan}\Delta\overline{\mathbf{U}} = -\overline{\mathbf{R}} \tag{18}$$

where $\Delta\overline{\mathbf{U}}$ is the vector containing all nodal corrections of displacement components in the macroscopic mesh,

$$\overline{\mathbf{K}}_{tan} = \int_{\overline{\Omega}} \overline{\mathbf{B}}^T \overline{\mathbf{C}}_{tan}\left(\overline{\mathbf{u}}^p\right)\overline{\mathbf{B}}d\Omega \tag{19}$$

and

$$\overline{\mathbf{R}} = -\int_{\overline{\Omega}} \overline{\mathbf{B}}^T\left[\overline{\boldsymbol{\sigma}}\left(\overline{\mathbf{u}}^p\right)\right]d\Omega + \int_{\partial\overline{\Omega}_F} \mathbf{N}^T\overline{\mathbf{F}}^* d\Omega, \tag{20}$$

where $\overline{\mathbf{C}}_{tan}$ is the matrix form of $\overline{\mathbb{C}}_{tan}$, $\left[\overline{\boldsymbol{\sigma}}\right]$ is the vector form of $\overline{\boldsymbol{\sigma}}$, and $\overline{\mathbf{B}}$ and $\overline{\mathbf{N}}$ are matrices of FEM shape functions and shape functions derivatives related to the macroscopic mesh. The components of $\overline{\mathbb{C}}_{tan}$ can be evaluated numerically by perturbation using:

$$\left(\overline{\mathbb{C}}_{tan}\right)_{ijkl}(\overline{\boldsymbol{\varepsilon}}) = \frac{\partial\overline{\sigma}_{ij}}{\partial\overline{\varepsilon}_{kl}}(\overline{\boldsymbol{\varepsilon}}) \simeq \frac{\overline{\sigma}_{ij}\left(\overline{\boldsymbol{\varepsilon}} + \boldsymbol{\delta}^{(kl)}\right) - \overline{\sigma}_{ij}\left(\overline{\boldsymbol{\varepsilon}}\right)}{\theta} \tag{21}$$

with

$$\boldsymbol{\delta}^{(kl)} = \frac{\theta}{2}\left(\mathbf{e}_k \otimes \mathbf{e}_l + \mathbf{e}_l \otimes \mathbf{e}_k\right) \tag{22}$$

where θ is a small parameter such that $\theta << \left|\overline{\boldsymbol{\varepsilon}}\right|$ and $\mathbf{e}_k,\ k = 1, 2, 3$ are unitary basis vectors. We can see that at each point of the macroscopic mesh, determining $\overline{\boldsymbol{\sigma}}$ requires solving one nonlinear problem over the RVE while determining $\overline{\mathbf{C}}_{tan}$ requires solving 6 additional calculations in 3D. We can provide an estimation of the total number of RVE calculations to be performed within a whole FE^2 simulation as follows:

$$N_{tot} \simeq (1+\beta) \times N_{int} \times N_e \times N_{iter} \times N_{evol}, \tag{23}$$

where $\beta = 3$ in 2D and $\beta = 6$ in 3D, N_{int} is the number of integration points per element in the macroscopic mesh, N_e is the number of elements in the macroscopic mesh, N_{iter} is the average number of macroscopic Newton iterations before convergence at each time step, and N_{evol} is the number of time steps for the whole simulation (loading steps in a quasi-static simulation). As an illustration, considering a 3D mesh containing 100,000 elements (which is a rather coarse mesh for practical applications), $N_{int} = 1$, $N_{iter} = 4$ and 10 time steps, $N_{evol} = 10$, we obtain $N_{tot} = 28{\times}10^6$ nonlinear calculations on the RVE during the whole simulation. This simple example shows that for most applications, the classical FE^2 method is not applicable, unless very coarse meshes at both micro and macro levels are used. To drastically reduce this complexity, the k-means FE^2 method is proposed in the next section.

3.2 | k-means FE^2

One bottleneck in the above FE^2 procedure is the total number of nonlinear problems associated with the number $N_{int} \times N_e$ of macroscopic integration points. In practical calculations, several RVEs are subjected to very close macroscopic strains $\overline{\boldsymbol{\varepsilon}}$. During a typical multiscale simulation, many redundant problems are then solved unnecessarily. Then, the idea is to identify clusters of integration points in the macro mesh according to the k-means algorithm described in section 2 and to associate a single RVE to each cluster $\mathcal{S}^k$. The macroscopic strain is defined for the cluster $\mathcal{S}^k$ by:

$$\overline{\boldsymbol{\varepsilon}}^k = \frac{1}{\sum_{i\in\mathcal{S}^k} V^i\omega^i}\sum_{i\in\mathcal{S}^k}\overline{\boldsymbol{\varepsilon}}^i V^i\omega^i \tag{24}$$

where $\overline{\boldsymbol{\varepsilon}}^i$ is the macroscopic strain at one integration point i in the cluster $\mathcal{S}^k$, V^i is the volume of the element containing the integration point i, and ω^i is its associated Gauss weight. Then, after solving the RVE problem, the corresponding macroscopic stress $\overline{\boldsymbol{\sigma}}^k$ is affected to all integration points in $\mathcal{S}^k$. The k-means algorithm has negligible cost as compared to the RVE calculations. Then, the new number of nonlinear RVE problems to be solved within the k-means FE^2 method can be evaluated as:

$$N_{tot} \simeq (1+\beta) \times N_{clust} \times N_{iter} \times N_{evol}, \tag{25}$$

where N_{clust} is the average number of clusters along the whole simulation. If $N_{clust} << N_{int} \times N_e$, a drastic time reduction can be achieved.

3.3 | Clusters updating in presence of internal variables

At each loading step, the k-means algorithm is used to identify the K clusters of macroscopic Gauss points in the macro mesh. It is worth noting that between two iterations of the macro Newton scheme, the number and distributions of clusters might evolve. In the case of nonlinear behavior with internal variables at the micro scale, precautions must be taken. The most general approach would be to include the internal variable state $\boldsymbol{\alpha}(\mathbf{x})$ in the clustering analysis. However $\boldsymbol{\alpha}(\mathbf{x})$ is a field defined in all point $\mathbf{x} \in \Omega$ of the RVE and its description would involve a high-dimensional vector. Due to this issue, we choose not to include the micro internal variable state $\boldsymbol{\alpha}(\mathbf{x})$ in the cluster analysis, but to assign the RVEs that do not belong to the same cluster from one iteration to another to a new sub-cluster. Then, at one new macro Newton iteration, all RVEs belonging to one sub-cluster have close macroscopic strain $\overline{\boldsymbol{\varepsilon}}$ and similar internal variables distributions $\boldsymbol{\alpha}(\mathbf{x})$. In the following we define the *clusters* as the sets of RVEs grouped according to their macroscopic strain values $\overline{\boldsymbol{\varepsilon}}$ as defined in section 2 and the *sub-clusters* as those grouped according to both strain $\overline{\boldsymbol{\varepsilon}}$ and internal variables states $\boldsymbol{\alpha}(\mathbf{x})$. Then, the actual number of RVE computations to be performed is associated with the number of sub-clusters.

To summarize:

- The clusters define regions of the macroscopic mesh with close strain states;
- Sub-clusters are regions of the macroscopic mesh with similar strain state and history of internal variables. Then, each sub-cluster is associated with one RVE calculation;
- The number of K clusters is defined by the user and does not change through the iterations. At each iteration, the intersection between the previous and the actual sets of clusters defines the new sub-clusters, whose number may increase during the simulation.

Fig. 3 shows an illustrative example of the evolution of clusters and the number of RVEs (sub-clusters) in a schematic macroscopic structure, where each square represents a Gauss point. Fig. 3(a) displays the position of 3 initial clusters (red, blue and green) which are identical to the sub-clusters. Fig. 3(b)(c)(d) depicts an evolution of the 3 clusters at respective $p+1$, $p+2$, $p+3$-th iterations. The number of sub-clusters in the macroscopic structure has evolved, so that at the $p+1$-th iteration, cluster 1 contains 3 sub-clusters (3 RVEs) with different internal variable distributions, and cluster 3 comprises two sub-clusters (2 RVEs) with different internal variable distributions. At the end of the $p+3$-th iteration, the total number of sub-clusters is updated to 13. This procedure of updating the number of RVEs is integrated in the k-means FE^2 method. This approach has the drawback of splitting the clusters into sub-clusters at each change of cluster distributions. However, we will show in the numerical applications that this number still remains much lower than N_{int}.

It is important to note that the present approach does not depend on the number or type of internal variables within the RVE. As the generation of new sub-clusters is purely algorithmic (intersections of current and previous clusters), no a priori mechanical analysis is required. New generated sub-clusters then have the same strain history when they are created, but are subsequently submitted to different strain histories. As shown in the numerical examples, we can then apply this approach to arbitrary local history-dependent behaviors (damage, plasticity, etc.).

The k-means FE^2 procedure has been implemented in the finite element code Cast3m[47]. A summary of the k-means FE^2 is provided in Algorithm 2.

3.4 | Additional remarks

- Note that for nonlinear behaviors without internal variables, the number of sub-clusters remains identical to the one of clusters and even larger gains of computational times can be achieved.
- The above approach can be combined with parallel computing to perform all RVE simulations associated to sub-clusters on different processors.
- If model reductions or surrogate model methods are available to replace the nonlinear RVE calculations, they can be combined with this framework.

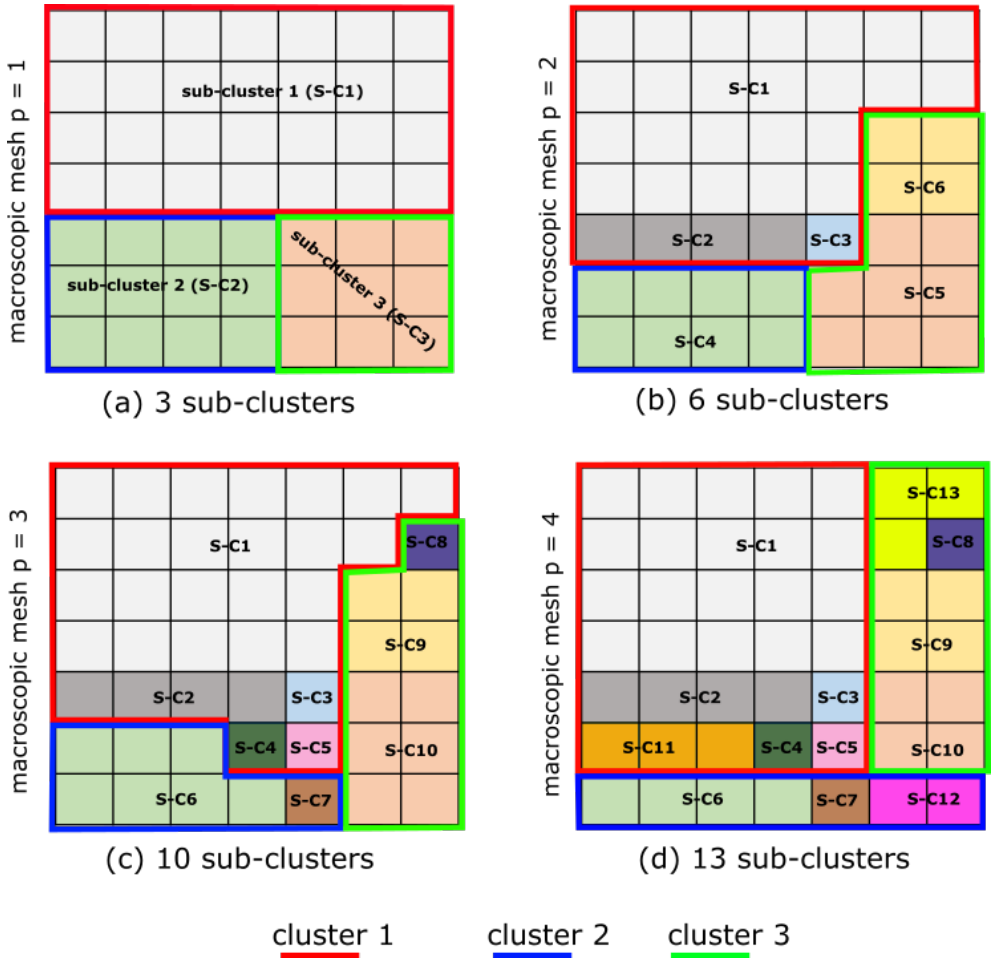


FIGURE 3 Schematic evolution of the clusters and number of RVE problems to be solved during a simulation. The square domains represent the elements of a macrcoscopic mesh. At first iteration of the Newton scheme (a), clusters are determined on the sole definition of macroscopic strains. At the second iteration (b), the clusters may be modified. The elements intersecting old and new clusters are splits into new "sub-clusters". Each sub-cluster is associated wih a nonlinear RVE calculation. For the next iterations (c) and (d), new sub-clusters can be formed.

4 | NUMERICAL EXAMPLES

In this section, we present several numerical examples to demonstrate the following features of the method and its interests for nonlinear multiscale computations: (a) the convergence of the method with respect to the number of clusters; (b) the validation of the method with respect to direct FE^2 simulations; (c) the low dependence of the number of clusters on the macro mesh, (d) the computational gains compared with FE^2 method and (e) the potential of the method for arbitrary micro nonlinear behaviors with internal variables.

4.1 | Material modeling at the micro scale

Small strains are assumed. We consider in the following microstructures composed of a nonlinear quasi-brittle matrix in which are distributed linear elastic or thermoelastic inclusions. As a result, the effective behavior of the material is nonlinear. It is computed by the multiscale method described in the previous sections. Linear and thermoelastic inclusions are described by the following constitutive law:

$$\boldsymbol{\sigma}(\mathbf{x}) = \mathbb{C}(\mathbf{x}) : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}(\Delta h)) \tag{26}$$

Algorithm 2 k-means FE^2

For $n = 1 : N$ (Loop over all time steps)
  $n = n + 1$
  Initialize: $\overline{\mathbf{u}}_0^{n+1} = \overline{\mathbf{u}}^n$
  While $E_{rr} > TOL$ (Macro Newton Loop).
    $p = p + 1$ (iterations)
    Input: $\overline{\boldsymbol{\varepsilon}}^e(\overline{\mathbf{u}}_j^{n+1}) \quad e = 1, 2...N_e$.
    Define clusters $\mathcal{S}^k \quad k = 1, 2, ..., K$ using Algorithm 1.
    Define sub-clusters $\mathcal{S}b_s^k$; ($s = 1$ in the first time step, see section 3.3).
    Compute $\overline{\boldsymbol{\varepsilon}}^k$ according to (24) .
    For $k = 1 : K$ (Loop over all clusters) .
      Input $\overline{\boldsymbol{\varepsilon}}^k$
        For $s = 1 : S$ (Loop over all sub-clusters in $\mathcal{S}^k$)
          Solve the nonlinear problem (9)-(12).
          Output: $\overline{\boldsymbol{\sigma}}^s, \overline{\mathbf{C}}_{tan}$.
          $\overline{\boldsymbol{\sigma}}^e(\overline{\mathbf{u}}_p^{n+1}) = \overline{\boldsymbol{\sigma}}^s, \quad \forall e \in \mathcal{S}b_s^k$.
          $\overline{\mathbf{C}}_{tan}^e(\overline{\mathbf{u}}_p^{n+1}) = \overline{\mathbf{C}}_{tan}^s, \quad \forall e \in \mathcal{S}b_s^k$.
          Compute $\left(\overline{\mathbf{K}}_{tan}^e\right)_{p+1}^{n+1}, \left(\overline{\mathbf{R}}^e\right)_{p+1}^{n+1}$.
          Assemble $\left(\overline{\mathbf{K}}_{tan}^e\right)_{p+1}^{n+1} \quad in \quad \left(\overline{\mathbf{K}}_{tan}\right)_{p+1}^{n+1}$.
          Assemble $\left(\overline{\mathbf{R}}^e\right)_{p+1}^{n+1} \quad in \quad \left(\overline{\mathbf{R}}\right)_{p+1}^{n+1}$.
        End.
    End.
    Solve $\left(\overline{\mathbf{K}}_{tan}\right)_{p+1}^{n+1} \Delta\overline{\mathbf{U}} = -\left(\overline{\mathbf{R}}\right)_{p+1}^{n+1}$.
    $\overline{\mathbf{u}}_{p+1}^{n+1} = \overline{\mathbf{u}}_p^{n+1} + \Delta\overline{u}$.
    $\mathbf{E}_{rr} = ||\overline{\mathbf{R}}_{p+1}^{n+1}||$.
    $\overline{\mathbf{u}}_p^{n+1} = \overline{\mathbf{u}}_{p+1}^{n+1}$.
  End.
  $\overline{\mathbf{u}}^n = \overline{\mathbf{u}}_p^{n+1}$.
End.

where $\mathbb{C}(\mathbf{x})$ is the elastic fourth-order tensor at point $\mathbf{x}$, and $\boldsymbol{\epsilon}(\Delta h)$ is an eigenstrain whose evolution is assumed to be given as a function of a parameter Δh associated with the evolution of the system. The matrix is assumed to be quasi-brittle and described by the phase field model for fracture, which is briefly reviewed in Appendix A. Then, here, the field of internal variables within the RVE which induces the creation of new sub-clusters (see section 3.3) is the damage field, $\boldsymbol{\alpha}(\mathbf{x}) \equiv d(\mathbf{x})$.

4.2 | Perforated plate in traction: local quasi-brittle behavior

The objective of this first example is to demonstrate the accuracy of the method and its convergence with respect to the number of clusters. We consider a plate of dimension $80 \times 80 \times 3$ mm^3 with a hole of radius $r = 15$ mm at its center. Due to symmetry, only a quarter of the structure is taken into account. The geometry and the boundary conditions are depicted in Fig. 4 (a). The mesh is composed of 401 linear tetrahedral elements (see Fig. 4 (b)). On the upper end of the sample, a displacement along the z-direction is prescribed, while the degrees of freedom in the other directions are blocked.

In this example, the RVE is composed of aggregates embedded in a cement matrix. A view of aggregates (inclusions) is depicted in Fig. 5 (a). The mesh is generated by using the python script Combs[48] operating in the CAD code Salome[49], see e.g.[50,51] and references therein. The mesh contains a total of 28042 tetrahedral elements. The volume fraction of inclusions is $f = 10\%$. The inclusions are assumed to be linear elastic with the following properties: $E^i = 80GPa$, $\nu^i = 0.19$, which denote the Young's modulus and the Poisson's ratio of inclusions, respectively. The matrix is assumed to be isotropic quasi-brittle and

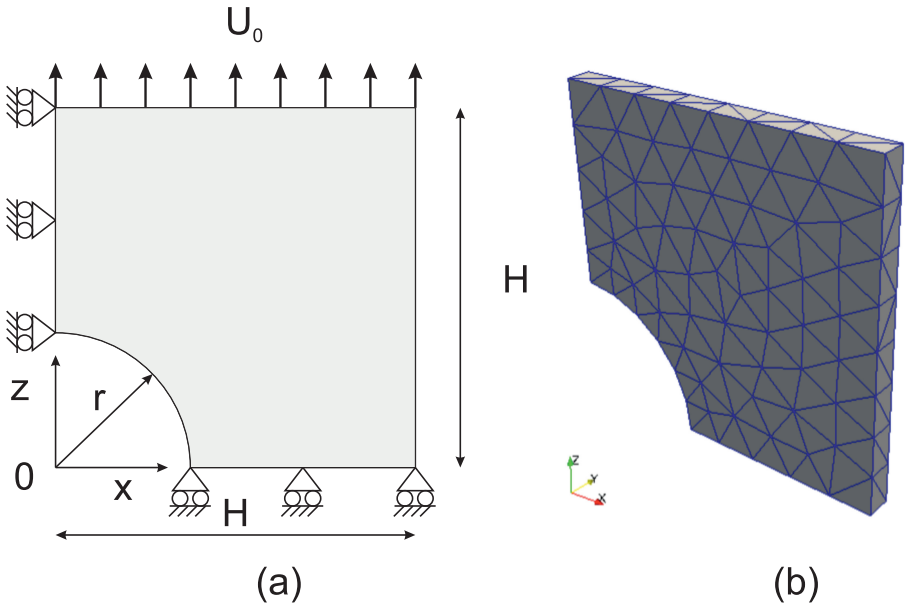


FIGURE 4 Perforated plate macro structure: (a): geometry and boundary conditions; (b) mesh.

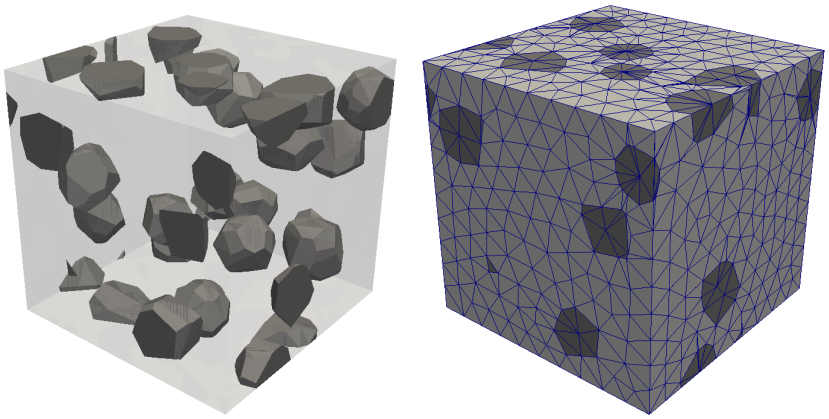


FIGURE 5 (a) RVE of a sand-cement microstructure. The aggregates are depicted in dark grey; (b) RVE mesh.

described by the phase field damage model provided in Appendix A with $E^m = 20$ GPa, $\nu^m = 0.2$ and $g_c^m = 0.12$ N/mm denoting the Young's modulus, Poisson's ratio and toughness of the matrix, respectively. The length scale parameter is chosen as $\ell = 2.8 \times 10^{-2}$ mm.

The k-means FE^2 method is applied for different numbers of clusters and its convergence is verified regarding the maximum von Mises stress in the macroscopic mesh with respect to the number of clusters in Fig. 6. We can note that 9 clusters are sufficient here to reach the convergence with regard to the chosen criterion, and that this converged solution is close to the reference FE^2 solution. In Fig. 7, we depict the force-displacement curve of the macro structure with respect to the number of clusters, again illustrating the convergence to the reference solution when increasing the number of clusters.

Figure 8 shows the evolution of the clusters during the simulation, respectively at the first loading step and at the final loading step. It is worth noting that as stated in section 3.3, the position of the clusters evolves during the simulation with the macro strain field.

We show the convergence of the local macro stress with respect to the number of clusters at the final loading step in Fig. 9, by depicting the von Mises stress fields in the structure for $k = 5, 6, 8$ and comparing it with the reference FE^2 solution.

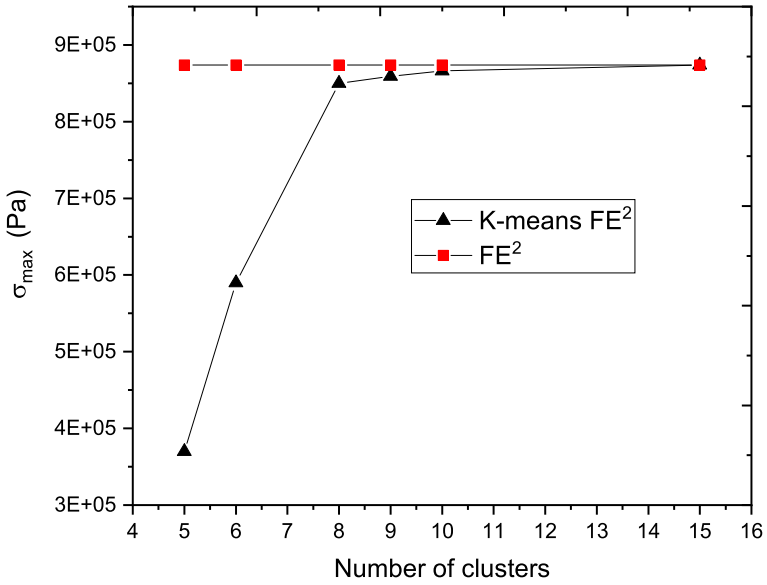


FIGURE 6 Convergence of the k-means FE^2 solution (von Mises stress) with respect to the number of clusters.

Finally, we report in Table 1 the total number of nonlinear RVE problems to be solved using the k-means FE^2 for different numbers of clusters, as well as the corresponding computational times, and compare it to the full FE^2 method. We can note a drastic reduction of computational times, with a ratio of about 15 for 8 clusters, which corresponds to the converged solution in Fig. 7.

k	Total number of calculated RVE problems (number of sub-clusters)	Final computation time (hours)
5	234	3.85 h
6	291	4.87 h
8	481	8.02 h
9	540	9.01 h
10	611	10 h
FE^2	7426	123 h

TABLE 1 Total number of nonlinear RVE problems and computational times on 32-cores LInux PC with 256 Go RAM with respect to the number of clusters for k-means and FE^2 solutions.

4.3 | 3-point bending beam: local quasi-brittle behavior

In this next example, we show an additional feature of the method: the low dependence of the number of clusters with respect to the mesh size. For this purpose, we consider a homogeneous beam of dimensions $200 \times 40 \times 2$ mm^3 in 3-point bending, as depicted in Fig. 10 (a). On $(x = 0, z = 0)$ mm, all degrees of freedom are blocked, whereas on $(x = 200, z = 0)$ mm the z and y displacements are blocked. A displacement in the $z-$ direction is prescribed at the middle of the beam. To test the influence of the macroscopic mesh refinement on the number of obtained clusters, 3 macroscopic meshes have been tested, including, respectively, 262, 431 and 786 tetrahedral elements (see Fig. 10). In this example, the RVE consists into a spherical elastic

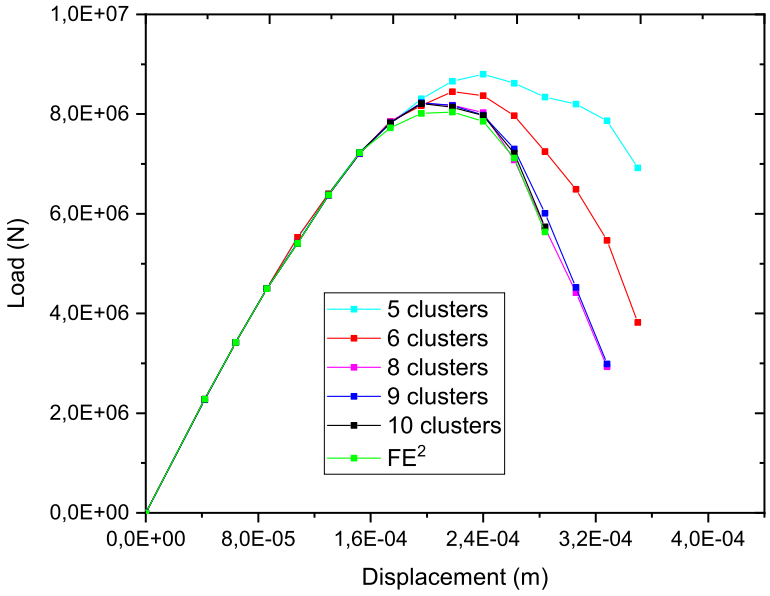
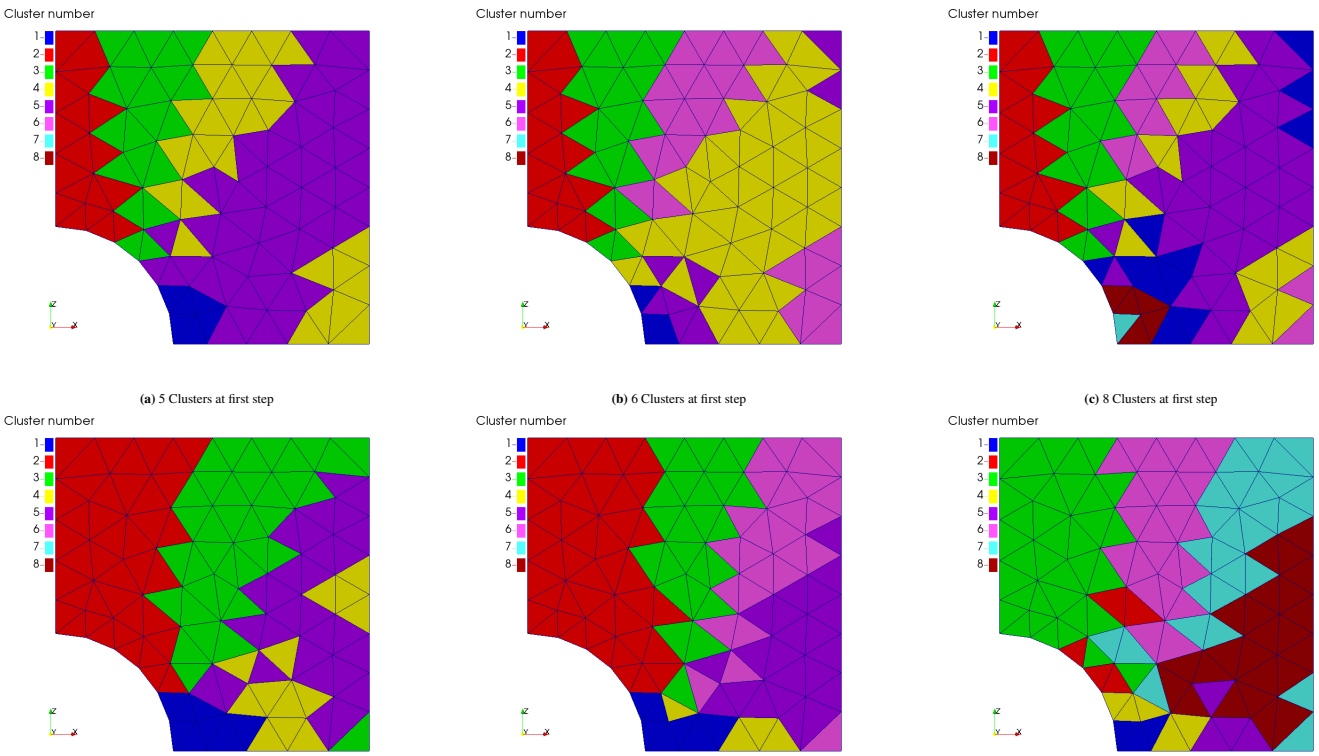


FIGURE 7 Perforated plate: convergence of the load-displacement curve with respect to the number of clusters k and reference FE^2 solution.



(a) 5 Clusters at first step

(b) 6 Clusters at first step

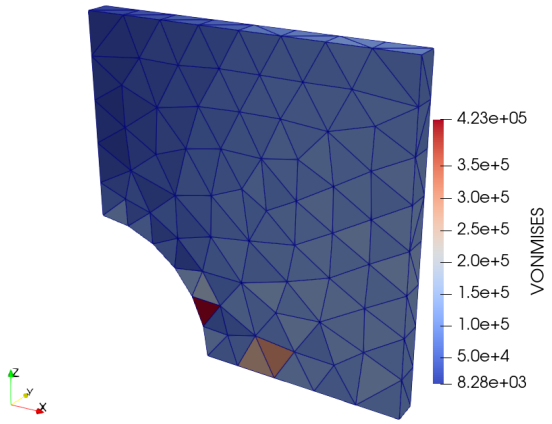
(c) 8 Clusters at first step

(d) 5 Clusters at final step

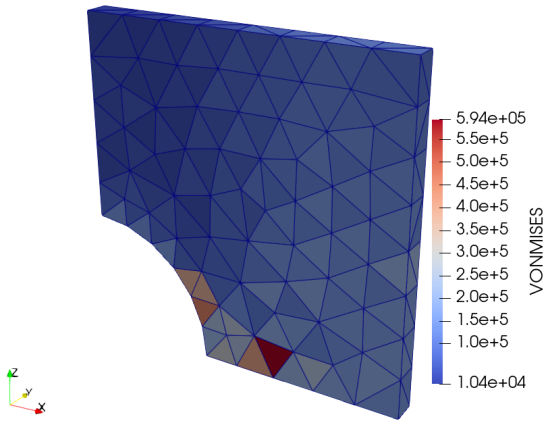
(e) 6 Clusters at final step

(f) 8 Clusters at final step

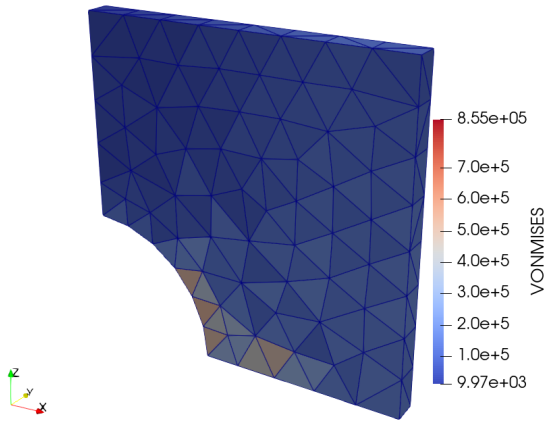
FIGURE 8 Cluster positions for the medium mesh at the first step (a) (b) (c) , and the final step (d) (e) (f) in the macroscopic structure for $K = 5, 6, 8$.



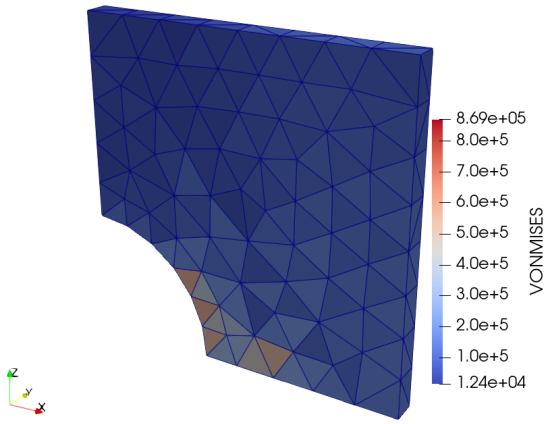
(a) von Mises stress: 5 Clusters



(b) von Mises stress: 6 Clusters



(c) von Mises stress: 8 Clusters



(d) von Mises stress: FE^2

FIGURE 9 von Mises stress in the macroscopic structure for $K = 5, 6, 8$ clusters, k-means solutions (a), (b) and (c), and FE^2 reference solution (d).

inclusion embedded into a quasi-brittle matrix. The dimensions of the RVE are $1 \times 1 \times 1$ mm^2. The inclusion is centered in the cubic RVE and has a radius $r = 0.14$ mm. The mechanical properties of the inclusion and the matrix are the same than in the previous example.

The k-means FE^2 method is applied for the different meshes and for different numbers of clusters. Convergence of the solutions with respect to the number of clusters for the different meshes is presented in Fig. 11, where the maximum von Mises stress is indicated at the final loading step of the simulations. For meshes 1 and 2, we have verified that the converged k-means solutions are in good agreement with the reference FE^2 solution. We did not perform this verification for the finest mesh, as the corresponding FE^2 computational costs was evaluated as too high (see table 2). As another observation, the total number of clusters required to converge to the reference solution does not increase significantly with the macro mesh size. This constitutes another important asset of the present method, even though this is obviously dependent on the regularity of the macro strain field. As another remark, it is worth pointing out that in the applications of the k-means FE^2 involving local damage, numerical regularization should be included at the macroscopic scale to guarantee convergence with respect to the macro mesh. For the sake of simplicity we did not implement such process here. Extensions with techniques like nonlocal damage model [52], gradient

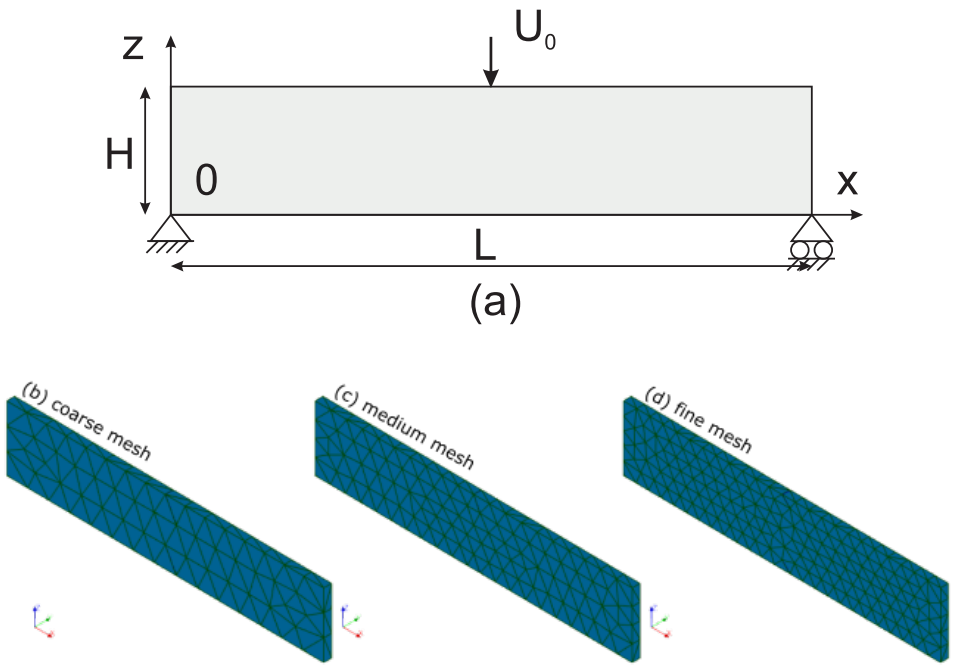


FIGURE 10 3-point bending beam; (a) geometry and boundary conditions; (b) coarse mesh; (c) medium mesh; (d) fine mesh.

enhanced damage[53], or gradient damage[54] could be considered in future studies. In addition, the existence of an RVE in the case of softening behavior is still a matter of debate in the literature and appropriate averaging techniques or analyses (see e.g.[55]) may be required in future extension of the present method in the case of local damage. As there is no macro regularization here, the 3 different meshes studied above are considered as 3 distinct cases and we do not intend to show a convergence with respect to the macro mesh in this example.

Figs. 12(a), (b) and (c) show the load-displacement curve of the beam subjected to 3-point bending for the 3 meshes and different numbers of clusters, and Fig. 12(d) shows the load-displacement curve of the 3 structures at convergence (K=8). According to these results the convergence of the model is reached for 8 clusters for the 3 meshes. The convergence with respect to the mesh refinement is also appreciated.

In Fig. 13, we illustrate the evolution of the clusters for the medium mesh when 5, 6 and 8 clusters are considered, for initial and final loading steps. As mentioned in section 3.3, the position of clusters changed in the macroscopic structure at the initial and final step, therefore the number of nonlinear RVE problem (sub-clusters) evolved.

In Fig. 14, we compare the macroscopic damage fields on the medium mesh for different numbers of clusters and the reference FE^2 solution. To better visualize the crack, a pseudo macro damage field D is defined as $D = 1 - \sqrt{\frac{\overline{\mathbb{C}}_{1111}(t)}{\overline{\mathbb{C}}^0_{1111}}}$, where t denotes pseudo time (loading evolution). We can see a crack initiating in the middle of the beam and propagating from the bottom to the top. The shape of the damaged area depends on the number of clusters K, and converges to the reference solution (FE^2) when $K = 9$.

Finally, we compare the computational times on the same Linux machine as in the previous example in Table 2, as a function of RVE calculations and number of clusters. Results show that the number of RVEs and thus the computational times are directly impacted by the mesh density in FE^2, while it affects in a minor way the computational times in k-means FE^2. The results show that keeping 8 clusters (corresponding to a converged solution), a speed-up ratio of almost 30 in comparison with FE^2 can be achieved for a similar accuracy. For the finest mesh, we only estimate the computational times, as the duration of the full FE^2 is not affordable. Then, the present technique allows solving problems with mesh sizes which are too costly in direct FE^2. Note that in the present work we have chosen damage as a local nonlinear behavior, but any other arbitrary local behaviors can be considered, like elastoplasticity or viscoplasticity.

As a final remark, if no FE^2 reference solution is available, the low computational costs of the method allow to use it for checking the convergence with respect to the number of clusters. In that case, the total computational costs must be calculated as the sum of the times related to several successive k-means analyses, with increasing numbers of clusters K. For example, in

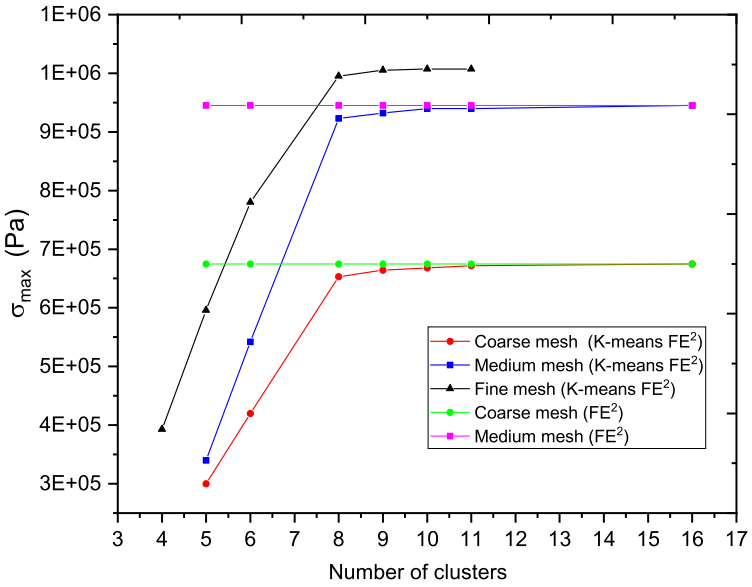


FIGURE 11 Convergence of the k-means FE^2 solutions with respect to the number of clusters for different meshes. The symbols indicate the maximum von Mises stress at the final loading step.

the present case, considering the 3 successive cases $K = 5$, $K = 6$, $K = 8$, the gains are still 5.4, 7.1 and 10.8 for meshes 1, 2 and 3, respectively, then yet an order of magnitude as compared to a direct FE^2 computation.

Number of clusters	Coarse mesh		Medium mesh		Fine mesh	
	Total number of nonlinear RVE problems (number of sub-clusters)	Final computation time (hours)	Total number of nonlinear RVE problems (number of sub-clusters)	Final computation time (hours)	Total number of nonlinear RVE problems (number of sub-clusters)	Final computation time (hours)
5	100	2.6 h	144	3.4 h	168	3.8 h
6	142	3.67 h	183	4 h	250	5.16 h
8	218	4.02 h	279	5.65 h	343	6.71 h
9	254	5 h	340	6.6 h	438	8 h
FE^2	3406	56 h	5603	93.38 h	10218 (estimated)	231.12 h (estimated)

TABLE 2 Comparisons of computational costs for both k-means FE^2 and FE^2 methods for the 3 meshes and different numbers of clusters.

4.4 | Perforated plate in traction: local elastoplastic behavior

In this example, the k-means FE^2 method is employed to solve the same macroscopic structure than in section 4.2 with the same boundary conditions, but with an elastoplastic behavior within the RVE. At the microscopic scale, a porous RVE of dimensions ($1 \times 1 \times 1$ mm^3) is considered, with a pore in the center of radius r = 0.15 mm. The behavior of the matrix is elastoplastic with linear isotropic hardening under the J_2-flow theory. In that context, the constitutive law is given by:

$$\boldsymbol{\sigma} = \mathbb{C} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) \tag{27}$$

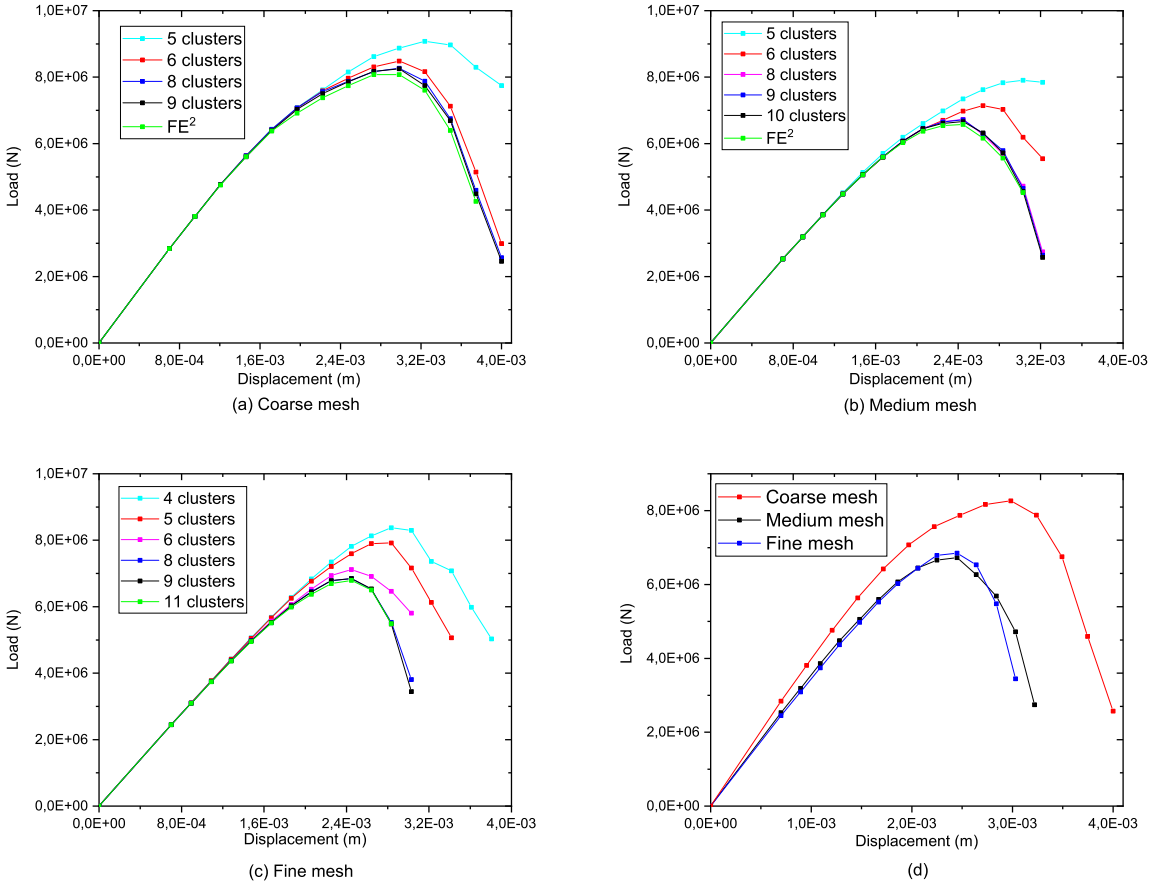


FIGURE 12 (a)-(c): Load-displacement curves of the 3 meshes with different number of clusters and FE^2 reference solution (except for (c)) and (d): comparison of the curves for the three meshes at convergence with respect to the number of clusters ssumed to be obtained for 8 clusters.

where $\boldsymbol{\varepsilon}^p$ is the plastic strain and where $\mathbb{C}$ is the elastic tensor, assumed to be isotropic. The elastic domain is defined by

$$f(\boldsymbol{\sigma}, p) = J_2(\boldsymbol{\sigma}) - \sigma_Y - Hp \leq 0 \tag{28}$$

where $J_2(\boldsymbol{\sigma}) = \left(\frac{3}{2}\mathbf{s} : \mathbf{s}\right)$, $\mathbf{s} = \boldsymbol{\sigma} - \frac{1}{3}Tr(\boldsymbol{\sigma})$, $Tr(.)$ being the trace operator, σ_Y is the yield stress, H is the linear hardening modulus and p the cumulated equivalent plastic strain expressed by:

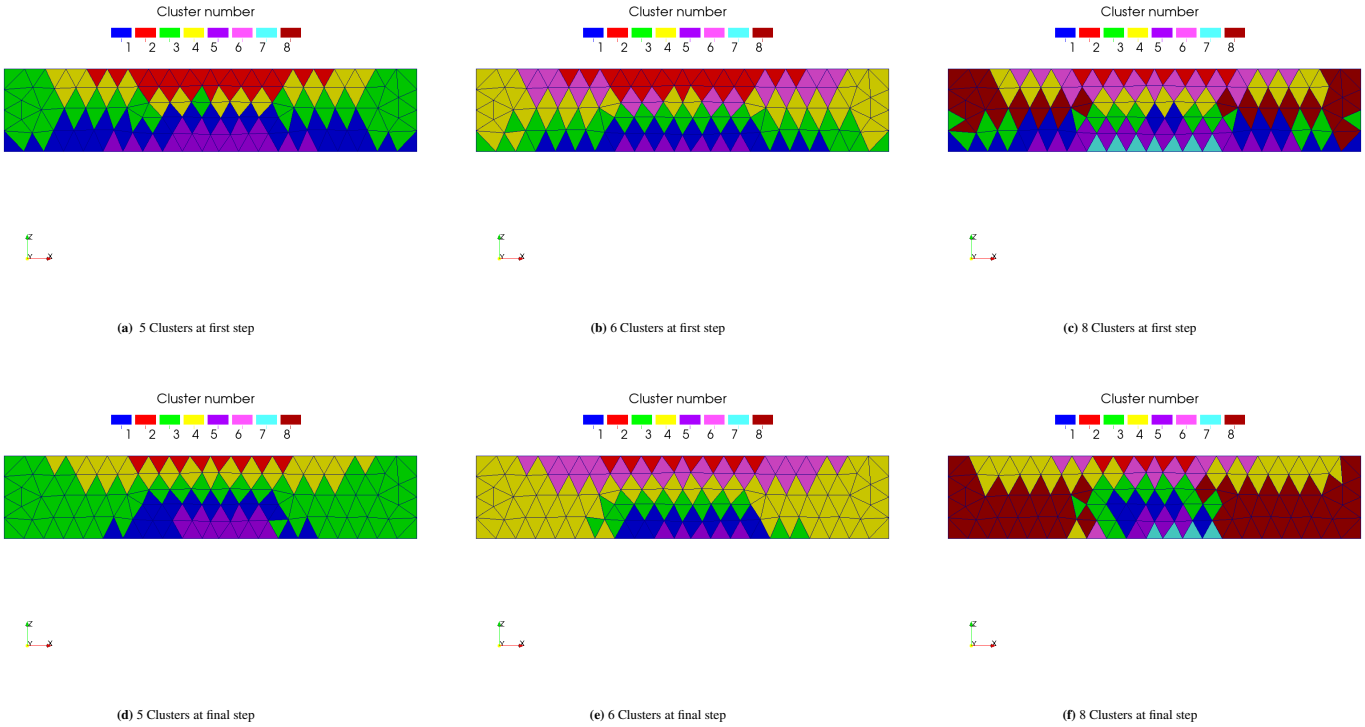
$$p(t) = \int_0^t \dot{p}(\tau) d\tau \tag{29}$$

with $\dot{p} = \left(\frac{2}{3}\dot{\boldsymbol{\varepsilon}}^p : \dot{\boldsymbol{\varepsilon}}^p\right)^{1/2}$. The evolution of the plastic strain $\boldsymbol{\varepsilon}^p$ is given by the normality rule:

$$\dot{\boldsymbol{\varepsilon}}^p = \dot{p}\frac{\partial f}{\partial \boldsymbol{\sigma}} \tag{30}$$

The mesh contains a total of 30452 tetrahedral elements. The mechanical properties of the matrix are as follows : $E^m = 15GPa$, $\nu^m = 0.2$, $\sigma_Y = 60MPa$ and $H = 1GPa$.

The k-means FE^2 method is applied for different numbers of clusters. Here, the field of internal variables within the RVE which induces the creation of new sub-clusters (see section 3.3) is the plastic strain field, $\boldsymbol{\alpha}(\mathbf{x}) \equiv \boldsymbol{\varepsilon}^p(\mathbf{x})$. The convergence is verified with respect to the maximum von Mises stress in the macroscopic structure for different number of clusters in Fig. 15(a). Convergence with respect to the number of clusters is reached for 9 clusters. In Fig. 15(b), the force-displacement curve of the macro-structure for different number of clusters is presented.



(a) 5 Clusters at first step **(b)** 6 Clusters at first step **(c)** 8 Clusters at first step

(d) 5 Clusters at final step **(e)** 6 Clusters at final step **(f)** 8 Clusters at final step

FIGURE 13 Cluster positions at the first step (a) (b) (c) , and at the final step (d) (e) (f) in the structure 2 for $K = 5, 6, 8$.

Figure 16 shows the von Mises stress field in the macroscopic structure for $K = 7$, 8, and 9 clusters and for the reference solution FE^2 at the final loading step, where again the convergence with respect to the number of clusters can be appreciated.

Table 3 summarizes the total number of nonlinear RVE problems solved by the k-means FE^2 method and the corresponding computation times for k = 6,7,8, and 9 clusters, as compared to a full FE^2 computation. A significant gain in terms of computational time is noticed, with a ratio of 17 for 9 clusters which corresponds to the converged solution.

k	Total number of calculated RVE problems (number of sub-clusters)	Final computation time (hours)
7	873	13.7 h
8	1062	17.87 h
9	1155	19.02 h
FE^2	19862	331 h

TABLE 3 Elastoplastic problem : Total number of nonlinear RVE problems and computational times on 32-cores Linux PC with 256 Go RAM with respect to the number of clusters for k-means and FE^2 solutions.

4.5 | Industrial example: waste package bahavior

The objective of this last example is to show the potential of the present method for industrial applications.

We are concerned here with contaminated wastes generated by industrial processes, containing different types of materials including metallic particles. In order to immobilize these wastes, a solidification process may be used to aggregate them into packages that have to fulfill minimal mechanical characteristics. The choice of the matrix material mainly depends on the type

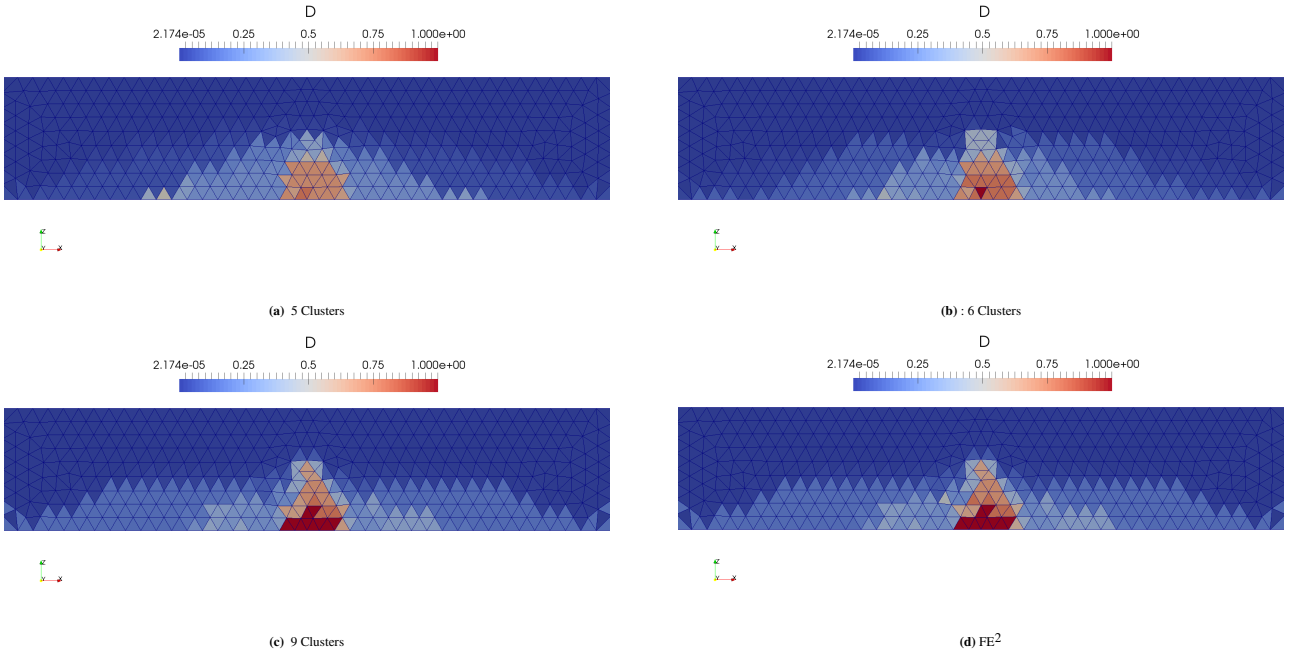


FIGURE 14 Macroscopic damage $D = 1 - \sqrt{\overline{C}_{1111}/\overline{C}^{0}_{1111}}$ for (a) $K = 5$; (b) $K = 6$, and (c) $K = 9$ clusters; (d) reference (FE^2) solution.

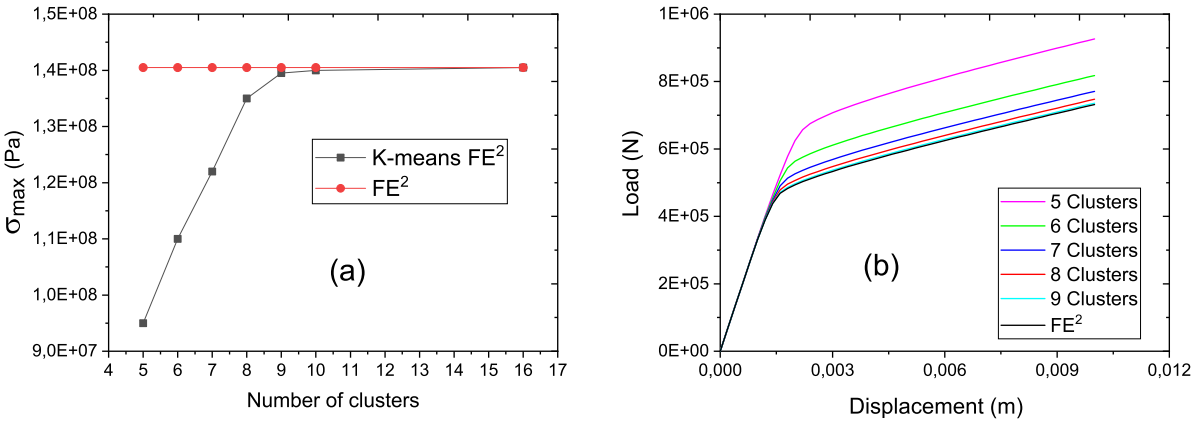
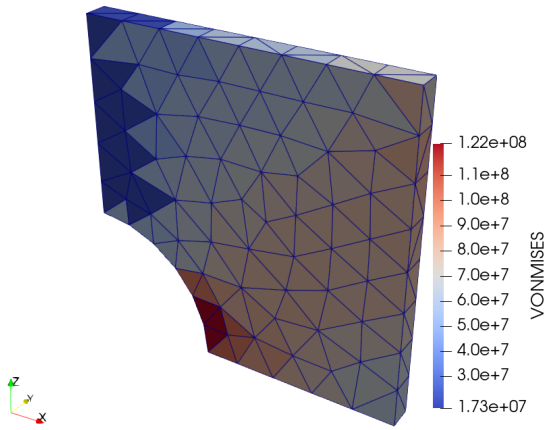


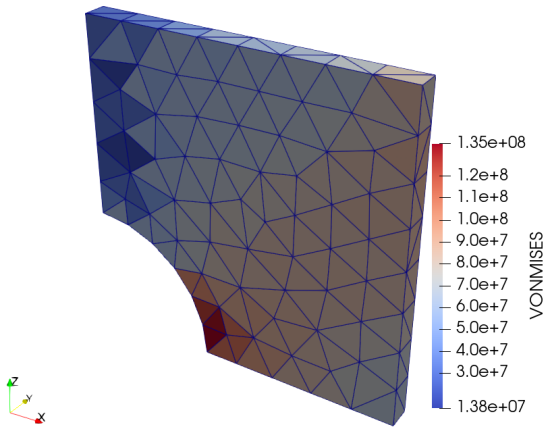
FIGURE 15 (a) Convergence of the k-means FE^2 solution (maximum von Mises stress) with respect to the number of clusters (b) convergence of the load-displacement curve with respect to the number of clusters k and reference FE^2 solution.

and the behavior of the particles in contact with the matrix. We consider a waste package where the matrix is made of hydrated cement paste, used for its low porosity and low permeability to liquids and gases in the hardened state, and then for its resistance to aggressive environments. However, the presence of free-water in the pores induces a progressive corrosion of reactive metal inclusions[56]. The corrosion reaction forms an expansive metal oxide layer around the metal inclusions, usually accompanied with a significant release of hydrogen gas[57,58,59]. Moreover, the thickness of the developed corrosion product layer can be important with respect to the metal inclusion, and the transformation process may in theory be complete if reactants are available (the whole inclusion can be transformed into corrosion products). This volume expansion associated with the formation of the corrosion products (metal oxide) can generate spurious micro-cracks in the cement matrix[59,60].

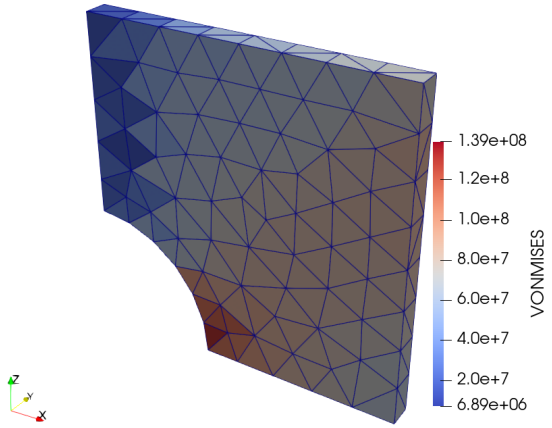
The studied macro structure is depicted in Fig. 17 where only a quarter is considered due to symmetry. The constitutive material is composed of a matrix of cement paste in which are dispersed non-reactive inclusions (graphite) and metallic particles.



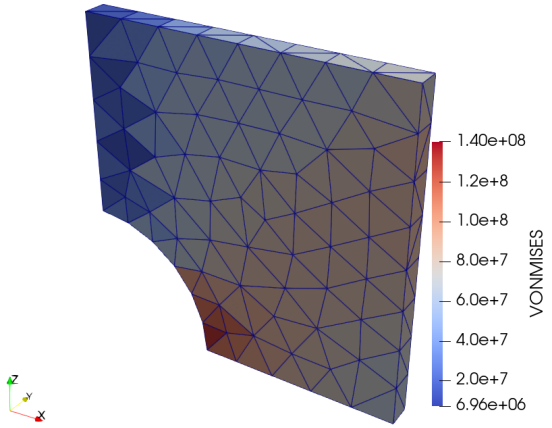
(a) von Mises stress: 7 Clusters



(b) von Mises stress: 8 Clusters



(c) von Mises stress: 9 Clusters



(d) von Mises stress: FE^2

FIGURE 16 von Mises stress in the macroscopic structure for $K = 7, 8, 9$ clusters, k-means solutions (a), (b) and (c), and FE^2 reference solution (d).

This mixture is solidified within a cylindrical container (see Fig. 17). The thickness of the external metallic shell is 15 mm in the cylindrical area and 20 mm in flat bottom area.

The RVE corresponding to the mixture is depicted in Figs. 18. Three cases are considered: (a) a single metallic inclusion in a cement paste matrix with equivalent initial radius of 3.5 mm representing a volume fraction of 0.05 %; (b) 10% in volume of graphite and 0.05% of metallic inclusions and (c) 20% of graphite and 0.05% of metallic inclusions. The metallic inclusions have an equivalent radius $r = 2.5$ mm in the last 2 cases. The graphite inclusions have polyhedral shapes. The size of the cubic RVE is 100 mm .

At the micro scale, the behavior of the graphite inclusions is assumed to be linear elastic, with properties provided in Table 4. The cement paste matrix is assumed to be quasi-brittle and described by the damage phase field model given in Appendix A. The behavior of uranium is linear thermoelastic, with linear properties described in Table 4 and with a constitutive law in the form:

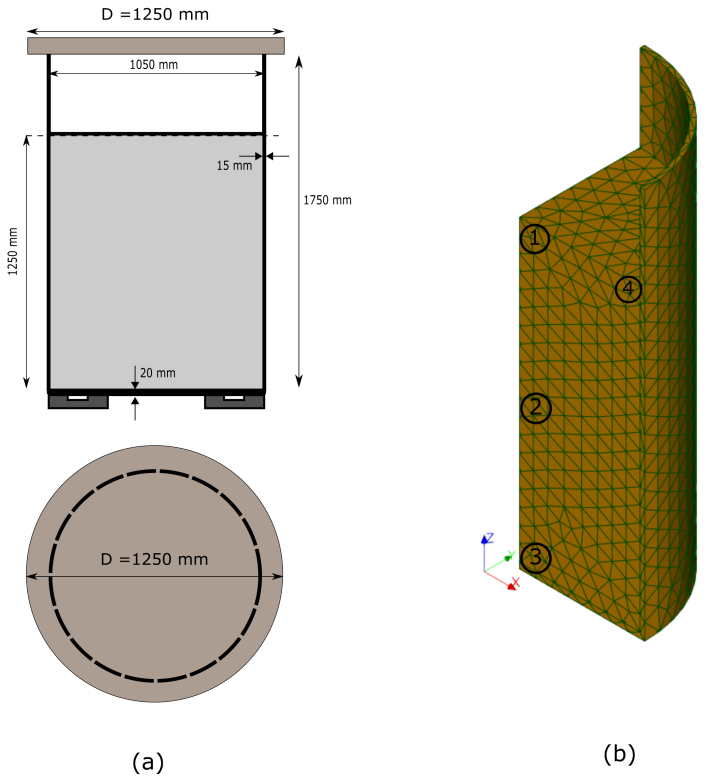


FIGURE 17 Waste package structure: (a) geometry and (b) macro mesh.

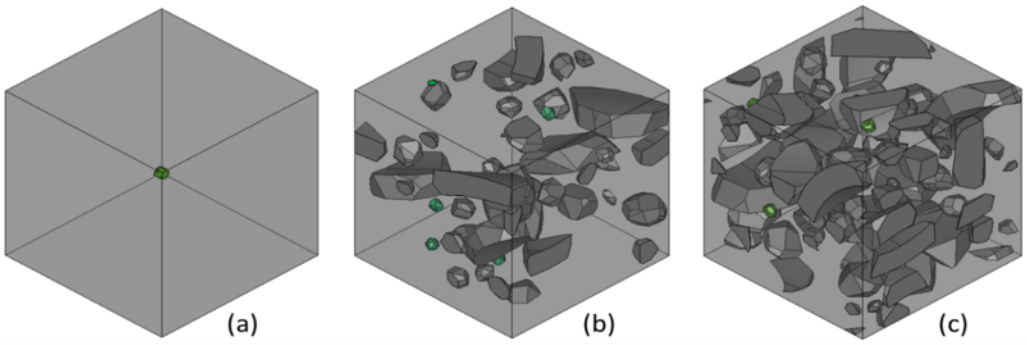


FIGURE 18 3 RVEs in the waste package structure problem; a) a single metallic inclusion in a cement paste matrix; (b) graphite (10 % and metallic inclusion 0.05%) and (c) graphite (20 % and metallic inclusion 0.05 %).

	Young Modulus	Poisson Coefficient
Cement paste matrix	20 GPa	0.2
Metallic inclusions	205 GPa	0.19
Graphite	15 GPa	0.25
Metal container	210 GPa	0.25

TABLE 4 Elastic parameters of the different phases.

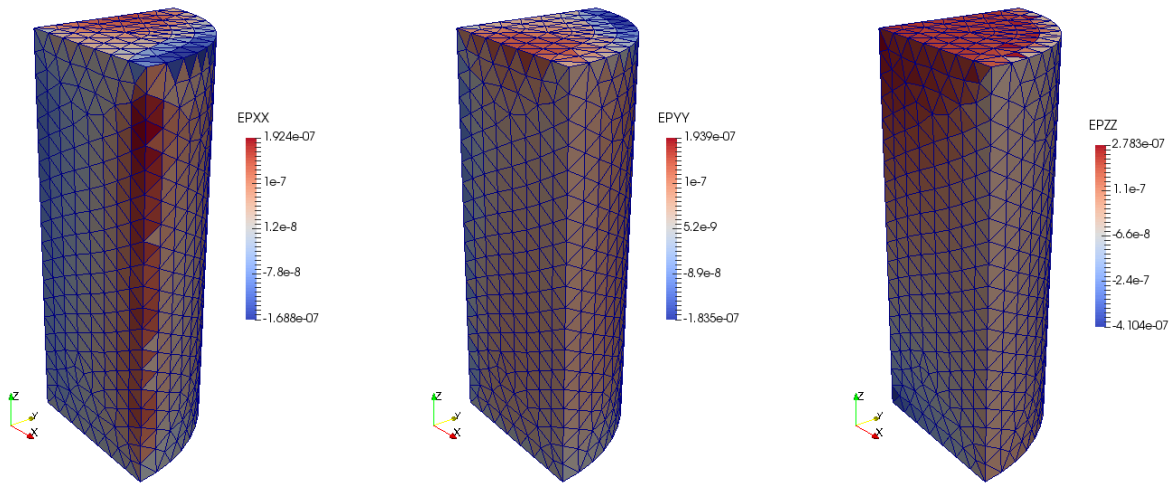


FIGURE 19 Macroscopic strain components under effective expansion $\varepsilon^* = 9.1 \times 10^{-7}$.

$$\boldsymbol{\sigma} = \mathbb{C} : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^*) \tag{31}$$

where $\boldsymbol{\varepsilon}^* = \beta_{eff}\Delta h$ is an eigenstrain associated with the swelling of the oxidized metallic inclusions, with $\beta_{eff} = 0.182$ m^{-1} (details on the calculation of such coefficient can be found e.g. in[50]). In Eq. 31, Δh is the thickness of the corrosion product layer. In our simulations, Δh was taken as the loading parameter. When Δh increases, the local strains near the inclusions within the RVE create damage which modifies the macro behavior. In what follows, we study the evolution of the local stiffness within the structure for increasing values of the corrosion thickness using the proposed k-means FE^2.

For illustrative purposes, Fig. 19 shows the macroscopic strain components with the effective expansion of $\varepsilon^* = 9.1 \times 10^{-7}$.

4.5.1 | Application 1

In this example the convergence of the k-means FE^2 with respect to the number of clusters in the macro structure is tested. The microstructure illustrated in figure 18 (a) is considered in the microscopic computations. The convergence of the k-means FE^2 method is verified by two different ways: (i) convergence with respect to the maximum von Mises stress within the structure at the final loading step and (ii) convergence with respect to the effective stiffness in chosen macro elements (when the value becomes quasi-constant as a function of Δh). Elements ①, ②, ③, and ④ are considered (see Fig. 17 (b)). Fig. 20 shows the position of clusters in the structure for $K = 5, 6, 8$ clusters at the final loading step.

Fig. 21 shows the decrease of effective stiffnesses as a function of the thickness of the corrosion product layer Δh for different numbers of clusters. For elements ① and ④, convergence is verified for 6 clusters. A faster convergence in element ① than in the element ④ is observed, which can be explained by the size of the cluster to which the element ① belongs, which remains unchanged for the different k values (see the red cluster in Fig. 20). As a comparison, an estimated computational time for this problem using the full FE^2 method would take 759 hours (45570 RVE calculations), while for $k = 8$ clusters the k-means FE^2 only uses 8.9 hours (538 RVE calculations).

4.5.2 | Application 2

In this second case, we consider both swelling metallic inclusions and graphite inclusions (Fig. 18). The RVEs meshes in Figs. 18 (b) and (c) contain 53706 and 73389 tetrahedral elements, respectively. Here $K = 8$ clusters have been used. Fig.22 shows the effective stiffness as a function of the thickness of the corrosion product layer Δh. When $\Delta h = 0$, the effective elastic moduli are the one of the undamaged material. A decrease of the effective stiffnesses is observed with the increase of the corrosion reaction in all elements of the package. In addition, the effective behavior of the material becomes progressively anisotropic because of the local distribution of graphite inclusions. Furthermore, a larger decrease is observed for all stiffnesses in the element ④ as compared to the elements ①, ②, and ③ in the two microstructures (b) and (c).

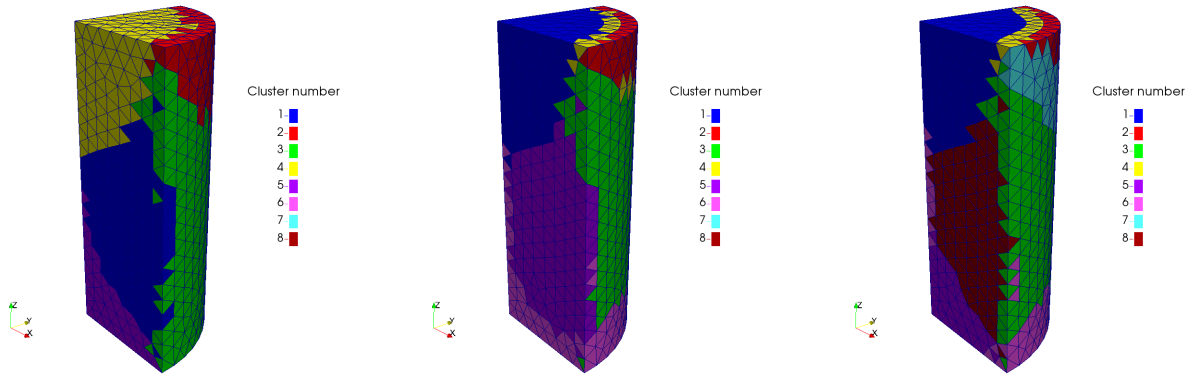


FIGURE 20 Clusters position for ($k = 5, 6, 8$) at the final loading step.

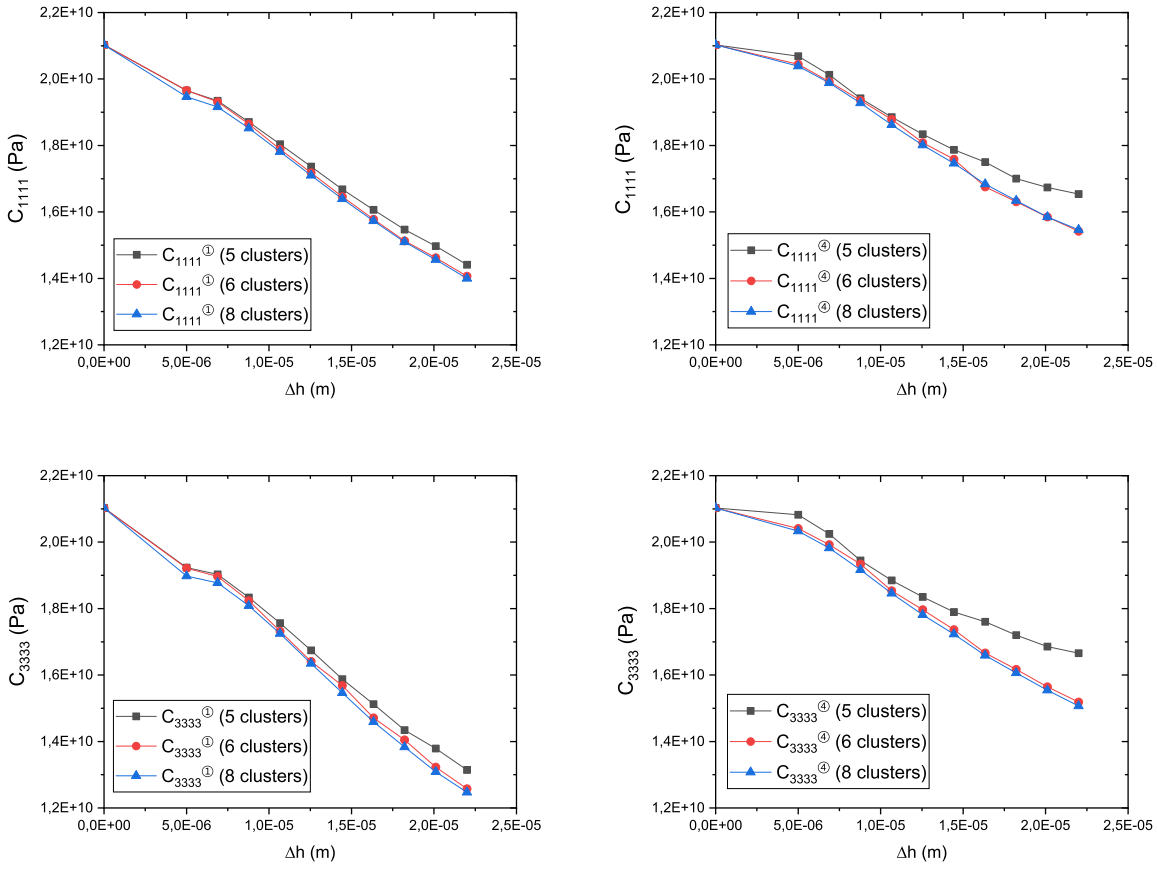


FIGURE 21 Effective stifnesses $\overline{C}_{1111}$ and $\overline{C}_{3333}$ as a function of Δh for the elements ① and ④ (see Fig. 17 (b)) for different numbers of clusters K=5,6 and 8.

The damage distribution of microstructures (b) and (c) in the elements ①, ③, and ④ is shown in Fig. 23, in which as expected, cracks appear around the metallic inclusions, and then they propagate within the microstructure. In addition, the damage is localized where the metallic inclusions are close to each other, and has a tendency to propagate in the direction of the nearest inclusion.

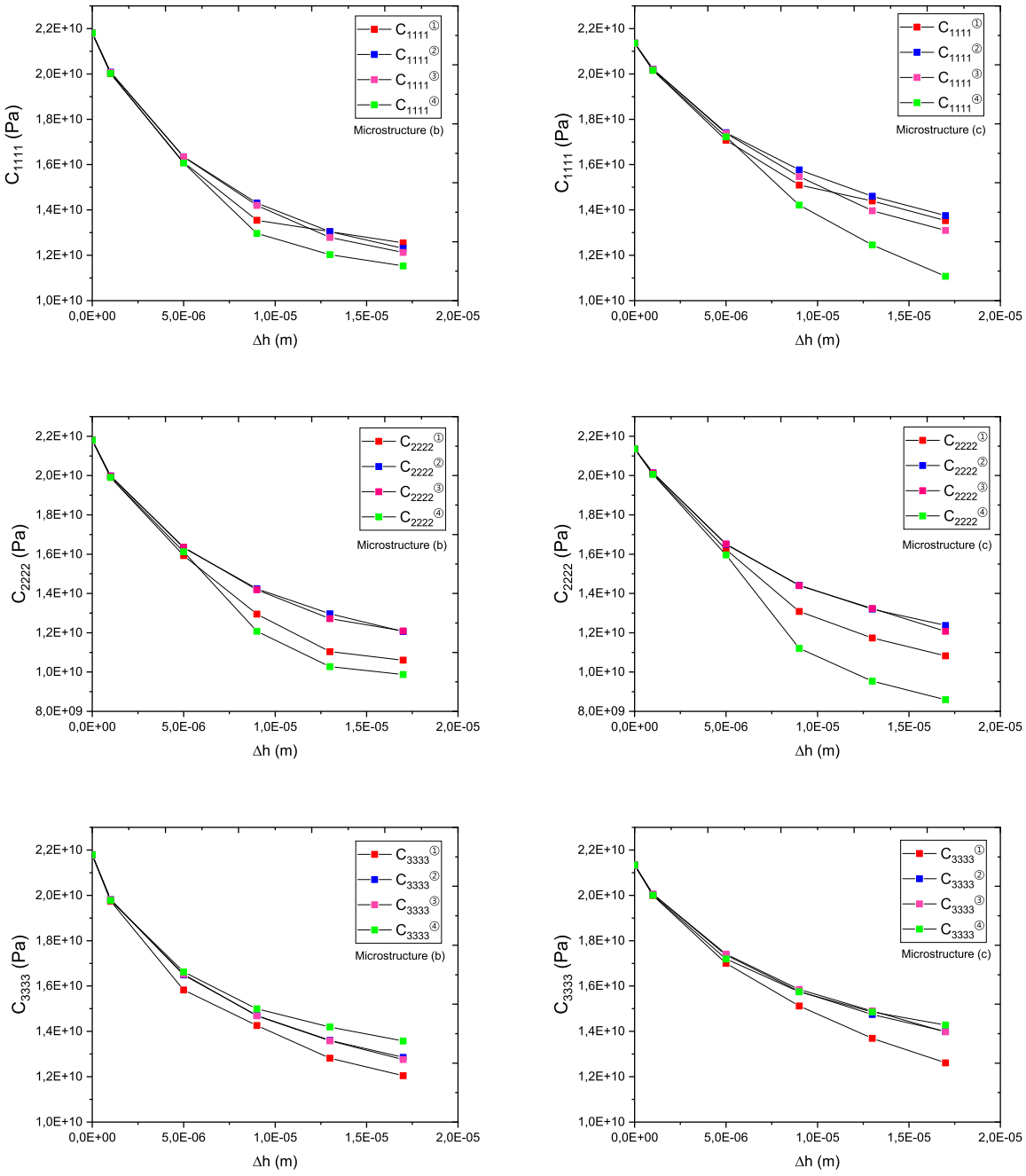


FIGURE 22 Effective stifnesses $\overline{C}_{1111}$, $\overline{C}_{2222}$ and $\overline{C}_{3333}$ as a function of Δh for the microstructures (b) and (c).

5 | CONCLUSION

In this work, we have presented a new machine-learning based multiscale method, called k-means FE^2, to solve general nonlinear multiscale problems with internal variables and loading history-dependent behaviors, without use of surrogate models. In contrast with machine learning multiscale methods where a surrogate model is constructed to describe the micro nonlinear effective behavior, we reduce here the macro scale problem by constructing clusters of Gauss points in the structure which are assumed to be in the same mechanical state. A k-means clustering - machine learning technique is employed to select the Gauss

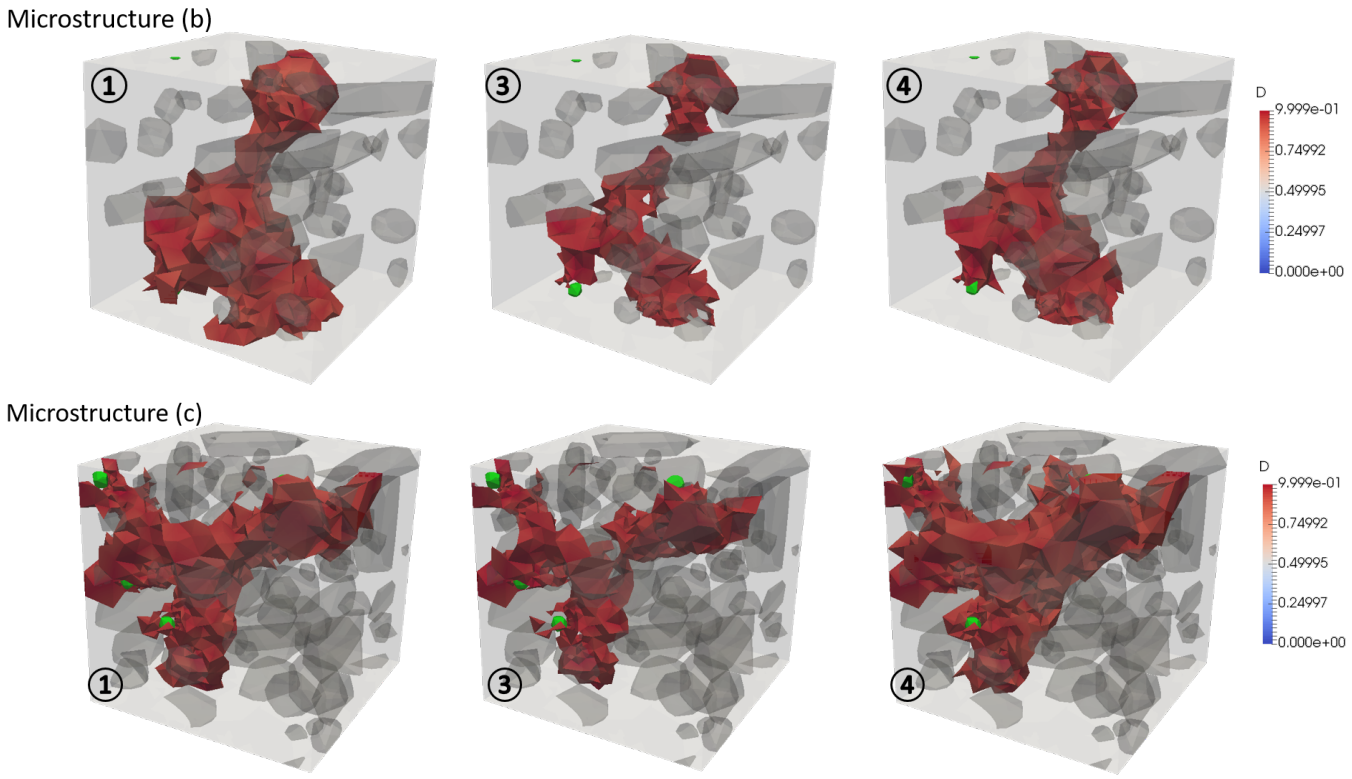


FIGURE 23 Local damage in RVEs designated by ①, ③ and ④ (see Fig. 17(b) for the definition of these elements), for microstructures (b) and (c) (see Fig. 18).

points based on their strain state and sets of internal variable. Then, for all Gauss points in a cluster, only one micro nonlinear problem is solved, and its response in terms of effective mechanical properties is transferred to all points of the cluster. The operation is repeated until convergence is reached at both macro and micro scales. The method presents several advantageous features. First, its solution is convergent with respect to the number of clusters. This is expected since in the limit of one cluster for one single macro mesh element, we recover the direct FE^2 solution. We have shown in the different numerical examples that even for a low number of clusters, a very accurate solution can be achieved as compared to the full FE^2 solution taken as a reference. In addition, we have shown that the total number of clusters, which is directly related to the computational time, is only weakly dependent on the mesh refinement in the macro structure. In fact, it only depends on the regularity of the fields at the macro level. We have demonstrated accelerations of FE^2 calculations up to a factor 50 in typical applications involving local fracture within a Representative Volume Element. It is worth noting that arbitrary nonlinear behaviors can be considered at the micro level including internal variables and loading history-dependent behaviors. In contrast, in surrogate-based machine learning methods, handling loading history-dependent local behaviors is delicate. Finally, the convergence of the method with respect to the number of clusters allows determining the necessary number of clusters without any knowledge about the final solution by gradually increasing the number of clusters and checking the convergence. The method has been applied to heterogeneous structures with local quasi-brittle behavior described by damage, and to a nuclear waste storage structure of industrial interest. As a final remark, this framework can be complementary to parallel computing and to machine-learning surrogate models to reduce the RVE calculations for nonlinear multiscale calculations.

6 | ACKNOWLEDGEMENTS

The financial support of Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA) and of Orano, France, is gratefully aknowledged.

7 | DATA AVAILABILITY STATEMENT

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

References

1. Shakiba M., Parson N., Chen X.G.. Modeling the Effects of Cu content and deformation variables on the high-temperature flow behavior of dilute Al-Fe-Si alloys using an artificial neural network. *Materials.* 2016;9(7):536.

2. Chojaczyk A., Teixeira A.P., Neves L.C., Cardoso J.B., Soares C.G.. Review and application of artificial neural networks models in reliability analysis of steel structures. *Structural Safety.* 2015;52:78–89.

3. Liu G.R., Xu Y.G., Wu Z.P.. Total solution for structural mechanics problems. *Computer methods in applied mechanics and engineering.* 2001;191(8-10):989–1012.

4. Waszczyszyn Z., Ziemiański L.. Neural networks in mechanics of structures and materials–new results and prospects of applications. *Computers & Structures.* 2001;79(22-25):2261–2276.

5. RehmanTahira Z.U., Mandal P.. Artificial neural network prediction of buckling load of thin cylindrical shells under axial compression. *Engineering Structures.* 2017;152:843–855.

6. Rao C., Liu Y.. Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization. *Computational Materials Science.* 2020;184:109850.

7. Lißner J., Fritzen F.. Data-driven microstructure property relations. *Mathematical and Computational Applications.* 2019;24(2):57.

8. Ryckelynck D., Goessel T., Nguyen F.. Mechanical dissimilarity of defects in welded joints via Grassmann manifold and machine learning. *Comptes Rendus Mécanique.* 2020;348(10-11):911–935.

9. Kirchdoerfer T., Ortiz M.. Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering.* 2016;304:81–101.

10. Carrara P., Lorenzis L., Stainier L., Ortiz M.. Data-driven fracture mechanics. *Computer Methods in Applied Mechanics and Engineering.* 2020;372:113390.

11. Mozaffar M., Bostanabad R., Chen W., Ehmann K., Cao J., Bessa M.. Deep learning predicts path-dependent plasticity. *Proceedings of the National Academy of Sciences.* 2019;116(52):26414–26420.

12. F. D. González, Chinesta , Cueto E.. Learning non-Markovian physics from data. *Journal of Computational Physics.* 2021;428:109982.

13. Feyel F.. Multiscale FE^2 elastoviscoplastic analysis of composite structure. *Computational Material Science.* 1999;16(1-4):433-454.

14. Feyel F., Chaboche J.-L.. FE^2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. *Computer Methods in Applied Mechanics and Engineering.* 2000;183(3-4):309–330.

15. Yvonnet J., Gonzalez D., He Q.-C.. Numerically explicit potentials for the homogenization of nonlinear elastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering.* 2009;198:2723-2737.

16. Le B.A., Q.-C. J. Yvonnet, He . Computational homogenization of nonlinear elastic materials using neural networks. *International Journal for Numerical Methods in Engineering.* 2015;104(12):1061–1084.

17. Fritzen F., Kunc O.. Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *European Journal of Mechanics A/Solids.* 2017;.

18. Lu X., Giovanis D., Yvonnet J., Papadopoulos V., Detrez F., Bai J.. A data-driven computational homogenization method based on neural networks for the nonlinear anisotropic electrical response of graphene/polymer nanocomposites. *Computational Mechanics.* 2019;64:307–321.

19. Li B., Zhuang X.. Multiscale computation on feedforward neural network and recurrent neural network. *Frontiers of Structural and Civil Engineering.* ;:1–14.

20. Logarzo H. J., Capuano G., Rimoli J.J.. Smart constitutive laws: Inelastic homogenization through machine learning. *Computer Methods in Applied Mechanics and Engineering.* 2021;373:113482.

21. Nguyen-Thanh V.M., Nguyen L.T.K., Rabczuk T., Zhuang X.. A surrogate model for computational homogenization of elastostatics at finite strain using high-dimensional model representation-based neural network. *International Journal for Numerical Methods in Engineering.* 2020;121(21):4811–4842.

22. Wang K., Sun W.. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. *Computer Methods in Applied Mechanics and Engineering.* 2018;334:337–380.

23. Ghavamian F, Simone A. Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network. *Computer Methods in Applied Mechanics and Engineering.* 2019;357:112594.

24. Wu L., Zulueta K., Major Z., Arriaga A., Noels L.. Bayesian inference of non-linear multiscale model parameters accelerated by a Deep Neural Network. *Computer Methods in Applied Mechanics and Engineering.* 2020;360:112693.

25. Peigney M.. A Fourier-based machine learning technique with application in engineering. *International Journal for Numerical Methods in Engineering.* 2021;122(3):866–897.

26. Rocha I.B.C.M., Kerfriden P., Meer F.P.. On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning. *Journal of Computational Physics: X.* 2020;:100083.

27. Liu Z., Bessa M.A., Liu W.K.. Self-consistent clustering analysis: an efficient multi-scale scheme for inelastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering.* 2016;306:319–341.

28. Liu Z., Fleming M., Liu W.K.. Microstructural material database for self-consistent clustering analysis of elastoplastic strain softening materials. *Computer Methods in Applied Mechanics and Engineering.* 2018;330:547–577.

29. Rocha I.B.C.M., Kerfriden P., Meer F.P.. Micromechanics-based surrogate models for the response of composites: A critical comparison between a classical mesoscale constitutive model, hyper-reduction and neural networks. *European Journal of Mechanics-A/Solids.* 2020;82:103995.

30. Avery P., Huang D.Z., He W., Ehlers J., Derkevorkian A., Farhat C.. A computationally tractable framework for nonlinear dynamic multiscale modeling of membrane woven fabrics. *International Journal for Numerical Methods in Engineering.* ;.

31. Wu L., Kilingar N.G., Noels L., others . A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths. *Computer Methods in Applied Mechanics and Engineering.* 2020;369:113234.

32. Bhattacharjee S., Matouš K.. A nonlinear manifold-based reduced order model for multiscale analysis of heterogeneous hyperelastic materials. *Journal of Computational Physics.* 2016;313:635–653.

33. Bhattacharjee S., Matouš K.. A nonlinear data-driven reduced order model for computational homogenization with physics/pattern-guided sampling. *Computer Methods in Applied Mechanics and Engineering.* 2020;359:112657.

34. Kaunda Rennie. New artificial neural networks for true triaxial stress state analysis and demonstration of intermediate principal stress effects on intact rock strength. *Journal of Rock Mechanics and Geotechnical Engineering.* 2014;6(4):338–347.

35. MacQueen J.. Some methods for classification and analysis of multivariate observations. In: :281–297Oakland, CA, USA; 1967.

36. Fayyad U.M., Piatetsky-Shapiro G., Smyth P., Uthurusamy R.. Advances in knowledge discovery and data mining. In: American Association for Artificial Intelligence; 1996.

37. Gersho A., Gray R.M.. *Vector quantization and signal compression*. Springer Science & Business Media; 2012.

38. Duda R.O., Hart P.E., Stork G.D.. *Pattern classification and scene analysis*. Wiley New York; 1973.

39. Hartigan J.A., Wong M.A.. AK-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics).* 1979;28(1):100–108.

40. Ozdemir I., Brekelmans W.A.M., Geers M.G.D.. Computational homogenization for heat conduction in heterogeneous solids. *International Journal for Numerical Methods in Engineering.* 2008;73(2):185–204.

41. Geers M.G.D, Kouznetsova V.G., Brekelmans W.A.M.. Multi-scale computational homogenization: Trends and challenges. *Journal of Computational and Applied Mathematics.* 2010;234(7):2175-2182.

42. Terada K., Kikuchi N.. A class of general algorithms for multi-scale analysis of heterogeneous media. *Computer Methods in Applied Mechanics and Engineering.* 2001;190:5427-5464.

43. Kouznetsova V.G., Geers M.G.D., Brekelmans W.A.M.. Multi-scale constitutive modeling of heterogeneous materials with gradient enhanced computational homogenization scheme. *International Journal for Numerical Methods in Engineering.* 2002;54:1235-1260.

44. Ghosh S., Lee K., Raghavan P.. A multilevel computational model for multi-scale damage analysis in composite and porous media. *International Journal of Solids and Structures.* 2001;38:2335-2385.

45. Geers M.G.D., Yvonnet J.. Multiscale Modeling of Microstructure-Property Relations. *MRS Bulletin.* 2016;41(08):610-616.

46. Yvonnet J.. *Computational Homogenization of Heterogeneous Materials with Finite Elements*. Springer Nature; 2019.

47. *http://www-cast3m.cea.fr/.* 2021.

48. Bourcier C, Dridi W, Chomat L, Laucoin E, Bary B, Adam E. Combs: open source python library for RVE generation. Application to microscale diffusion simulations in cementitious materials. In: :02107EDP Sciences; 2014.

49. *Salome software, https://www.salome-platform.org/.* 2021.

50. Benaimeche M.A., Bary B.. Analytical and numerical prediction of mechanical properties of heterogeneous materials with expansive inclusions: Application to waste cementation. *International Journal of Mechanical Sciences.* 2020;169:105318.

51. Bary B., Bourcier C., Helfer T.. Analytical and 3D numerical analysis of the thermoviscoelastic behavior of concrete-like materials including interfaces. *Advances in Engineering Software.* 2017;112:16–30.

52. Pijaudier-Cabot G., Bažant Z.. Nonlocal damage theory. *J. Eng. Mech..* 1987;113:1512-1533.

53. Peerlings R.H.J., Borst R., Brekelmans W.A.M., Vree H.P.J.. Gradient-enhanced damage for quasi-brittle materials. *International Journal for Numerical Methods in Engineering.* 1996;39(39):3391-3403.

54. Marigo Jean-Jacques, Maurini Corrado, Pham Kim. An overview of the modelling of fracture by gradient damage models. *Meccanica.* 2016;51(12):3107–3128.

55. Nguyen Vinh Phu, Lloberas-Valls Oriol, Stroeven Martijn, Sluys Lambertus Johannes. On the existence of representative volumes for softening quasi-brittle materials–a failure zone averaging scheme. *Computer Methods in Applied Mechanics and Engineering.* 2010;199(45-48):3028–3038.

56. Bart F., Coumes C. Cau, Frizon F., Lorente S.. *Cement-based materials for nuclear waste storage*. Springer Science & Business Media; 2012.

57. Banos A., Hallam K.R., Scott T.B.. Corrosion of uranium in liquid water under contained conditions with a headspace deuterium overpressure. Part 2: The ternary U+ H2O (l)+ D2 system. *Corrosion Science.* 2019;152:261–270.

58. Haschke J.M.. Corrosion of uranium in air and water vapor: consequences for environmental dispersal. *Journal of alloys and compounds.* 1998;278(1-2):149–160.

59. Stitt C.A., Paraskevoulakos C., Banos A., et al. In-situ, time resolved monitoring of uranium in BFS: OPC grout. Part 2: Corrosion in water. *Scientific reports.* 2018;8(1):1–12.

60. Stitt CA, Paraskevoulakos C, Banos A, et al. In-situ, time resolved monitoring of uranium in BFS: OPC grout. Part 1: Corrosion in water vapour. *Scientific reports.* 2017;7(1):1–9.

61. Francfort G.A., Marigo J.J.. Revisiting brittle fracture as an energy minimization problem. *Journal of the Mechanics and Physics of Solids.* 1998;46(8):1319-1342.

62. Bourdin B., Francfort G.A., Marigo J.J.. *The Variational Approach to Fracture*. Springer-Verlag, Berlin; 2008.

63. Pham K., Marigo J.-J.. The variational approach to damage: I. The foundations. *C. R. Mecanique.* 2010;338:191-198.

64. Bourdin B., Francfort G.A., Marigo J.J.. Numerical experiments in revisited brittle fracture. *Journal of the Mechanics ad Physics of Solids.* 2000;48:797-826.

65. Buliga M.. Energy Minimizing Brittle Crack Propagation. *J. Elast.*. 1999;52:201-238.

66. Amor H., Marigo J.-J., Maurini C.. Regularized formulation of the variational brittle fracture with unilateral contact: Numerical experiments. *Journal of the Mechanics and Physics of Solids.* 2009;57(8):1209-1229.

67. Miehe C., Hofacker M., Welschinger F.. A phase field model for rate-independent crack propagation: Robust algorithmic implementation based on operator splits. *Computer Methods in Applied Mechanics and Engineering.* 2010;199:2776-2778.

68. He Q.-C., Shao Q.. Closed-Form Coordinate-Free Decompositions of the Two-Dimensional Strain and Stress for Modeling Tension–Compression Dissymmetry. *Journal of Applied Mechanics.* 2019;86(3):031007.

69. Nguyen T.T., Yvonnet J., Waldmann D., He Q.-C.. Implementation of a new strain split to model unilateral contact within the phase field method. *International Journal for Numerical Methods in Engineering.* 2020;121(21):4717–4733.

70. Nguyen T.T., Yvonnet J., Zhu Q.-Z., Bornert M., Chateau C.. A phase field method to simulate crack nucleation and propagation in strongly heterogeneous materials from direct imaging of their microstructure. *Engineering Fracture Mechanics.* 2015;139:18-39.

How to cite this article: X , X , X , X , and X (X), X, *X*, *X*.

APPENDIX

A BRIEF REVIEW OF PHASE FIELD METHOD FOR QUASI-BRITTLE FRACTURE

The phase field method[61,62,63,64,65,66,67] is adopted here at the microscale to model the damage through microcracking within the microstructure. The formulations are briefly reviewed in this appendix.

For an elastic cracked body defined in a domain $\Omega \subset \mathbb{R}^3$ containing sharp cracks denoted collectively as Γ, the total energy of the system is defined as:

$$E = \int_{\Omega} \Psi(\boldsymbol{\epsilon}, \Gamma) d\Omega + g_c \int_{\Gamma} d\Gamma - \int_{\partial\Omega_F} \mathbf{F}^* \cdot \mathbf{u} d\Gamma \tag{A1}$$

where $\Psi(\boldsymbol{\varepsilon}, \Gamma)$ is the elastic strain density function and g_c is the critical energy release rate in the sense of Griffith. In Eq. (A1), $\partial\Omega_F$ denotes the portion of the boundary $\partial\Omega$ where tractions $\mathbf{F}^*$ are prescribed. The above energy form can be replaced by a regularized one, given by:

$$E = \int_{\Omega} \Psi(\boldsymbol{\varepsilon}, d) d\Omega + g_c \int_{\Omega} \gamma(d, \nabla d) d\Omega - \int_{\partial\Omega_F} \mathbf{F}^* \cdot \mathbf{u} d\Gamma, \tag{A2}$$

where γ is a crack density function, whose model can be chosen among several possible forms, leading to a class of shapes for the regularized damage field near the crack (see e.g.[67]). In this regularized framework, the cracks are no more described by surfaces but by a smooth field $d(\mathbf{x})$. For an isotropic elastic solid, Ψ can be decomposed according to:

$$\Psi(\boldsymbol{\varepsilon}) = g(d)\Psi^+ + \Psi^- \tag{A3}$$

to only affect the damage to traction modes, and where $g(d)$ is a degradation function such that $g(0) = 1$, $g(1) = 0$ and $g'(1) = 0$ and $\Psi^+(\boldsymbol{\varepsilon}^+)$ and $\Psi^-(\boldsymbol{\varepsilon}^-)$ denote parts of the strain density related to tensile and compressive parts of the strain tensor, respectively, which are defined here following Miehe[67] by:

$$\Psi^{\pm}(\boldsymbol{\varepsilon}) = \lambda(\langle Tr(\boldsymbol{\varepsilon})\rangle_{\pm})^2/2 + \mu Tr\{(\boldsymbol{\varepsilon}^{\pm})^2\}. \tag{A4}$$

This form allows avoiding interpenetration when the cracks are closed without any special algorithm for auto-contact, which renders the implementation very simple. In (A4), λ and μ denote the elastic Lamé's constants in each phase. The operator $\langle . \rangle_{\pm}$ is defined as $\langle x \rangle_{\pm} = (x \pm |x|)/2$, $\boldsymbol{\varepsilon}^+$ is the tensile part while $\boldsymbol{\varepsilon}^-$ is the compression part of the strain tensor obtained by the spectral decomposition:

$$\boldsymbol{\varepsilon}^{\pm} = \sum_{i=1}^{n} \langle Tr(\boldsymbol{\varepsilon})\rangle_{\pm} \mathbf{v}_i \otimes \mathbf{v}_i \tag{A5}$$

in which $\mathbf{v}_i$ are the eigenvectors of the strain tensor $\boldsymbol{\varepsilon}$. Other decompositions have been proposed as in[66,68,69].

Variational approach to damage in tandem with regularization, called in the literature "phase field method", implies the minimization of the total energy with respect to the displacement field $\mathbf{u}$ and the minimization of the energy with respect to the scalar field d describing the crack surface in a smooth manner. This second minimization is subjected to an inequality constraint $\dot{d} \geq 0$. To formulate this minimization problem in a simpler setting, a time-stepping $\mathcal{T} = \left\{t^0, t^1, ..., t^n, t^{n+1}, ..., t^N\right\}$ can be introduced. At each time step t^{n+1}, the problem is to find the displacement fields $\mathbf{u}^{n+1}$ and d^{n+1} such that

$$\mathbf{u}^{n+1}, d^{n+1} = \underset{\substack{\mathbf{u}\in\mathcal{K}_A \\ 0\leq d^n \leq d^{n+1} \leq 1}}{\operatorname{Argmin}} E \tag{A6}$$

where $\mathcal{K}_A$ is a set of kinematically admissible fields. One possible algorithm to solve this problem is to use sequential solving of both minimization problems as

$$D_{\delta\mathbf{u}} E = 0 \tag{A7}$$

$$D_{\delta d} E = 0, \;\; 0 \leq d^n \leq d^{n+1}, \tag{A8}$$

where $D_{\delta\mathbf{v}} f(\mathbf{u})$ is the Gateaux derivative, defined as:

$$D_{\delta\mathbf{v}} f(\mathbf{u}) = \left\{ \frac{f}{d\alpha} \left(f\left(\mathbf{u} + \alpha\, \delta\mathbf{v}\right)\right) \right\}_{\alpha=0}. \tag{A9}$$

The first equation (A7) defines the mechanical problem while the second one (A8) defines the phase field problem. These two problems are coupled as both involve the fields $\mathbf{u}$ and d.

Eq. (A7) can be developed according to:

$$\int_{\Omega} \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}, d) \,:\, \boldsymbol{\varepsilon}(\delta\mathbf{u}) d\Omega - \int_{\partial\Omega_F} \overline{\mathbf{F}} \cdot \delta\mathbf{u} d\Gamma = 0, \tag{A10}$$

where

$$\frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}, d) = \boldsymbol{\sigma}. \tag{A11}$$

For $g(d) = \left((1-d)^2 + k\right)$, where k is a small scalar parameter introduced to avoid ill-conditioning when elements are fully damaged, and with Ψ defined as in (A3) we obtain:

$$\boldsymbol{\sigma} = \left((1-d)^2 + k\right)\left\{\lambda \langle Tr\boldsymbol{\varepsilon}\rangle_+ \mathbf{1} + 2\mu\boldsymbol{\varepsilon}^+\right\} + \lambda \langle Tr\boldsymbol{\varepsilon}\rangle_- \mathbf{1} + 2\mu\boldsymbol{\varepsilon}^-. \tag{A12}$$

It yields the classical weak form of the mechanical problem as follows: find $\mathbf{u} \in \mathcal{D} = \left\{\mathbf{v} | \mathbf{v} = \overline{\mathbf{u}} \text{ on } \partial\Omega_u, \mathbf{v} \in H^1(\Omega)\right\}$ such that:

$$\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\delta\mathbf{u})\, d\Omega = \int_{\partial\Omega_F} \mathbf{F}^* \cdot \delta\mathbf{u}\, d\Gamma \tag{A13}$$

for all $\delta\mathbf{u} \in H_0^1(\Omega) = \left\{\mathbf{v} | \mathbf{v} = 0 \text{ on } \partial\Omega_u, \mathbf{v} \in H^1(\Omega)\right\}$, where $\partial\Omega_u$ denotes the portion of $\partial\Omega$ where Dirichlet boundary conditions are prescibed. The Euler-Lagrange equation (strong form) associated with Eq. (A13) is given by:

$$\nabla \cdot \boldsymbol{\sigma} = 0, \quad \boldsymbol{\sigma}\mathbf{n} = \mathbf{F}^* \text{ over } \partial\Omega_F, \quad \mathbf{u} = \mathbf{u}^* \text{ over } \partial\Omega_u, \tag{A14}$$

where $\mathbf{u}^*$ are prescribed displacements. The first equation in (A8) can be developed as:

$$\int_\Omega \frac{\partial\Psi}{\partial d}\delta d\, d\Omega + g_c \int_\Omega D_{\delta d}\gamma d\Omega = 0. \tag{A15}$$

Choosing γ as

$$\gamma(d, \nabla d) = \frac{1}{2\ell} d^2 + \frac{\ell}{2}\nabla d \cdot \nabla d \tag{A16}$$

we obtain, after some calculations (see e.g. [70]):

$$\int_\Omega \left(2\left[\Psi^+\right] + \frac{g_c}{\ell}\right) d\delta d + g_c \ell \nabla d \cdot \nabla(\delta d) d\Omega = \int_\Omega 2\left[\Psi^+\right] \delta d\, d\Omega. \tag{A17}$$

The associated Euler-Lagrange equations to (A17) are given by:

$$\left(2\left[\Psi^+\right] + \frac{g_c}{\ell}\right) d - \ell\, g_c \Delta d = 2\left[\Psi^+\right],$$

$$\nabla d \cdot \mathbf{n} = 0 \text{ over } \partial\Omega, \tag{A18}$$

where Δd denotes the Laplacian operator. We follow the classical Miehe's approach [67] to enforce the irreversibility condition by introducing a history function $\mathcal{H}$ which substitutes Ψ^+ to handle loading and unloading and defined as:

$$\mathcal{H}(\mathbf{x}, t) = \max_{\tau \in [0,t]} \left\{\tilde{\Psi}^+(\mathbf{x}, \tau)\right\}. \tag{A19}$$

It leads to:

$$\left(2\mathcal{H} + \frac{G_c}{\ell}\right) d - \ell G_c \Delta d = 2\mathcal{H}, \quad \nabla d \cdot \mathbf{n} = 0 \text{ on } \partial\Omega, \tag{A20}$$

and to the corresponding weak form: find $d \in H^1(\Omega)$ such that

$$\int_\Omega \left(2\left[\mathcal{H}\right] + \frac{g_c}{\ell}\right) d\delta d + g_c \ell \nabla d \cdot \nabla(\delta d) d\Omega = \int_\Omega 2\left[\mathcal{H}\right] \delta d\, d\Omega \tag{A21}$$

for all $\delta d \in H_1(\Omega)$. The problem can be solved by classical Finite Elements (see e.g. [70]).