Universitat Politècnica de Catalunya
Escola Tècnica Superior d'Enginyers
de Camins, Canals i Ports

DEPARTAMENT D'ENGINYERIA CIVIL I AMBIENTAL (DECA)

INTERNATIONAL CENTER FOR NUMERICAL METHODS IN ENGINEERING

PhD Thesis:

COMPUTATIONAL MULTISCALE MODELING OF FRACTURE PROBLEMS AND ITS MODEL ORDER REDUCTION

By

Manuel Alejandro Caicedo Silva

Advisors:

Prof. Xavier Oliver

Dr. Alfredo E. Huespe

Barcelona, January 2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola de Doctorat

Curso académico: 2016-2017

Acta de calificación de tesis doctoral

Nombre y apellidos
MANUEL ALEJANDRO CAICEDO SILVA

Programa de doctorado
ANALISIS ESTRUCTURAL

Unidad estructural responsable del programa
DEPARTAMENT D'ENYINYERIA CIVIL I AMBIENTAL (DECA)

Resolución del Tribunal

Reunido el Tribunal designado a tal efecto, el doctorando / la doctoranda expone el tema de su tesis doctoral titulada COMPUTATIONAL MULTISCALE MODELING OF FRACTURE PROBLEMS AND ITS MODEL ORDER REDUCTION.

Acabada la lectura y después de dar respuesta a las cuestiones formuladas por los miembros titulares del tribunal, éste otorga la calificación:

☐ NO APTO ☐ APROBADO ☐ NOTABLE ☐ SOBRESALIENTE

(Nombre, apellidos y firma) Presidente/a		(Nombre, apellidos y firma) Secretario/a
(Nombre, apellidos y firma) Vocal	(Nombre, apellidos y firma) Vocal	(Nombre, apellidos y firma) Vocal

BARCELONA, _____ de _____ de 2017

El resultado del escrutinio de los votos emitidos por los miembros titulares del tribunal, efectuado por la Comisión Permanente de la Escuela de Doctorado, otorga la MENCIÓN CUM LAUDE:

☐ SÍ ☐ NO

(Nombre, apellidos y firma) Presidente/a de la Comisión Permanente de la Escuela de Doctorado	(Nombre, apellidos y firma) Secretario/a de la Comisión Permanente de la Escuela de Doctorado

BARCELONA, _____ de _____ de 2017

RESUMEN

Esta Tesis se enfoca en el modelamiento numérico de la fractura y su propagación en materiales heterogéneos sujetos a degradación, mediante modelos multiescala jerárquicos basados en la técnica FE^2, abordando a su vez la problemática del coste computacional excesivo mediante el desarrollo, la implementación y validación de un conjunto de herramientas computacionales basadas en técnicas de modelos reducidos.

Para la modelización de problemas de fractura, se desarrolló, implementó y validó, un modelo multiescala de fallo con las siguientes características:

- En la macroescala, se adaptaron los últimos avances de la *Aproximación de Discontinuidades Fuertes del Continuo (CSDA)*, hasta el momento desarrollados para modelos monoescala. Se formula un nuevo elemento finito con alta capacidad de capturar y modelar localización de deformaciones en bandas que pueden interceptar al elemento finito en direcciones arbitrarias. Para evaluar la dirección de propagación de falla se utiliza la técnica del *crack-path field* (Oliver et al., 2014).
- En la microescala, en aras de usar mecanismos de fallo que sean simples de implementar y calcular, y de cara al posterior desarrollo de una formulación de modelos reducidos, se propone el uso de elementos cohesivos tipo banda, equipados con un modelo constitutivo de daño isótropo regularizado, capaz de representar la decohesión del material. Estos elementos cohesivos son distribuidos entre los diferentes componentes de la microestructura, y en sus fronteras, cumpliendo a su vez el papel de elementos de interfase.

Se verificó la objetividad de los resultados del modelo con respecto al tamaño de la *celda de fallo*, y al tamaño del elemento finito de la macroestructura. De igual forma, también se verificó la consistencia del modelo multiescala por medio de la comparación de resultados obtenidos con el mismo, y soluciones obtenidas mediante *Simulaciones Numéricas Directas (DNS)*.

En cuanto a la reducción del coste computacional en los modelos multiescala jerárquicos del tipo FE^2. Se propone reformular el problema de valores de contorno de la microescala, mediante el uso de dos técnicas sucesivas de reducción, definidas como *Modelo de Orden Reducido MOR* y *Modelo de Orden Hiper-reducido (HROM ó HPROM)*, respectivamente.

En primer lugar, para el Modelo de Orden Reducido, el problema de elementos finitos estándar (*de alta fidelidad*) de la microestructura, es proyectado y resuelto

en un subespacio de menor dimensión mediante el método de la *Descomposición Ortogonal Propia (POD)*. En segundo lugar, para el desarrollo del Modelo de Orden Hiper-reducido, dos técnicas han sido estudiadas y desarrolladas, a saber: los métodos de *interpolación* y los métodos de *Cuadratura de Orden Reducido ROQ* (An et al., 2009). Las bases reducidas para la proyección de las *variables primales*, son calculadas por medio de la *Descomposición en Valores Singulares SVD* de *snapshots* captados de trayectorias de *entrenamiento* previamente definidas.

Para modelar problemas de materiales heterogéneos caracterizados por relaciones constitutivas que poseen endurecimiento material, las fluctuaciones de desplazamiento y las tensiones de la microestructura fueron seleccionadas como variables primales para la primera y segunda reducción, respectivamente. En este caso, la segunda reducción se realizó por medio de la interpolación del campo de tensiones. Sin embargo, se puede observar que el operador de proyección de las tensiones, al ser calculado a partir de la selección de *snapshots* de pasos numéricamente convergidos, produce un malcondicionamiento del problema. Este malcondicionamiento es estudiado en profundidad y es corregido para dar lugar a una solución robusta y consistente.

Para la reducción en modelos de fractura, se propuso como punto de partida la formulación multiescala de fallo desarrollada en esta Tesis. Al igual que en la modelización de problemas de materiales que poseen endurecimiento material, se preserva el esquema de dos reducciones sucesivas.

Teniendo en cuenta el caracter discontinuo de las deformaciones en la microescala en problemas de fractura. Se propone la descomposición de la celda de fallo en dos partes, un dominio *cohesivo* que contiene la totalidad de bandas cohesivas, y un dominio *regular* compuesto por el dominio disjunto remanente de la microcelda. Cada uno de estos dominios tiene un tratamiento individual. El modelo de la microescala, es reformulado como un problema de *punto de silla* en el cual se busca minimizar el potencial de energía libre, sujeto a restricciones para que cumpla los postulados básicos de la modelización multiescala.

En una primera reducción, se propone como variable primal el campo de deformaciones fluctuantes, que por medio del método de la *Descomposición Ortogonal Propia (POD)*, se proyecta y resuelve el problema en un espacio de dimensión reducida. La segunda reducción se basa en integrar los términos que resultan del problema variacional de equilibrio mediante una *Cuadratura de Orden Reducido (ROQ)*, conformada por un conjunto de puntos de integración, considerablemente inferior con respecto a la cantidad de puntos de integración requeridos por la cuadratura de Gauss convencional en el modelo de *alta fidelidad*.

Esta metodología de hiperreducción mediante cuadraturas de orden reducido (ROQ), provó ser considerablemente mas eficiente y robusta que los métodos de interpolación, siendo además, aplicable a problemas multiescala con endurec-

imiento.

Para la validar los modelos reducidos desarrollados en esta Tesis, se realizaron múltiples pruebas variando la cantidad de bases para ambas reducciones, observando la convergencia del modelo reducido con respecto a la respuesta del modelo de alta fidelidad, incrementando la cantidad de modos y puntos de integración. Igualmente se puede concluir que, para errores admisibles (inferiores al 5%), las aceleraciones del costo computacional involucrado respecto al tiempo requerido por el modelo de alta fidelidad (speed-up) son del orden de 110 veces. Lo que constituye aceleraciones considerablemente superiores a lo reportado por la literatura.

ABSTRACT

This Thesis focuses on the numerical modeling of fracture and its propagation in heterogeneous materials by means of hierarchical multiscale models based on the FE^2 method, addressing at the same time, the problem of the excessive computational cost through the development, implementation and validation of a set of computational tools based on reduced order modeling techniques.

For fracture problems, a novel multiscale model for propagating fracture has been developed, implemented and validated. This multiscale model is characterized by the following features:

- At the macroscale level, were adapted the last advances of the *Continuum Strong Discontinuity Approach (CSDA)*, developed for monoscale models, devising a new finite element exhibiting good ability to capture and model strain localization in bands which can be intersect the finite element in random directions; for failure propagation purposes, the adapted *Crack-path field* technique (Oliver et al., 2014), was used.
- At the microscale level, for the sake of simplicity, and thinking on the development of the reduced order model, the use of cohesive-band elements, endowed with a regularized isotropic continuum damage model aiming at representing the material decohesion, is proposed. These cohesive-band elements are distributed within the microscale components, and their boundaries.

The objectivity of the solution with respect to the *failure cell* size at the microscale, and the finite element size at the macroscale, was checked. In the same way, its consistency with respect to *Direct Numerical Simulations (DNS)*, was also tested and verified.

For model order reduction purposes, the microscale *Boundary Value Problem (VBP)*, is rephrased using *Model Order Reduction* techniques. The use of two subsequent reduction techniques, known as: *Reduced Order Model (ROM)* and *HyPer Reduced Order Model (HPROM or HROM)*, respectively, is proposed.

First, the standard microscale finite element model *High Fidelity (HF)*, is projected and solved in a low-dimensional space via *Proper Orthogonal Decomposition (POD)*. Second, two techniques have been developed and studied for multiscale models, namely: a) *interpolation* methods, and b) *Reduced Order Cubature (ROQ)* methods (An et al., 2009). The reduced bases for the projection of the *primal variables*, are computed by means of a judiciously training, defining a set of pre-defined

training trajectories.

For modeling materials exhibiting hardening behavior, the microscale displacement fluctuations and stresses have been taken as *primal variables* for the first and second reductions, respectively. In this case, the second reduction was carried out by means of the stress field interpolation. However, it can be shown that the stress projection operator, being computed with numerically converged snapshots, leads to an ill-possed microscale reduced order model. This ill-poseddness is deeply studied and corrected, yielding a robust and consistent solution.

For the model order reduction in fracture problems, the developed multiscale formulation in this Thesis was proposed as point of departure. As in hardening problems, the use of two successive reduced order techniques was preserved.

Taking into account the discontinuous pattern of the strain field in problems exhibiting softening behavior. A domain separation strategy, is proposed. A *cohesive* domain, which contains the cohesive elements, and the *regular* domain, composed by the remaining set of finite elements. Each domain has an individual treatment. The microscale Boundary Value Problem (BVP) is rephrased as a *saddle-point* problem which minimizes the potential of free-energy, subjected to constraints fulfilling the basic hypotheses of multiscale models.

The strain flucuations are proposed as the primal variable for the first reduction, where the high fidelity model is projected and solved into a low-dimensional space via *POD*. The second reduction is based on integrating the equilibrium equations by means of a *Reduced Order Quadrature (ROQ)*, conformed by a set of integration points considerably smaller than the classical Gauss quadrature used in the *high fidelity* model.

This methodology had been proven to be more robust and efficient than the interpolation methods, being applicable not only for softening problems, but also for hardening problems.

For the validation of the reduced order models, multiple test have been performed, changing the size of the set of reduced basis functions for both reductions, showing that convergence to the high fidelity model is achieved when the size of reduced basis functions and the set of integration points, are increased. In the same way, it can be concluded that, for admissible errors (lower than 5%), the reduced order model is $\sim$ 110 times faster than the high fidelity model, considerably higher than the speedups reported by the literature.

CONTENTS

ACRONYMS

BVP	Boundary Value Problem
CSDA	Continuum Strong Discontinuity Approach
DNS	Direct Numerical Simulation
EFEM	Embedded Finite Element Methodology
EIM	Empirical Interpolation Method
FE	Finite Element Method
FE^2	FE$\times$FE hierarchical multiscale technique
HF	High Fidelity
HPROM	High-Performance Reduced Order Model
HROM	Hyper-Reduced Order Model
MOR	Model Order Reduction
POD	Proper Orthogonal Decomposition
ROM	Reduced Order Model
ROQ	Reduced Optimal Quadrature
RUC	Repeating Unit Cell
RVE	Representative Volume Element
SVD	Singular Value Decomposition

Part I

RESEARCH SUMMARY

1

INTRODUCTION

1.1 State of the Art

1.1.1 *Multiscale modeling of heterogeneous materials*

During the last decades, a large variety of multiscale strategies focusing on the study and analysis of the mechanical behavior of heterogeneous materials, have been proposed by the computational mechanics community. Based on the work of (Böhm, 2013), these strategies may be divided into three main groups:

- *Mean-Field approaches*: based on the seminal contributions developed by (Eshelby, 1957) and (Mori and Tanaka, 1973). In these approaches, the microfields within each constituent of an heterogeneous material, are approximated by their phase averages, typically, *phase-wise* uniform stress and strain fields are employed. Recently, the application of these approaches to nonlinear modeling in composites has become a subject of active research.

- *Bounding Methods:* Variational principles are used to obtain upper and lower bounds of the overall mechanical properties (e. g. elastic tensors, secant moduli, yield thresholds, among others). Bounding Methods are closely related with Mean-Field approaches, because many analytical bounds are obtained on the basis of phase-wise constant fields.

  The formal treatments were provided by, i. e., (Nemat-Nasser and Hori, 1999; Bornert, 1999; Ponte Castañeda and Suquet, 1998; Markov, 2000; Milton, 2002; Torquato, 2002). Two of the most relevant results of this kind of models are, the *upper bounds* of (Voigt, 1889), and the *lower bounds* of (Reuss, 1929). Posteriorly, Hill extends those bounds to tensorial entities, particularly, for constitutive tangent tensors (Hill, 1952).

- *Representative Volume Element (RVE) based approaches:* these approximations are based on studying discrete microstructures, aiming at evaluating the microscale fields, fully accounting for the interactions between different phases. *Homogenization* is used as a strategy to upscale the resulting averaged variables. This homogenization strategy uses *representative volumes*, which copy as much as possible the material heterogeneities. These representative volume elements must be large enough to capture a statistically

representative solution of the material behavior, but, also their size must be limited in comparison with the macroscopic characteristic length - (separation of scales).

The hypothesis of these approaches are properly fulfilled if there is a marked scale separation between the phenomena observed at the macroscale, and the ones observed at the microscale. However, nowadays, new approaches have been proposed to overcome this limitation.

Models based on the existence of a RVE can be divided into two main groups:

- *Hierarchical models:* the RVE, subjected to consistent boundary conditions[1] is used to obtain a detailed microscale response. The link between scales is reached by means of an energetic identity, such as the Hill-Mandel Principle of Macro-Homogeneity (Hill, 1965; Mandel, 1971), or even on more general approaches, like the Principle of Multiscale Virtual Power (Blanco et al., 2016). In the context of the Finite Element Method (FEM), this methodology is known as FE×FE hierarchical multiscale technique (FE^2) (Feyel and Chaboche, 2000).

  In virtue of the potential applications in microstructures with complex morphologies, the FE^2 technique is suitable to deal with problems facing material instabilities, like fracture processes. In this sense, some approaches have been proposed (Belytschko et al., 2008; Song and Belytschko, 2009; Nguyen et al., 2010a,b; Matous et al., 2008; Coenen et al., 2012; Toro et al., 2014), among others. One of the main contributions of this Thesis, is an alternative approach with marked differences with respect to the previous ones.

- *Concurrent models:* the RVE is embedded into the macroscale geometry, and the corresponding kinematic compatibility is guaranteed via Lagrange multipliers, similarly to Domain Decomposition approaches.

  Like the above mentioned models, concurrent models are widely applied. In approaches based on the Finite Element Method, the embedded meshes are not necessarily compatible. However, its computational viability is only for cases with small scale separation, this feature becomes into its main disadvantage. Some concurrent approaches in the field of fracture mechanics have been recently developed, i. e., (Lloberas-Valls et al., 2012).

1.1.2 *Fracture mechanics*

1.1.2.1 *Monoscale Fracture Approaches*

The study and analysis of fracture in solids has been a topic of research since the last century. The seminal works on this topic were focussed on the Elastic

1 Consistent in the sense that, all possible boundary conditions have to be compatible with the strains obtained at the macroscale.

Fracture Mechanics. However, its generalization to nonlinear material behavior is a non trivial task.

Starting from the concept of fracture energy, which has become a central issue in nonlinear fracture mechanics modeling, several techniques have been developed:

- *Cohesive models:* Based on the introduction of interfaces embedded into a continuum medium. These interfaces admit the development of displacements discontinuities. Cohesive forces across the interfaces act opposing to the crack opening, diminishing as material degradation takes place. The energy necessary to produce a crack is equal to the fracture energy. Some applications of this kind of models are found in (Pandolfi et al., 1999; Molinari et al., 2007; Toro et al., 2016; Rodrigues et al., 2016).
- *Continuum regularized models:* Characterized by a continuum constitutive law displaying a softening response. These kind of approaches are subjected to material instabilities and bifurcation processes, causing ill-possedness of the problem from the mathematical point of view. As a remedy to this flaw, a constitutive law regularization (localization limiters) is introduced, ensuring mesh objective solutions. Some proposed models can be found in (Pijaudier-Cabot and Bazant, 1987; Tvergaard and Needleman, 1995; Pijaudier-Cabot and Bazant, 1987; Aifantis, 1984; de Borst and Mühlhaus, 1992; Peerlings et al., 1996, 2001, 2002; Steinmmann and Willam, 1991; Muhlhaus and Vardoulakis, 1987), among others.
- *Continuum Strong Discontinuity Approach (CSDA):* This approach establishes a link between *cohesive models* and *continuum models*. Its fundamentals have been presented in the seminal work (Simo et al., 1993), posteriorly improved and applied to many applications in static and dynamic scenarios (Oliver, 2000; Oliver et al., 2002; Oliver and Huespe, 2004b,a; Oliver et al., 2014). This approach provides an unified theory, which goes from the continuum description to the degradation and posterior material failure exhibiting displacement discontinuities (cracks). In this context, the continuum constitutive model subjected to a kinematics inducing displacement discontinuities represents also a "projected" cohesive law on the crack surface.

1.1.2.2 *Multiscale Fracture Approaches*

The study of heterogeneous materials subjected to softening, and, therefore, to degradation and failure, through multiscale approaches brings additional challenges. The fundamental reason lies in two aspects: (a) it becomes imperative the use of regularized constitutive theories at both scales in order to ensure the well-possednes of the multiscale problem. (b) The *size effect*, intrinsically related to the fracture energy, and extensively studied by (Bazant and Planas, 1998). As a result of this, the homogenized stress tensor, in the post-citrical regime, becomes extremely sensible to the RVE size.

The second issue is the existence of the RVE (Gitman et al., 2007), and the fulfillment of the basic hypothesis in multiscale modeling.

The necessity to develop specific homogenization techniques, becomes a starting point for obtaining consistent multiscale formulations. (Belytschko et al., 2008) has proposed a methodology that excludes the localization domain in the homogenization process. More recently, (Song and Belytschko, 2009) proposed a predetermined size of RVEs. (Matous et al., 2008) describes a novel methodology, based on the existence of a macroscopic adhesive interface, which links the macroscale jump of displacements with an equivalent jump at the microscale, imposed by consistent boundary conditions.

Recent works (Nguyen et al., 2010a,b), describe the material failure by means of nonlocal gradient theories. In this kind of approaches, and, in contrast with other alternatives, the homogenization of the stress field during the post-critical regime, is carried out at the localization zone (this zone corresponds to a subdomain of the RVE). However, other authors claimed some inconsitencies related to this kind of approaches, particularly, about the fact that kinematics at the macroscale is not equivalent to the kinematics modeled at the microscale.

1.1.3 *Model Order Reduction*

In general, the FE^2 method involving fine space-time discretization and time-dependent homogenization procedures, involves an enormous computational cost, being even larger when facing fracture mechanics problems.

The idea of exploiting the combination of dimensionality reduction and multiscale modeling is certainly not new. A survey of the related literature reveals that, over the last decade, researchers from various scientific disciplines dealing with multiscale problems have begun to consider the model reduction as a potential route –complementary to improvements in software and hardware power –to diminish the often unaffordable cost of multiscale simulations. In the specific context of homogenization-based multiscale methods, the application of model reduction techniques has been addressed by several authors, namely, (Ganapathysubramanian and Zabaras, 2004; Yvonnet and He, 2007; Boyaval, 2007; Monteiro et al., 2008; Nguyen, 2008). The strategy adopted in all these works for constructing a cost efficient model of the micro-cell is the standard reduced basis method. The gist of this strategy is to project the governing equations onto a low-order subspace spanned by carefully chosen bases (Amsallem et al., 2009).

1.1.3.1 *Reduced basis techniques (ROM)*

Reduced basis methods, in its standard form, suffer from an important limitation when handling nonlinear problems: they reduce notably the number of degrees of freedom –and thus the pertinent equation solving effort–, yet the computational cost associated to the evaluation of the internal forces and jacobians

at quadrature points remains the same. Standard reduction methods prove, be effective only when dealing with micro-cells whose constituents obey simple constitutive laws (linear elasticity). In a general inelastic case, the calculation of the stresses at each gauss point is, on its own, a computationally expensive operation and dominates the total cost of the computation. As a consequence, the speed up provided by standard model reduction methods in nonlinear scenarios is practically negligible, and may not compensate the cost associated to the offline construction of the reduced-order bases.

1.1.3.2 *High-performance reduced order modeling techniques (HPROM)*

The origin of the first effective proposal on this issue can be traced back to the seminal work of (Barrault et al., 2004), who suggested to approximate the nonlinear term in the reduced-order equations by a linear combination of a few, carefully chosen basis functions. In the spirit of a offline/online strategy, in the standard reduced basis approach, these spatial bases are computed offline from full-order snapshots of the non-linear term, whereas the corresponding parameter-dependent modal coefficients are determined online by interpolation at a few (as many as basis functions), judiciously pre-selected spatial points. As in classical reduced bases methods, the efficiency of this second or collateral reduction is predicated on the existence of a moderate number $M \ll N$ (N is the original dimension of the problem) of basis functions whose span accurately approximate the manifold induced by the parametric dependence of the nonlinear contribution. The interpolation method developed by (Barrault et al., 2004) is known as the Empirical Interpolation Method (EIM); the main ingredients of this method are: a) the use of a greedy algorithm to generate a set of maximally independent bases from the collection of snapshots of the nonlinear term, on the one hand; and b) the recursive selection – also via a greedy algorithm – of spatial locations where the error between the full-order bases and their reconstructed counterparts is maximum[2].

In solution methods in which the governing equations are used in its variational form (as in the Finite Element Method (FE)), the reduction of the computational complexity arising from nonlinearities can be, alternatively, achieved by approximating the integrals in which the offending nonlinear function appears, rather than the function itself, as done in the interpolatory and least-square reconstruction techniques discussed above. Based on this observation, (An et al., 2009) propose a quadrature scheme devised for fast-run integration of the subspace spanned by a representative set of snapshots of the nonlinear integrand.

In what follows, we shall consider as equivalent the appellations *HPROM* and *Hyper-Reduced Order Model (HROM)* to refer to reduced basis methods combined with interpolatory or least-square reconstruction schemes.

2 Maximum in the sense that, the selected points have to be taken from components in which the error between the high fidelity and the High-Performance Reduced Order Model (HPROM) solutions is greater.

1.1.3.3 *Reduction Order Modeling in fracture problems*

The development of reduced models for non-homogeneous materials has been tackled in numerous previous contributions, such as (Michel et al., 2001), where the proposed reduction techniques are based on Fourier's transforms, or (Yvonnet and He, 2007), where a reduced model is applied the homogenization analysis of hyperelastic solids subjected to finite strains. Also, the work in (Ryckelynck, 2009) develops a hyper-reduced model of a monoscale analysis which consider nonlinear material behavior. However, the existing literature barely considers reduced order modeling of non-smooth problems, as is the case of fracture, where discontinuous displacements occur. The multiscale case, when fracture takes place at both scales of the problem, makes the task even much harder. Indeed, only very few contributions have been presented in the literature about this topic, see for example: (Oskay and Fish, 2007), which follows an eigendeformation-based methodology, or (Zhang and Oskay, 2016; Kerfriden et al., 2013) that resort to global–local approaches.

The previous approaches combine projection techniques and, in some cases, empirical criteria to integrate the equilibrium equations in the domain. However, these are ussually ad-hoc techniques, that had been applied to problems with relatively simple crack propagation schemes. Currently, some researchers consider the effective model order reduction of fracture processes, an insolvable problem. This Thesis will reconsider this statements, by developing a robust HPROM formulation, for multiscale fracture problems resulting in high computational speedups.

1.2 Adopted Approach

The approach adopted in this work, uses a FE method and multiscale hierarchical models. Particularly the FE^2 approach is used, which involves two scales (*macroscale* and *micro/meso-scale*) both discretized via finite elements. *Infinitesimal strain* setting, and *first-order homogenization* are assumed.

For fracture modeling purposes, the CSDA is adapted to the multiscale setting, and used for modeling propagating fracture at the macroscale level. At the microscale level, the use of predefined cohesive bands, distributed within the components and its interfaces, is proposed. These cohesive bands are endowed with regularized continuum damage models, which induce the crack initiation and propagation.

The Model Order Reduction techniques used in this Thesis, are based on the *Proper Orthogonal Decomposition (POD)*, defining the projection of the *full order* model into a low-dimension small space, and, on the use of novel *interpolation* and *Reduced Optimal Quadrature (ROQ)* schemes to diminish the computational cost generated by the multiscale problem.

1.3 Objectives and Scope

The main objectives of this Thesis are:

- *To develop a consistent and minimally intrusive multiscale hierarchical approach for propagating fracture with proper transfer of energy across scales.*
- *To develop, implement and validate a set of computational tools to efficiently reduce the unaffordable computational cost associated to the* FE^2 *approach.*

1.4 Outline

The remainder of this manuscript is organized in four chapters. Chapter 2 is devoted to the derivation of the multiscale model for propagating fracture, including a brief introduction to the fundamentals of the computational homogenization used in the proposed approach. Chapter 3 deals with the derivation of reduced order models for multiscale, smooth and non-smooth (fracture), problems, as well as some numerical results obtained from the developed models. Chapter 4 provides some concluding remarks and identifies areas for future research. In Appendix A, the participations in national and international conferences, and specialized workshops are listed. In Appendix B, a short summary of the supporting papers is presented. Finally, in Appendix C, the scientific publications supporting this Thesis, and co-authored by the author, are annexed.

2

MULTISCALE MODELING APPROACH TO FRACTURE PROBLEMS

2.1 Computational Homogenization

In the context of two scale (macroscale – micro/mesoscale) problems, computational homogenization of materials is generally regarded as a way of obtaining point-wise stress–strain constitutive models at the macroscale, accounting for complex micro/mesoscopic material morphology.

The homogenization approach used in this Thesis –commonly know as *first-order* homogenization– is only valid for materials that display either *statistical homogeneity* or *spatial periodicity*.

In consequence, depending on the morphology and random distribution of constituents at the microscale, the definition and existence of a *representative sample* RVE plays an important role in the material characterization of heterogeneities at the macroscale.

This representative sample, hereafter denoted $\mathcal{B}_\mu \in \mathbb{R}^d (d = 2, 3)$, is assumed to exhibit several features. One of those corresponds to the *size indifference* property (Terada et al., 2000; Kouznetsova, 2002; Drago and Pindera, 2007), which states that if the size of this sample is increased, the response remains identical regardless the *admissible* boundary conditions on the RVE. The lower size limit for the RVE satisfying the size indifference property is represented by the *characteristic length-scale* denoted as h_μ, giving rise to the existence of the RVE, whereas in microstructures that display periodicity, is known as *Repeating Unit Cell (RUC)*, or simply *unit cell*. Furthermore, $\mathcal{B}_\mu$ has to be small enough to be regarded as a point at the macroscale (Gross and Seelig, 2011) (i. e., $h_\mu \ll L$, being L the characteristic length of the macroscale $\mathcal{B}$, see Fig. 1) this is the so-called *scale separation* hypothesis.

This section presents a summary of the multiscale variational formulation used in this Thesis. This approach is based on the following fundamental hypotheses:

- The infinitesimal strains setting is used.
- Quasiestatic problems are considered[1]. A monotonically increasing pseudo-time variable t is used, $t \in [0, t_f]$, being t_f the final time of the analysis. In the development of the multiscale model for fracture modeling, the incremental form of the equilibrium equations is used, due to the fact that the kinematic enhancement is modified along time. This issue is fully detailed in Sec. 2.2.
- The multiscale problem is restricted to two scales, although it can be easily extended to additional scales. The *macroscale*, usually identified as the structural scale, is denoted as $\mathcal{B}$, and its material points are denoted as $\boldsymbol{x}$. The representative sample $\mathcal{B}_\mu$ is defined as *meso/microscale*, in which, every material point is denoted by $\boldsymbol{y}$. In addition, for the sake of clarity, entities at the small scale, are identified by the subindex μ, see Fig. 1.
- The body at the macroscale, $\mathcal{B}$, is subjected to predefined force or displacement actions, applied along its boundary $\boldsymbol{\Gamma}$. This boundary $\boldsymbol{\Gamma}$ is supposed to be smooth by parts, and it can be splitted into two parts, $\boldsymbol{\Gamma}_D$ and $\boldsymbol{\Gamma}_N$ where $\boldsymbol{\Gamma}_D \cap \boldsymbol{\Gamma}_N = \emptyset$, representing the domains in which *Dirichlet* and *Newman* boundary conditions are imposed.
- Every point $\boldsymbol{x}$ at the macroscale is related with a corresponding heterogeneous microstructural representative domain (the RVE), assuming the existency of a *scale separation*, so that the representative length at the microscale h_μ is considerably smaller than that representing the macroscale L ($h_\mu \ll L$).
- The body at the macroscale, $\mathcal{B}$, is idealized as a statistically homogeneous media, in which the mechanical state at a generic point $\boldsymbol{x}$ is characterized by the strain and stress tensors, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$, respectively. The macroscale strain tensor $\boldsymbol{\varepsilon}$ is the input variable for the microscale, the corresponding output variables are the homogenized stress $\boldsymbol{\sigma}$ and the homogenized tangent constitutive $\mathbb{C}$ tensor ($\dot{\boldsymbol{\sigma}} = \mathbb{C} : \dot{\boldsymbol{\varepsilon}}$). In this way, the multiscale model can be interpreted as a constitutive model that, given a strain history, returns the stress and tangent constitutive tensors histories, accounting for the morphology and the interaction of the different components at the microscale, see Fig. 1.
- The *Hill-Mandel Variational Principle of Macro-Homogeneity* (Hill, 1965; Mandel, 1971), which states an equivalence between the virtual power densities between micro- and macroscale, and requires the adoption of specific kinematically admissible displacement fluctuations at the RVE, is adopted.

1 However, in one article supporting this Thesis, dynamic problems are also considered. See (Hernández et al., 2017)

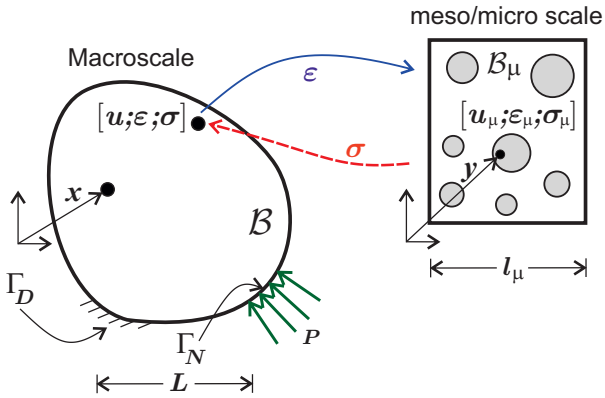


Figure 1: Macrosctructure with an embedded local microstructure.

2.1.1 *RVE kinematics and strain tensor*

In the context of the adopted first-order homogenization setting, the microscopic velocity field $\dot{\boldsymbol{u}}_\mu$ can be splitted as follows:

$$\dot{\boldsymbol{u}}_\mu = \dot{\boldsymbol{u}}(x,t) + \dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t)\cdot\boldsymbol{y} + \dot{\tilde{\boldsymbol{u}}}_\mu(\boldsymbol{y},t) \tag{2.1}$$

where $\dot{\boldsymbol{u}}$ stands for the velocity at the macroscale, $\dot{\boldsymbol{\varepsilon}}$ stands for the rate of infinitesimal macroscopic strain tensor, the term $\dot{\boldsymbol{\varepsilon}}\cdot\boldsymbol{y}$ is a velocity term that varies linearly with $\boldsymbol{y}$, and $\dot{\tilde{\boldsymbol{u}}}_\mu$ the *velocity fluctuations*. The decomposition of the rate of microscopic strain tensor $\dot{\boldsymbol{\varepsilon}}_\mu$ in the finite element framework yields, from the spatial differentiation of Eq. (2.1):

$$\dot{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{x},\boldsymbol{y},t) = \dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t) + \boldsymbol{\nabla}^s_{\boldsymbol{y}}\dot{\tilde{\boldsymbol{u}}}_\mu(\boldsymbol{y},t) \tag{2.2}$$

The starting point of multi-scale constitutive settings, is the assumption that the rate of macroscopic strain $\dot{\boldsymbol{\varepsilon}}$, at a point $\boldsymbol{x}$ of the macro-continuum, is the *volume average* of the rate of microscopic strain (2.2), over the RVE associated with $\boldsymbol{x}$. This assumption is also interpreted as the fact that the microscale deformations only influence the macroscale behavior through its volume average.

$$\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t) = \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu}\dot{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},t)\,dV \tag{2.3}$$

In virtue of (2.2) and (2.3), this condition is equivalent to impose the volume average of the symmetric gradient of the velocity fluctuations $\boldsymbol{\nabla}^s_{\boldsymbol{y}}\dot{\tilde{\boldsymbol{u}}}_\mu$ to vanish. This condition can be written using the Gauss theorem as a constraint over $\dot{\tilde{\boldsymbol{u}}}_\mu$, involving the whole volume of the RVE, as follows:

$$\dot{\tilde{\boldsymbol{u}}}_\mu \in \tilde{\mathcal{V}}^u_\mu \qquad \text{where} \qquad \tilde{\mathcal{V}}^u_\mu := \left\{ \dot{\tilde{\boldsymbol{u}}}_\mu \;\middle|\; \int_{\Gamma_\mu}\dot{\tilde{\boldsymbol{u}}}_\mu \otimes^s \boldsymbol{\nu}_\mu\,d\Gamma = \boldsymbol{0} \right\} \tag{2.4}$$

where $\tilde{\mathcal{V}}^u_\mu$ is defined as the space of admissible microscale velocity fluctuations in the RVE, Γ_μ stands for the boundary of the domain $\mathcal{B}$, and $\boldsymbol{\nu}_\mu$ is the unit normal

vector on Γ_μ. Equation (2.4) is also known as the *minimum constraint boundary condition*.

The actual set of kinematically admissible velocity fields $\tilde{\mathcal{V}}_\mu$, together with the associated space of *virtual kinematically* admissible velocities at the microscale, denoted by $\mathcal{W}_\mu^u$, play a fundamental role in the variational formulation of the equilibrium problem of the microscale. This space $\mathcal{W}_\mu^u$ can be defined as follows:

$$\mathcal{W}_\mu^u := \left\{ \boldsymbol{\eta} = \boldsymbol{v}_1 - \boldsymbol{v}_2 \quad | \quad \boldsymbol{v}_1, \boldsymbol{v}_2 \in \tilde{\mathcal{V}}_\mu \right\} \tag{2.5}$$

In virtue of (2.4), and the fact that $\tilde{\mathcal{V}}_\mu^u$ is itself a vector space, it can be concluded from (2.4) that:

$$\mathcal{W}_\mu^u = \tilde{\mathcal{V}}_\mu^u \tag{2.6}$$

Furthermore, the same arguments can be applied to the *total* form, and establish that any *kinematically admissible displacement fluctuation* $\tilde{\boldsymbol{u}}_\mu$ belongs also to $\tilde{\mathcal{V}}_\mu^u$.

2.1.2 *Hill-Mandel Principle of Macro-Homogeneity*

The scale bridging equations are completed by introducing the *Hill-Mandel Principle of Macro-Homogeneity* (Hill, 1965; Mandel, 1971). Based on physical arguments, this Principle states that the macroscopic stress power equates the volume average over the RVE of the microscopic stress power, making both, macroscale and microscale, continuum descriptors *energetically equivalent*. Thus, departing from:

$$\delta\boldsymbol{\varepsilon}_\mu = \delta\boldsymbol{\varepsilon} + \boldsymbol{\nabla}_{\boldsymbol{y}}^s \delta\tilde{\boldsymbol{u}}_\mu \qquad \forall \delta\boldsymbol{\varepsilon} \in \mathcal{E} \quad \forall \delta\tilde{\boldsymbol{u}}_\mu \in \tilde{\mathcal{V}}_\mu \tag{2.7}$$

where $\mathcal{E}$, stands for the space of all second order macroscopic strain tensor functions, Eq. (2.7) is similar to Eq. (2.2), but for admissible strain variations. Therefore, the following identity holds:

$$\boldsymbol{\sigma} \cdot \delta\boldsymbol{\varepsilon} \;=\; \frac{1}{\Omega_\mu} \int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu : \delta\boldsymbol{\varepsilon}_\mu \, d\mathcal{B}_\mu \qquad \forall \delta\boldsymbol{\varepsilon}_\mu \tag{2.8}$$

In particular, taking $\delta\tilde{\boldsymbol{u}}_\mu = \boldsymbol{0}$, and $\forall \delta\boldsymbol{\varepsilon} \in \mathcal{E}$, yields:

$$\boldsymbol{\sigma} = \frac{1}{\Omega_\mu} \int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\boldsymbol{y}, t) \, d\mathcal{B}_\mu \tag{2.9}$$

where, $\boldsymbol{\sigma}$ stands for the *macroscopic stress tensor*, which turns out to be as the volume average of the microscopic stress $\boldsymbol{\sigma}_\mu$. Equation (2.9) is also fulfilled in rate form. In addition to Eq. (2.9), the following condition emerges from the variational equation (2.8) solving for $\delta\boldsymbol{\varepsilon} = \boldsymbol{0}$:

$$\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu : \boldsymbol{\nabla}_{\boldsymbol{y}}^s \delta\tilde{\boldsymbol{u}}_\mu \, d\mathcal{B}_\mu = \boldsymbol{0} \qquad \forall \delta\tilde{\boldsymbol{u}}_\mu \in \tilde{\mathcal{V}}_\mu \tag{2.10}$$

Eq. (2.10) defines the variational microscale equilibrium problem (or *microscale virtual power principle*).

2.2 Multiscale Fracture Mechanics issues

In Computational Fracture Mechanics, hierarchical multiscale methods involve additional issues. In particular:

- The existence of the RVE has been questioned in the literature, arguing that for fracture cases, the material loses its statistical homogeneity (Gitman et al., 2007; Nguyen et al., 2010a).
- The fact that the homogenized constitutive model lacks an internal length (Bazant, 2010), raising similar issues than in classical phenomenological monoscale problems.

Additionally, mesh-bias dependence, and the proper fracture energy dissipation issues (Rots, 1988) via regularized constitutive models (Oliver, 1989; Oliver et al., 2002; Oliver and Huespe, 2004a; Oliver et al., 2015) are also crucial issues to be considered at each scale.

Along this section, the most important aspects of the proposed multiscale approach are summarized. This multiscale approach is fully detailed in Paper #2 in Sec. B.2.

2.2.1 *Multiscale modeling setting*

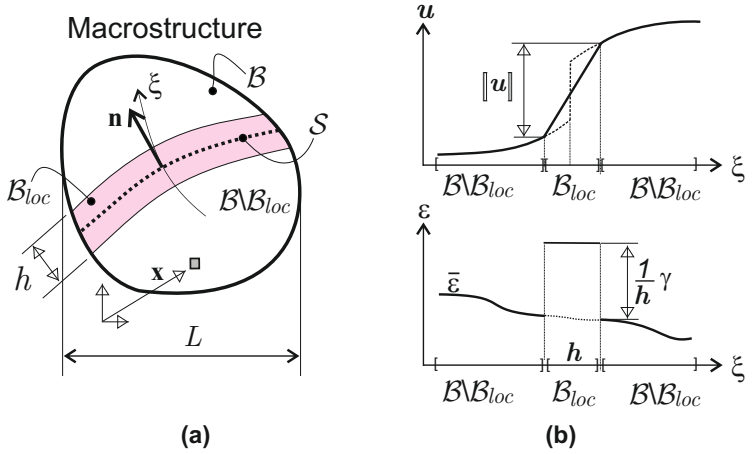


Figure 2: Macroscopic (Structural scale) body $\mathcal{B}$ (a) subdivision in a non-smooth domain $\mathcal{B}_{\text{loc}}(t)$, and a smooth domain $\mathcal{B} \setminus \mathcal{B}_{\text{loc}}(t)$ (b) h-regularized displacement and strain discontinuity kinematics.

2.2.1.1 *Macroscale Model*

Considering the body $\mathcal{B}$, at the macroscale (see Fig. 2) it is assumed that material points, $\boldsymbol{x}$, of the macroscopic body belong, at the current time t, to either one of the two subdomains:

- Domain $\mathcal{B} \setminus \mathcal{B}_{\text{loc}}(t)$: the set of points at the macroscale, exhibiting smooth behavior. The infinitesimal strain field $\boldsymbol{\varepsilon}(\boldsymbol{x}, t)$ is described in rate form, as:

$$\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x}, t) = (\boldsymbol{\nabla}_{\boldsymbol{x}} \otimes \dot{\boldsymbol{u}}(\boldsymbol{x}, t))^{s} \equiv \boldsymbol{\nabla}^{s}_{\boldsymbol{x}} \dot{\boldsymbol{u}}(\boldsymbol{x}, t) \qquad \forall \boldsymbol{x} \in \mathcal{B} \setminus \mathcal{B}_{\text{loc}}(t) \tag{2.11}$$

  where $\boldsymbol{u}(\boldsymbol{x}, t)$ is the macroscopic displacement field, t stands for the time or *pseudo-time* parameter, and $(\cdot)^{s}$, stands for the symmetric counterpart of $(\cdot)$.

- Domain $\mathcal{B}_{\text{loc}}(t)$: the set of points exhibiting material failure and, therefore, a non-smooth behavior. The strain field at these points is assumed to be captured by a h-regularized *strong-weak* discontinuity kinematics, h being the width of the corresponding strain localization band (see Fig. 2-(b)).

See: ***Sec.2.1*** ***Paper #2***

$$\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x}, t) = \dot{\tilde{\boldsymbol{\varepsilon}}}(\boldsymbol{x}, t) + \frac{\kappa_{\mathcal{B}_{\text{loc}}}(\boldsymbol{x})}{h} \dot{\boldsymbol{\gamma}}(\boldsymbol{x}, t) = \dot{\tilde{\boldsymbol{\varepsilon}}}(\boldsymbol{x}, t) + \delta^{h}_{s}(\boldsymbol{x}) \dot{\boldsymbol{\gamma}}(\boldsymbol{x}, t) \quad \forall \boldsymbol{x} \in \mathcal{B}_{\text{loc}}(t) \tag{2.12}$$

  In Eq. (2.12), $\dot{\tilde{\boldsymbol{\varepsilon}}}(\boldsymbol{x}, t)$ stands for the regular (smooth) counterpart of the strain, $\kappa_{\mathcal{B}_{\text{loc}}}$ is a colocation (*characteristic*) function on $\mathcal{B}_{\text{loc}}$ (See. Fig. 3), so that the term $\kappa_{\mathcal{B}_{\text{loc}}}(\boldsymbol{x})$ becomes a h-regularized Dirac's delta function shifted to the center-line, $S(t)$ (the macroscopic discontinuity-path at the current time t, as shown in Fig. 2-(a)). Thus, in Eq. (2.12), the term $\delta^{h}_{s}(\boldsymbol{x}) \dot{\boldsymbol{\gamma}}(\boldsymbol{x}, t)$ corresponds to the non-smooth (discontinuous and h-regularized) localized counterpart of the strains; a space-discontinuous second order tensor for the *weak-discontinuity case*.

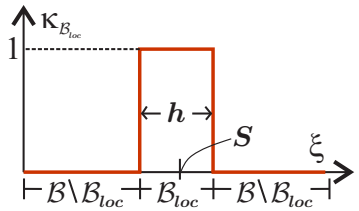


Figure 3: Colocation function $\kappa_{\mathcal{B}_{\text{loc}}}(\boldsymbol{x})$

2.2.1.2 *Microscale Model*

Assuming that the fracture at the macroscale has arisen, in turn, by the appearance of failure mechanisms at the microscale level, originated by some type of material failure. The next step is to endow the microscale model with

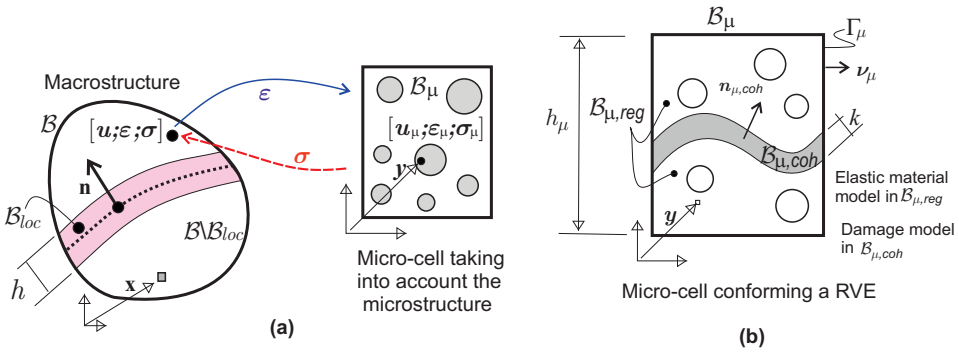


Figure 4: Outline of the multiscale model for propagating fracture: **a)** macro and micro scales; **b)** microcell model accounting for material failure.

mechanisms to capture the onset and propagation of this material failure. Therefore, without introducing further details, it is considered that the microstructure shall be able to capture some dominant failure mechanisms of the material.
For this purpose, a micro *failure cell* $\mathcal{B}_\mu$, of characteristic size h_μ, is considered to exist at every material point $\boldsymbol{x} \in \mathcal{B}$. It accounts for the material morphology at the microscale (voids, inclusions etc.). In addition, it is endowed with a set of *cohesive bands* ($\mathcal{B}_{\mu,\text{coh}} \subset \mathcal{B}_\mu$) of very small width $k \ll h_\mu$, whose position and other geometric properties (typically the normal n_μ, see Fig. 4) are predefined. At the current time t, the activation (de-cohesion) of a number of those bands, defines the current subset of active bands $\mathcal{B}_{\mu,\text{act}}(t) \subset \mathcal{B}_{\mu,\text{coh}} \subset \mathcal{B}_\mu$ which constitutes the "active" microscopic failure mechanism, for the considered point $\boldsymbol{x} \in \mathcal{B}$.

In principle, there is no intrinsic limitation on the number of the "candidate" cohesive bands to be considered at the failure cell. On one hand, their number and spatial position have to be sufficient to capture the dominant material failure mechanisms at the macroscale. On the other hand, the associated computational cost sets a limitation on the number of such bands. In this context, the following domains at the microscale are considered (see Fig. 4):

- Domain $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}}$: the set of points $\boldsymbol{y}$, which do not belong to the cohesive bands. They are compelled to exhibit a smooth behavior described by a *Continuum hardening model*, typically:

$$\dot{\boldsymbol{\sigma}}_\mu = \Sigma^{\text{hard}}(\dot{\boldsymbol{\varepsilon}}_\mu) \equiv \mathbb{C}_\mu^{\text{hard}} : \dot{\boldsymbol{\varepsilon}}_\mu \qquad \forall \boldsymbol{x} \in \mathcal{B}; \quad \forall \boldsymbol{y} \in \mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}} \tag{2.13}$$

  where $\boldsymbol{\sigma}_\mu$ and $\boldsymbol{\varepsilon}_\mu$ stand, respectively, for the stress and strain fields at the microscale point, $\boldsymbol{y}$, of the failure cell (corresponding to the macroscale point $\boldsymbol{x}$), being $\mathbb{C}_\mu^{\text{hard}}$ the microscopic inelastic constitutive tensor derived from the hardening constitutive model.

- Domain $\mathcal{B}_{\mu,\text{coh}}$: the set of microscale cohesive bands. As for the material behavior, the disctintion of two situations has to be made in this case:

– The failure cell, $\mathcal{B}_\mu$ is associated to a non-smooth material point at the macroscale ($\boldsymbol{x} \in \mathcal{B}_{\text{loc}}$). Without limiting the use of any other inelastic constitutive model, this domain is endowed with an *isotropic continuum damage model*, exhibiting inelasticity with regularized strain softening only for tensile stress - *tensile-damage continuum damage model* (Oliver, 1995b; Faria et al., 1998; Oliver et al., 2005). Its constitutive response is represented in a general form as:

See: ***Sec.2.2*** ***Paper #2***

$$\dot{\boldsymbol{\sigma}}_\mu = \Sigma^{\text{inelas}}(\dot{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\mu}) \qquad \forall \boldsymbol{x} \in \mathcal{B}_{\text{loc}}(t); \quad \forall \boldsymbol{y} \in \mathcal{B}_{\mu,\text{coh}} \tag{2.14}$$

where $\boldsymbol{\mu}$ stands for a set of *internal variables* accounting for the inelastic behavior evolution.

– The failure cell, $\mathcal{B}_\mu$ is associated to a smooth material point at the macroscale, ($\boldsymbol{x} \in \mathcal{B} \setminus \mathcal{B}_{\text{loc}}(t)$). In this case, the inelastic model, in Eq. (2.14) is enforced to behave instantaneously elastic at the cohesive bands domain, $\mathcal{B}_{\mu,\text{coh}}$, i. e.:

$$\dot{\boldsymbol{\sigma}}_\mu = \Sigma^{\text{elastic}}_{\text{inst}}(\dot{\boldsymbol{\varepsilon}}_\mu) = \mathbb{C}^{\text{elastic}}_{\text{inst}} : \dot{\boldsymbol{\varepsilon}}_\mu \quad \forall \boldsymbol{x} \in \mathcal{B} \setminus \mathcal{B}_{\text{loc}}(t); \quad \forall \boldsymbol{y} \in \mathcal{B}_{\mu,\text{coh}} \tag{2.15}$$

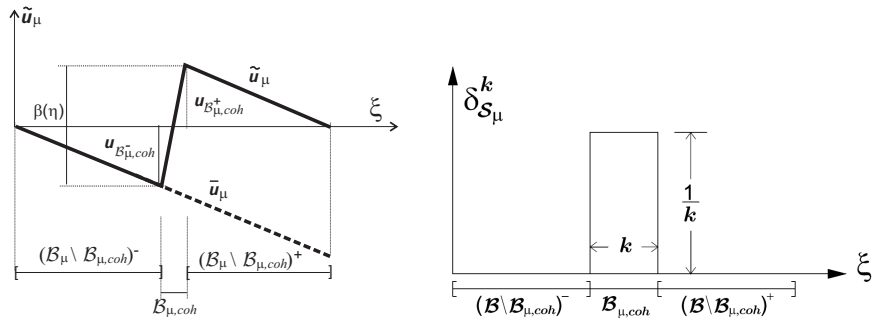
where, in continuum damage models, $\mathbb{C}^{\text{elastic}}_{\text{inst}} = (1 - d_\mu) \cdot \mathbb{C}_\mu$ with $\dot{d}_\mu = 0$, and being $\dot{d}_\mu$ the rate of the damage internal variable (a scalar for isotropic damage cases).

An advantage of this methodology, in the previous setting, is that the same failure cell morphology is considered to represent the microstructure at every macroscopic point of $\mathcal{B}$, both for $\boldsymbol{x} \in \mathcal{B}_{\text{loc}}(t)$ and $\boldsymbol{x} \in \mathcal{B} \setminus \mathcal{B}_{\text{loc}}(t)$. The only difference is the considered constitutive behavior at the cohesive bands $\mathcal{B}_{\mu,\text{coh}}$, defined in Eqs. (2.13),(2.14) and (2.15).

Displacement fluctuations in the CSDA: Considering Eq. (2.1), with a local coordinate system (ξ, η) aligned with the domain $\mathcal{B}_{\mu,\text{coh}}$ (see Fig. 4), and, exhibiting the de-cohesive behavior allocated to the cohesive bands, the smooth part of the microscopic displacement fluctuation field, $\bar{\boldsymbol{u}}_\mu$, can be expressed as:

$$\begin{aligned}
\dot{\bar{\boldsymbol{u}}}_\mu(\xi,\eta,t) &= \dot{\tilde{\boldsymbol{u}}}_\mu(\xi,\eta,t) - \mathcal{H}_{\mathcal{B}_{\mu,\text{coh}}}(\xi)\dot{\boldsymbol{\beta}}_\mu(\eta,t) && (a)\\
\mathcal{H}_{\mathcal{B}_{\mu,\text{coh}}}(\xi) &= \begin{cases} 0 & \forall \boldsymbol{y} \in (\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}})^- \\ \frac{\xi}{k} & \forall \boldsymbol{y} \in \mathcal{B}_{\mu,\text{coh}} \\ 1 & \forall \boldsymbol{y} \in (\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}})^+ \end{cases} && (b)\\
\dot{\boldsymbol{\beta}}_\mu(\eta(\boldsymbol{y}),t)\Big|_{\boldsymbol{y}\in\mathcal{B}_{\mu,\text{coh}}} &\equiv [\![\dot{\tilde{\boldsymbol{u}}}_\mu(\xi,\eta,t)]\!]^+_- && (c)
\end{aligned} \tag{2.16}$$

where $\mathcal{H}_{\mathcal{B}_{\mu,\text{coh}}}(\xi)$ is the k-regularized Heaviside function shifted to $\mathcal{B}_{\mu,\text{coh}}$, and $\dot{\boldsymbol{\beta}}_\mu(\eta,t)$ is a (smooth) function arbitrarily defined except for the restriction in Eq. (2.16)-(c), In Eq. (2.16) $[\![(\cdot)(\xi,\eta,t)]\!]^+_- \equiv (\cdot)(\xi,\eta,t)|_{\xi=k} - (\cdot)(\xi,\eta,t)|_{\xi=0}$, is the



(a) Microscopic displacement fluctuation field.

(b) k-regularized Dirac Delta function.

Figure 5: Cohesive Band behavior.

apparent jump of $(\cdot)(\xi,\eta,t)$ across the cohesive band.

Following these statements, the microscale displacement fluctuation is given by (see Fig. 5):

$$\dot{\tilde{\boldsymbol{u}}}_\mu(\xi,\eta,t) = \underbrace{\dot{\tilde{\boldsymbol{u}}}_\mu(\xi,\eta,t)}_{\text{smooth}} + \mathcal{H}_{\mathcal{B}_{\mu,\text{coh}}}(\xi)\dot{\boldsymbol{\beta}}_\mu(\eta,t) \tag{2.17}$$

Eq. (2.17) constitutes the displacement counterpart of a k-regularized strong discontinuity kinematics (Oliver, 1996a), and proves that the herein proposed cohesive-bands approach, is consistent with a k-regularized strong discontinuity at the cohesive domain $\mathcal{B}_{\mu,\text{coh}}$. In consequence, the corresponding microscopic strain fluctuation field is given by:

$$\boldsymbol{\nabla}^s_{\boldsymbol{y}}\dot{\tilde{\boldsymbol{u}}}_\mu(\boldsymbol{y},t) = \underbrace{\boldsymbol{\nabla}^s_{\boldsymbol{y}}\dot{\tilde{\boldsymbol{u}}}_\mu(\boldsymbol{y},t) + \mathcal{H}_{\mathcal{B}_{\mu,\text{coh}}}(\xi)\boldsymbol{\nabla}^s_{\boldsymbol{y}}\dot{\boldsymbol{\beta}}_\mu(\boldsymbol{y},t)}_{\dot{\tilde{\boldsymbol{\varepsilon}}}_\mu(\boldsymbol{y},t)} + \delta^k_{s_\mu}(\dot{\boldsymbol{\beta}}_\mu \otimes \boldsymbol{n}_\mu)^s \tag{2.18}$$

where $\delta^k_{s_\mu}$ stands for the k-regularized Dirac delta function, placed at the center line, S_μ, of $\mathcal{B}_{\mu,\text{coh}}$ (see Fig. 5-(b)). Thus, the rate of microscopic strain field $\dot{\boldsymbol{\varepsilon}}_\mu$ can be written in terms of the rate of macroscopic strain $\dot{\boldsymbol{\varepsilon}}$, and the rate of microscopic displacement fluctuations $\dot{\tilde{\boldsymbol{u}}}_\mu$, as follows:

$$\dot{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{x},\boldsymbol{y},t) = \underbrace{\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t) + \dot{\tilde{\boldsymbol{\varepsilon}}}_\mu(\boldsymbol{y},t)}_{\text{bounded}} + \underbrace{\delta^k_{s_\mu}(\dot{\boldsymbol{\beta}}_\mu \otimes \boldsymbol{n}_\mu)^s}_{\text{unbounded}} \tag{2.19}$$

From Eq. (2.19), it can be concluded, that the second term at the right-hand side becomes unbounded in the limit $k \to 0$. In multiscale modeling, this expression is equivalent to the one given, in phenomenological monoscale models, in the context of the *Continuum Strong Discontinuity Approach* (CSDA) of material failure (Oliver et al., 2002).

2.2.2 *Homogenized (induced) constitutive equation*

One of the most specific features of the proposed multiscale approach, is that the same homogenization setting is used in points of both domains, smooth ($\boldsymbol{x} \in \mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}}$), and non-smooth ($\boldsymbol{x} \in \mathcal{B}_{\mu,\text{coh}}$), coinciding with the approach presented in Sec. 2.1. Other approaches (Toro et al., 2014), redefine the failure cell along time, fulfilling conditions of material bifurcation induced by instabilities at the microscale. More complex approaches (Kouznetsova, 2002; Geers et al., 2010; Otero et al., 2015; Lesicar et al., 2015) propose the use of *second-order* computational homogenization schemes in order to get better accuracy in the prediction of high strain gradients. In this Thesis it is claimed the ability of the proposed approach to induce discrete failure in a *first-order homogenization* setting, giving rise to objective responses, and proper energy transfer through scales.

An issue appearing in this scenario, widely known in hierarchical multiscale approaches, is its high computational cost. In this context, the proposed model was also conceived to be combined with the use of *model order reduction techniques* (Paper #3) (Oliver et al., 2017b). These techniques have been deeply studied in this Thesis, and their main features are presented in Chapter 3.

In what follows, the consequences of the homogenization procedure based on the Hill-Mandel Principle of Macro-homogeneity are analyzed. The fact that the regularized strong discontinuities appear also at the microscale, being captured by the cohesive bands $\mathcal{B}_{\mu,\text{coh}}$, is one of the most relevant features of the proposed approach.

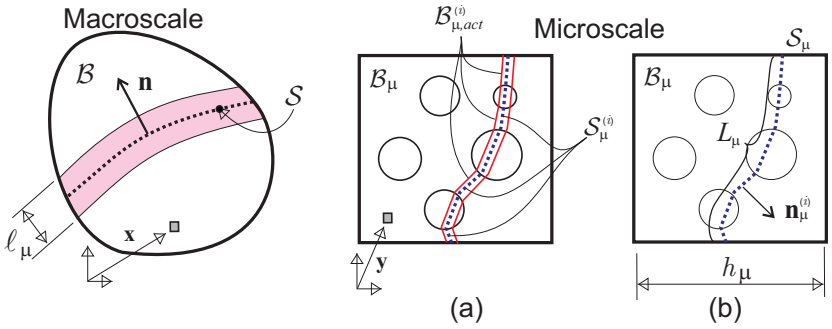


Figure 6: Multiscale model: **(a)** failure cell with activated failure mode; **(b)** geometrical characterization of the failure mode.

For the sake of generality, the RVE is considered composed by several components: a matrix, and randomly distributed inclusions and voids. In addition, a number of cohesive bands are considered defining the set $\mathcal{B}_{\mu,\text{coh}}$ (a sketch is presented in Fig. 6); those cohesive bands allow failure within the matrix, across the aggregates and at the matrix/aggregate interface.

See: ***Sec.2.4*** ***Paper #2***

Following the previous domain decomposition (smooth and non-smooth subdomains) in Sec. 2.2.1.2, the Eq. (2.9) can be integrated in the two separated subdomains:

$$\dot{\boldsymbol{\sigma}} = \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \dot{\boldsymbol{\sigma}}_\mu \, dV = \frac{1}{\Omega_\mu}\left[\int_{\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}}} \dot{\boldsymbol{\sigma}}_\mu \, dV + \int_{\mathcal{B}_{\mu,\text{coh}}} \dot{\boldsymbol{\sigma}}_\mu \, dV\right] \tag{2.20}$$

In consonance with the definition of $\dot{\boldsymbol{\sigma}}_\mu$ (in particular, the bounded behavior of the microscopic stress field), the second term on the right hand side can be neglected assuming a small enough width of the cohesive bands ($k \to 0$).

Finally, after some manipulations of Eq. (2.20), and following the definitions of microscale kinematics in Eq. (2.19), and the lemma in Eq. 23 in Paper #2 (Oliver et al., 2015), the resulting homogenized constitutive equation fulfills the following:

$$\dot{\boldsymbol{\sigma}}(\boldsymbol{x},t) = \bar{\mathbb{C}} : \left[\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t) + \dot{\boldsymbol{\chi}}(\boldsymbol{x},t) - \dot{\boldsymbol{\varepsilon}}^{(i)}(\boldsymbol{x},t)\right] \qquad \bar{\mathbb{C}} \equiv \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \mathbb{C}_\mu(\boldsymbol{y}) \, dV \tag{2.21}$$

$$\dot{\boldsymbol{\varepsilon}}^{(i)}(\boldsymbol{x},t) = \frac{1}{\Omega_\mu}\int_{S_\mu} (\dot{\boldsymbol{\beta}}_\mu \otimes \boldsymbol{n}_\mu)^s \, dS = \frac{1}{l_\mu}\overline{(\dot{\boldsymbol{\beta}}_\mu \otimes \boldsymbol{n}_\mu)^s}_{S_\mu} \tag{2.22}$$

$$\overline{(\dot{\boldsymbol{\beta}}_\mu \otimes \boldsymbol{n}_\mu)^s}_{S_\mu} = \frac{1}{L_\mu}\int_{S_\mu} (\dot{\boldsymbol{\beta}}_\mu \otimes \boldsymbol{n}_\mu)^s \, dS \qquad l_\mu(\boldsymbol{x},t) \equiv \frac{\Omega_\mu}{L_\mu} = \mathcal{O}(h_\mu) \tag{2.23}$$

where, l_μ stands for a *characteristic length*, depending on the activated microscopic failure pattern. The tensorial entities $\dot{\boldsymbol{\chi}}(\boldsymbol{x},t)$ and $\dot{\boldsymbol{\varepsilon}}^{(i)}(\boldsymbol{x},t)$, are inelastic strains, and play the same role than internal variables in phenomenological models. However, unlike them, here, their evolution is determined, at every macroscopic sampling point $\boldsymbol{x}$, by homogenized values of entities at the corresponding microscopic failure cell $\mathcal{B}_\mu$. This extends to non-smooth problems, some theoretical results already derived for smooth problems, see (Michel and Suquet, 2003, 2004). In addition, a characteristic length l_μ emerges naturally in Eq. (2.22), as the ratio between the measure of the failure cell (area in 2D and volume in 3D), and the measure (length/surface) of the activated microscopic failure mechanism. In consequence this length is of the order of the failure cell size. For a deeper review of the analytical results of this induced homogenized constitutive model, the reader is addressed to Sec. 2.4 in Paper #2.

See: ***Sec.2.4 Paper #2***

The role of the characteristic length, l_μ, *naturally derived* from the present formulation, is not only computational, but it has also other very relevant physical and mechanical implications. Consideration of such a characteristic length, for multiscale based approaches, has been claimed from the material mechanics community (Bazant, 2010), and sometimes introduced in a heuristic way in other approaches (Unger, 2013). This characteristic length depends on both the specific data of the problem and the local microscopic failure state. Through its consideration, the correct energy transfer between scales and mesh

size objectivity can be achieved.

In summary, Eq. (2.21) and Eq. (2.22) retrieve the format of a *constitutive model equipped with an internal length and with internal variables whose evolution is described by the microstructure behavior*. Although this model will never be used for computational purposes[2], it supplies relevant insights on the properties of the resulting homogenized constitutive model.

2.2.3 *Energy dissipation*

Let us consider, on one hand the fracture energy, $G_\mu^f(\boldsymbol{y})$ corresponding to points $\boldsymbol{y} \in \mathcal{B}_\mu$, defined as a material property specific for every compound of the heterogeneous RVE, and, on the other hand, the macroscale fracture energy $G^f(\boldsymbol{x})$, obtained as an output from the homogenization procedure. According to their definitions, those fracture energies can be computed in terms of fracture energy densities, in terms of the energy dissipation that takes place in bands with bandwidth k (at the microscale) and l_μ (at the macroscale), respectively.

$$g^f(\boldsymbol{x}) = \frac{G^f(\boldsymbol{x})}{l_\mu(\boldsymbol{x})} = \int_0^\infty \boldsymbol{\sigma}(\boldsymbol{x},t) : \dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t)\, dt \quad \forall \boldsymbol{x} \in \mathcal{B}_{\text{loc}} \tag{2.24}$$

$$g_\mu^f(\boldsymbol{x}) = \frac{G_\mu^f(\boldsymbol{x})}{k} = \int_0^\infty \boldsymbol{\sigma}_\mu(\boldsymbol{y},t) : \dot{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},t)\, dt \quad \forall \boldsymbol{y} \in \mathcal{B}_{\mu,\text{coh}} \tag{2.25}$$

See: ***Sec.2.6*** ***Paper #2***

In virtue of the Hill-Mandel Principle of Macro-Homogeneity, it can be concluded that the macroscopic fracture energy G^f is equivalent to the average of microscopic fracture energy $G_\mu^f(\boldsymbol{y})$, along the activated failure mechanism at the microscale S_μ. Replacing Eq. (2.24) into Eq. (2.8), and after some manipulations, the macroscopic fracture energy is given by the expression (Oliver et al., 2015):

$$\frac{G^f(\boldsymbol{x})}{l_\mu(\boldsymbol{x})} = \underbrace{\frac{L_\mu}{\Omega_\mu}}_{\frac{1}{l_\mu(\boldsymbol{x})}} \cdot \frac{1}{L_\mu} \int_{S_\mu} G_\mu^f(\boldsymbol{y})\, dS_\mu = \frac{1}{l_\mu(\boldsymbol{x})} \overline{G_\mu^f(\boldsymbol{y})}_{S_\mu} \tag{2.26}$$

$$G^f(\boldsymbol{x}) = \overline{G_\mu^f(\boldsymbol{y})}_{S_\mu} \tag{2.27}$$

where $\overline{G_\mu^f(\boldsymbol{y})}_{S_\mu}$ is the mean value of the microscopic fracture energy varying along the active failure path. Eq. (2.28) provides the relationship of fracture energies at both scales. In case of an homogeneous fracture energy at the active cohesive bands of the microscale, Eq. (2.26) translates into an exact equivalence of fracture energies along the scales, i. e.:

$$G^f(\boldsymbol{x}) = G_\mu^f \tag{2.28}$$

2 Instead, the homogenized value of the stress in Eq. (2.9) is point-wise used to evaluate the current macroscopic stress in terms of the corresponding macroscopic strain.

In the light of this result, it can be easily concluded that the fracture energies at the microscale determine, in average, the effective fracture energy at the macroscale. It is stressed the importance of the characteristic length l_μ in order to guarantee the proper dissipation between scales. For more details, the reader is addressed to Appendix B in Paper #2.

2.2.4 *Numerical aspects: finite element model*

The proposed multiscale formulation has been implemented in a Finite Element model following the setting of a FE^2 strategy. Accordingly, two *nested* finite element models are used:

- At the macroscale level, an Embedded Finite Element Methodology (EFEM) based on the CSDA is used, aiming at capturing propagating cracks. As described in (Oliver et al., 2014), this technology consists of the insertion, during specific stages of the simulation, of goal oriented specific strain fields via mixed finite element formulations. This allows modeling propagating cracks through the macroscale finite element mesh.
- A standard Finite Element model is used at the microscale level, combining standard elements endowed with continuum hardening constitutive models and cohesive-band elements endowed with regularized constitutive softening models. These are placed in the edges of every finite element, capturing the crack onset and strain localization, similar to the cohesive interface elements in (Pandolfi et al., 1999), and more recently in (Rodrigues et al., 2016). This approach benefits the simplicity of the algorithm and the non-intrusive character of its implementation.

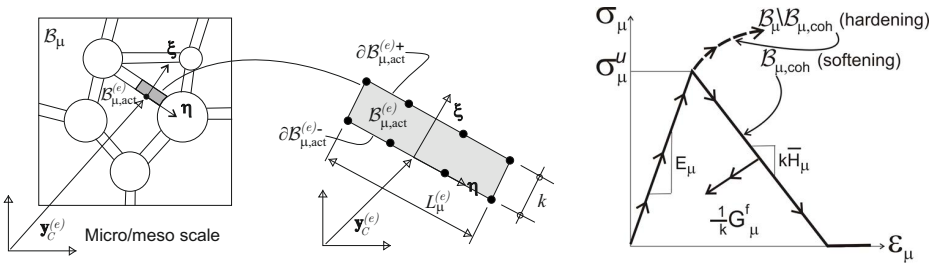
In what follows, these two finite element models are described.

2.2.4.1 *Failure cell finite element model*

Standard quadrilateral finite elements are adopted for the numerical simulation of the cell response. The cohesive bands $\mathcal{B}_{\mu,\text{coh}}$ are also modeled by quadrilateral isoparametric finite elements of very small thickness $k \ll h_\mu$ (high aspect ratio), as shown in Fig. 7-(a), endowed with constitutive models whose behavior is sketched in Fig. 7-(b) and defined through equations (2.13) to (2.15). The remaining finite elements of the cell are endowed with either elastic or inelastic hardening responses. Therefore, only elements on the cohesive bands can exhibit strain localization.

The corresponding nonlinear problem in the failure cell is then solved for the discretized version of the microscale displacement fluctuations, using Eq. (2.10). Dirichlet boundary conditions precluding rigid body motions, and minimal boundary conditions in Eq. (2.4), are also imposed.

Material failure propagates naturally through the RVE, strain localization takes place at the finite elements defining the cohesive bands. At every time step of the



(a) Cohesive bands (b) Constitutive equation

Figure 7: Multiscale model: finite element discretization at the microscale.

analysis, those finite elements who are in loading state, define the active set of cohesive bands $\mathcal{B}_{\mu,\text{act}}$ conforming the active failure mechanism.

2.2.4.2 *Finite element model at the macroscale: material failure propagation*

One of the most critical issues in computational modeling of material failure is the appropriate capture of the crack onset and propagation. *When does failure trigger at a given material point?* and *how does it propagate?*, these two questions are the cornerstone of material failure propagation algorithms.

At the microscale, where the morphology and the position of candidate propagation mechanisms are predefined, the two issues are of minor relevance due to the adopted simplified failure-bands model. However, at the macroscale, there is not a predefined failure path, and in principle, any material point may fail and propagate in any direction. To adequately solve the previous questions, the procedure for modeling onset and propagation of discontinuities recently developed for monoscale problems (Oliver et al., 2014) has been extended to the multiscale setting. The proposed methodology is based on the use of the following specific techniques:

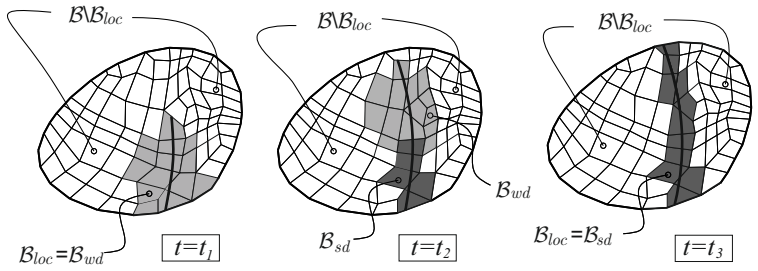


Figure 8: Evolution of the injection domains for three typical stages ($t_1 < t_2 < t_3$) of the discontinuity propagation.

- ***Strain injection techniques***: based on the use of goal oriented assumed-strain fields injected in selected domains, via mixed formulations (Simo and

Hughes, 1986; Simo and Riffai, 1990; Reddy and Simo, 1995; Zienkiewicz and Taylor, 2000). The standard (four points) Gauss quadrature rule, corresponding to full integration of two-dimensional quadrilaterals, is complemented with two additional sampling points placed at the barycenter of the element (see Fig. 9), termed *singular* and *regular* sampling points. These two additional quadrature points sample the stresses similarly to the standard Gauss points. Therefore, for the injected elements, numerical integration (typically evaluation of the incremental internal forces in terms of the stresses), is based on those two additional sampling points by defining the weight indicated in **Table 3** - in Appendix B in Paper #2.

As for propagation purposes, two different enhanced strain injection stages, are considered[3]:

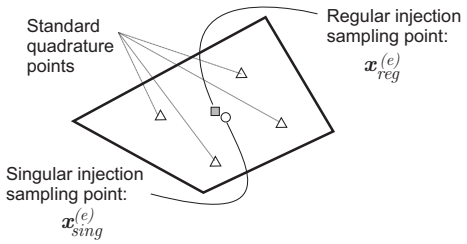


Figure 9: Sampling points involved in the numerical integration.

– In a *first stage, the* ***weak discontinuity stage***, embedded localization bands with bandwidth l_μ, at the macroscale, are incrementally injected (prior to development of displacement discontinuities) in an evolving subdomain. These embedded localization bands have no preferred orientation (they have an isotropic character), and exhibit a great ability to propagate material failure in the proper directions. This so-injected elements are used for a very short time in order to avoid stress locking effects. Once the crack propagation remains stable, and the crack path is well defined, the injection stage is switched to the second stage.

  The injected strain rate at element e, with n_{node} nodes, is the following:

$$\dot{\boldsymbol{\varepsilon}}^{(e)}(\boldsymbol{x},t) \equiv \underbrace{\sum_{i=1}^{n_{\text{node}}} \boldsymbol{\nabla} N_i(\boldsymbol{x}) \otimes \dot{\boldsymbol{u}}_i(t)}_{\text{regular}} + \underbrace{\zeta_{\mathcal{S}}^{h^{(e)},l_\mu^{(e)}}(\boldsymbol{x})\dot{\boldsymbol{\gamma}}^{(e)}(t)}_{\text{singular}} \tag{2.29}$$

  where N_i are the standard shape functions, $\dot{\boldsymbol{u}}(t)$, the macroscale nodal displacements, and $\zeta_{\mathcal{S}}^{h^{(e)},l_\mu^{(e)}}$ is the regularized *dipole-function* in the element (e). A description of the weak enhanced mode is presented in

3 To switch between stages, a set of *control variables* are defined, all those detailed in Sec. 3 and Appendix B in Paper #2

Fig. 10, In addition, the variational problem for the weak-discontinuity regime, in rate form, is presented in **Box A1** in the Appendix B in Paper #2.

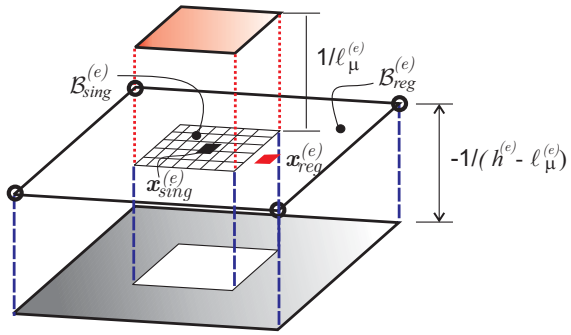


Figure 10: Weak discontinuity mode. Elemental regularized dipole function $\zeta_{\mathcal{S}}^{h^{(e)},l_\mu^{(e)}}$.

– In a *second stage, the* ***strong discontinuity stage***, the obtained crack path field, $\mathcal{S}$, is used to determine the appropriate placement of an elemental embedded strong discontinuity strain field, which is incrementally injected in the appropriated set of elements $\mathcal{B}_{\text{sd}}$.

  In the present multiscale context, the proposed second stage consists of the incremental injection of the following *elemental strong discontinuity mode*:

$$\dot{\boldsymbol{\varepsilon}}^{(e)}(\boldsymbol{x},t) \equiv \underbrace{\sum_{i=1}^{\text{node}} \boldsymbol{\nabla} N_i(\boldsymbol{x}) \otimes \dot{\boldsymbol{u}}_i(t)}_{\text{regular}} + \underbrace{\delta_s^{l_\mu^{(e)}} (\dot{\boldsymbol{\beta}}^{(e)} \otimes \boldsymbol{n}^{(e)})^S}_{\text{singular}} \tag{2.30}$$

  in terms of the ${l_\mu}^{(e)}$-regularized Dirac delta function $\delta_s^{l_\mu^{(e)}}$ (displayed in Fig. 11), being $\boldsymbol{n}^{(e)}$ the direction of the element normal provided by the solution of the discontinuous bifurcation problem presented in Sec. 2.5 in Paper #2. The resulting variational problem for the injection procedure is summarized in **Box A2** - Appendix B in Paper #2.

- ***Crack-path-field techniques***: their goal is the identification of the trace of the propagating crack by means of the so-called crack-path field. It is denoted as $\mu(\boldsymbol{x},t)$, and obtained from a selected localized scalar variable $\alpha(\boldsymbol{x},t)$, which identifies the crack path as the locus where $\alpha(\boldsymbol{x},t)$ takes its transversal maximum value. In order to define this locus $\mathcal{S}_t$, some alternatives have been developed in this multiscale framework, see Eqs. $55-56$ in Paper #2 (Oliver et al., 2015). The variational statement for the crack-path field model is fully detailed in **Box 3.1** in Paper #2.

The resulting procedure is a robust and efficient technique to model propagating material failure in a finite element setting. It is especially appropriate for

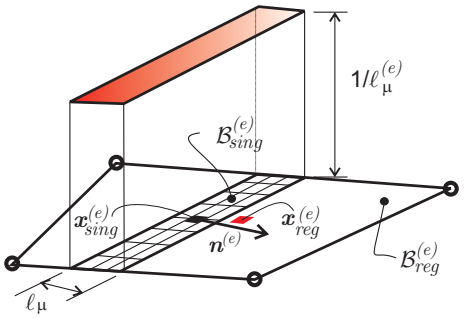


Figure 11: Strong discontinuity mode. Elemental regularized Dirac delta function $\delta_s^{l_\mu^{(e)}}$.

capturing material failure propagation in coarse meshes, in contraposition of the alternative extra elemental character techniques (i. e. *phase-field*, *gradient* or *non-local damage* models), where several elements span the localization band. In addition, its implementation in an existing finite element code has a little intrusive character.

In regards to the space and time integrations, as commented above, injection of weak-discontinuity and strong-discontinuity modes requires, in principle, specific integration rules in space, i. e.: a standard four-point Gauss quadrature rule, and two additional sampling points, for injected elements, $\mathcal{B}^{(e)} \subset \mathcal{B}_{\text{wd}}$ and $\mathcal{B}^{(e)} \subset \mathcal{B}_{\text{sd}}$ so that $\mathcal{B}_{\text{loc}} = \mathcal{B}_{\text{wd}} \cup \mathcal{B}_{\text{sd}}$. Since those domains evolve along time (see Fig. 8), some additional problems on the time-integration of the resulting equilibrium equations are found. To tackle this issue, in (Oliver et al., 2014) and (Oliver et al., 2015) is proven that defining some "equivalent" stress entities at the standard Gauss points, the spatial integration can be rephrased as a standard four Gauss points integration rule in the integration domain. This space-time integration rule is fully explained in Appendix B3 in Paper #2, and the corresponding stress evaluation is also summarized in Box A3.

3

MODEL ORDER REDUCTION IN MULTISCALE ANALYSIS

3.1 General Framework

Along this Thesis, some techniques for reducing the unaffordable computational cost inherent to the numerical simulation of multiscale fracture problems have been developed. Those techniques are combined to obtain a hyper-reduced order model HPROM, based on a two-stage strategy:

- First Stage: also termed Reduced Order Model (ROM), consists of a Galerkin projection, via *Proper Orthogonal Decomposition* (POD), onto a small space (*reduced-order manifold*), in which the set of modes conforming the projection basis are computed off-line.
- Second Stage: also called *HyPer-Reduced Order Model* (HPROM). Two different techniques have been developed here. The first one is based on *interpolation methods*, widely applied in problems exhibiting hardening behavior (see Paper #1) (Hernández et al., 2014b). The second one, based on a Reduced Optimal Quadrature (ROQ) rule, has been applied to fracture (non-smooth) problems. Similarly to the previous stage, the reduced modes functions used to reconstruct the state variables, are computed off-line.

In what follows, these techniques have been applied to the microscale Boundary Value Problem (BVP), while the finite element model at the macroscale remains as the standard one.

3.2 Reduced-order modeling (ROM) of the RVE problem

The model order reduction concept relies on the premise that, for any input parameter $\boldsymbol{\mu} \in \mathcal{D}$ governing the microscale displacement fluctuations $\tilde{\boldsymbol{u}}_{\mu}$, the solution can be approximated by a set of n linearly independent basis functions

$\boldsymbol{\Phi}$ approximately spanning the *primal variable*[1] space.

Following this idea, the *off-line stage* is devoted to determine via a POD technique, the reduced basis in which the High Fidelity (HF) solution is projected. Once this basis has been obtained, a subsequent *online stage* in the *reduced-space* is considered.

3.2.1 *Computation of the reduced basis functions*

Taking as a primal variable the displacement fluctuations, and departing from the problem depicted in Sec. 2.1, a first step consists of determining an approximation[2] of the finite element space of kinematically admissible microscale displacement fluctuations $\tilde{\mathcal{V}}_\mu^h$. This approximation is obtained as the span of the displacement fluctuation solutions obtained, for a judiciously chosen set of n_{hst} input strain *trajectories*, every trajectory being discretized into a number of steps n_{stp}. These set of finite element solutions are stored into the snapshot matrix $\mathbf{X}_u$ as column vectors:

$$\mathbf{X}_u = [\mathbf{U}^1, \mathbf{U}^2, \mathbf{U}^3, \cdots, \mathbf{U}^{n_{snp}}] \tag{3.1}$$

In consequence, the approximating space for $\tilde{\mathcal{V}}_\mu^h$, henceforth called the *snapshot space*, is then defined as:

$$\tilde{\mathcal{V}}_u^{snp} = \text{span}\{\tilde{\boldsymbol{u}}_\mu^1(\boldsymbol{y}), \tilde{\boldsymbol{u}}_\mu^2(\boldsymbol{y}), \tilde{\boldsymbol{u}}_\mu^3(\boldsymbol{y}), \cdots, \tilde{\boldsymbol{u}}_\mu^{n_{snp}}(\boldsymbol{y})\} \subseteq \tilde{\mathcal{V}}_\mu^h \tag{3.2}$$

where, $n_{snp} = n_{stp} \cdot n_{hst}$ is the total number of snapshots.

Once the snapshot matrix $\mathbf{X}_u$ has been computed, the *Elastic-Inelastic decomposition* technique is used to determine the *reduced basis functions*. The reason for it relies on the fact that the Singular Value Decomposition (SVD) applied to the whole matrix $\mathbf{X}_u$, may produce basis with a large number of elements, which makes difficult to retrieve the response of the RVE in some specific cases. Particularly, the elastic response[3], might request a much larger number of basis functions, this translating into a significant waste of computational cost.

To eliminate this shortcoming, in this Thesis, it is proposed a time partition of the space of snapshots $\tilde{\mathcal{V}}_u^{snp}$ into *elastic* ($\tilde{\mathcal{V}}_{u,el}^{snp}$), and *inelastic* ($\tilde{\mathcal{V}}_{u,inel}^{snp}$) subspaces.

$$\tilde{\mathcal{V}}_u^{snp} = \tilde{\mathcal{V}}_{u,el}^{snp} \oplus \tilde{\mathcal{V}}_{u,inel}^{snp} \tag{3.3}$$

See: ***App. B*** ***Paper #1***

obtaining the reduced basis as the combination (spatial sum) of both sub-bases. An orthonormal basis for $\tilde{\mathcal{V}}_{u,el}^{snp}$ is determined by taking a low number of elastic snapshots (at a minimum, $n_{snp}^e = 3$ for 2D problems, $n_{snp}^e = 6$ for 3D problems),

1 *Primal variable* is known as the selected variable to perform the reduction process.

2 In general, $\tilde{\mathcal{V}}_\mu^h$ cannot be precisely determined, such a task will require finite element analyses of the cell under *all conceivable* strain paths. Rather, one has to be content to construct an *approximation* of it.

3 Under an infinitesimal strain framework, this response is exactly recovered with only three basis (Hernández et al., 2014b).

and computing the corresponding orthonormal basis.

Once this set of elastic basis is known, the orthogonal projection of each snapshot $\tilde{\boldsymbol{u}}^k$ onto the orthogonal complement of $\tilde{\mathcal{V}}_{u,el}^{snp}$ is computed; with this new set of snapshots, the inelastic basis functions are obtained via SVD. Finally, the assembled basis results the following:

$$[\boldsymbol{\Phi}] = [\ \underbrace{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \boldsymbol{\Phi}_3}_{\text{elastic modes basis}} \qquad \underbrace{\boldsymbol{\Phi}_4, \boldsymbol{\Phi}_5, \boldsymbol{\Phi}_6, \cdots, \boldsymbol{\Phi}_{n_u}}_{\text{"essential" inelastic modes basis}}\ ] \tag{3.4}$$

and the reduced-order space $\tilde{\mathcal{V}}_u^*$, spanned by this base, is:

$$\tilde{\mathcal{V}}_u^* = \tilde{\mathcal{V}}_{u,el}^{snp} \oplus \tilde{\mathcal{V}}_{u,inel}^{snp,*} = \text{span}\left\{ \underbrace{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \boldsymbol{\Phi}_3}_{\text{elastic modes}} \quad \underbrace{\boldsymbol{\Phi}_4, \boldsymbol{\Phi}_5, \boldsymbol{\Phi}_6, \cdots, \boldsymbol{\Phi}_{n_u}}_{\text{"essential" inelastic modes}} \right\} \subseteq \tilde{\mathcal{V}}_\mu^h \tag{3.5}$$

Placing the m_e elastic modes in the first m_e positions, followed by the *essential*[4] inelastic modes, ensures the reduced-order model to deliver linear elastic solutions with the same accuracy than the HF solutions. For more details, the reader is encouraged to sent to the Appendix B in (Hernández et al., 2014b).

3.2.1.1 *Formulation of the reduced order model*

Once the reduced basis $[\boldsymbol{\Phi}]$ is computed, the online stage consists of solving the discrete version of the microscale equilibrium equation (via FE), projected onto the reduced-order space $\tilde{\mathcal{V}}_u^* \subseteq \tilde{\mathcal{V}}_\mu^h$ spanned by $[\boldsymbol{\Phi}]$. To this end, the test and trial functions, $\boldsymbol{\eta}$ and $\tilde{\boldsymbol{u}}_\mu$, are approximated by the following linear expansions:

*See: **Sec. 4 Paper #1***

$$\tilde{\boldsymbol{u}}_\mu(\boldsymbol{y}) \approx \tilde{\boldsymbol{u}}_\mu^*(\boldsymbol{y}) = \sum_{i=1}^{n_u} \boldsymbol{\Phi}_i(\boldsymbol{y}) \boldsymbol{c}_i \tag{3.6}$$

$$\boldsymbol{\eta}(\boldsymbol{y}) \approx \boldsymbol{\eta}^*(\boldsymbol{y}) = \sum_{i=1}^{n_u} \boldsymbol{\Phi}_i(\boldsymbol{y}) \boldsymbol{c}_i^* \tag{3.7}$$

where, $\tilde{\boldsymbol{u}}_\mu^*$ and $\boldsymbol{\eta}^*$ stand for the *low-dimensional* approximations of trial and test functions, respectively.

Introducing expressions (3.6) and (3.7) into the discrete version of the microscale BVP (see Sec. 4 in (Hernández et al., 2014b)), and multiplying the resulting expression by $\boldsymbol{\Phi}^T$ (Galerkin projection), it yields:

4 Essential based on a threshold given by an *a-priori* error estimation, see Sec. 9.4 in (Hernández et al., 2014b), thus, $\tilde{\mathcal{V}}_{u,inel}^{snp,*}$ corresponds to the truncated version of the full base with $n_u - 3$ dominant modes.

PROBLEM A (ROM) (*Microscale reduced problem via POD*):
Given the macroscale strain, $\boldsymbol{\varepsilon}$, and the reduced basis for displacement fluctuations $\boldsymbol{\Phi}$, find $\boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}$ satisfying:

$$\int_{\mathcal{B}_\mu} \mathbf{B}^{*T}(\boldsymbol{y})\boldsymbol{\sigma}_\mu(\boldsymbol{y}, \boldsymbol{\varepsilon} + \mathbf{B}^*\boldsymbol{c}, \boldsymbol{\mu})\, d\mathcal{B}_\mu \approx \sum_{i=1}^{n_g} \mathbf{B}^{*T}(\boldsymbol{y}_g, :)\boldsymbol{\sigma}_\mu(\boldsymbol{y}_g, :)\, w_g = \mathbf{0} \qquad (3.8)$$

where $\boldsymbol{c} = [c_1, c_2, \cdots, c_{n_u}] \in \mathbb{R}^{n_u}$ denotes the vector containing the coefficients associated to each basis function $\boldsymbol{\Phi}_i$, being $\boldsymbol{c}$ the basic unknowns for the standard reduced-order problem. $\mathbf{B}^*$ stands for the *reduced* strain-displacement matrix "B-matrix" defined as $\mathbf{B}^*(\boldsymbol{y}) = \mathbf{B}(\boldsymbol{y}) \cdot \boldsymbol{\Phi}(\boldsymbol{y})$. When using a Gauss quadrature integration scheme, $n_g = \mathcal{O}(n)$ is the total number of Gauss points of the mesh; w_g denotes the weight associated to the g-th Gauss point $\boldsymbol{y}_g$; $\mathbf{B}(\boldsymbol{y}_g, :)$ and $\boldsymbol{\sigma}_\mu(\boldsymbol{y}_g, :)$ stand for the reduced B-matrix and the stress vector at Gauss point $\boldsymbol{y}_g$, respectively (Hernández et al., 2014a).

3.2.2 *Specific issues in non-smooth (fracture) problems*

3.2.2.1 *Domain separation strategy*

Taking advantage of the unbounded character of the microscale strain field typically observed in this kind of problems, the failure cell is splitted into a *regular domain* (made of elastic matrix and possible inclusions) and a *cohesive domain* (cohesive bands exhibiting a softening cohesive behavior). Details on this issue can be found in Sec. 3.2.2 in Paper #3 (Oliver et al., 2017b).

3.2.3 *Formulation of the microscale saddle-point problem*

In addition to this proposal, the ROM of the failure cell is formulated in an unconventional manner, i. e.: *in terms of strain fluctuations rather than in terms of conventional displacement fluctuations*.

As it will be shown later, it is convenient to rephrase the original problem, posed in terms of displacement fluctuations (*PROBLEM-I* in Paper #3). The primary unknowns of the rephrased problem are now the microscale strain fluctuations instead of its displacement fluctuations, while the constrained original minimum problem (of the standard micro-cell BVP) is rewritten in terms of a Lagrange functional. The resulting formulation is a variationally consistent *saddle-point* formulation.

Considering the material free energy φ_μ for the isotropic damage model in $\mathcal{B}_{\mu,\text{coh}}$, the microscale stress field $\boldsymbol{\sigma}_\mu$ can be expressed as:

$$\boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}_\mu, \boldsymbol{\mu}) = \frac{\partial \varphi_\mu(\boldsymbol{\varepsilon}_\mu, \boldsymbol{\mu})}{\partial \boldsymbol{\varepsilon}_\mu} = \frac{\partial \varphi_\mu(\boldsymbol{\varepsilon} + \tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\mu})}{\partial \boldsymbol{\varepsilon}_\mu} \qquad (3.9)$$

complemented by the evolution laws of the internal variables (Simo and Hughes, 1998). Thus, in consonance with the hierarchical multiscale approach, the following parametrized functional can be defined:

$$\Pi_{[\varepsilon,\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda}) = \int_{\mathcal{B}_\mu} \varphi_\mu(\tilde{\boldsymbol{\varepsilon}}_\mu)\, d\mathcal{B}_\mu + \boldsymbol{\lambda} : \int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu\, d\mathcal{B}_\mu \tag{3.10}$$

where, $\boldsymbol{\lambda}(t) \in \mathbb{S}^{n\times n}$, is a symmetric second order tensor Lagrange multiplier enforcing condition (2.4) on the microscale strain fluctuations $\tilde{\boldsymbol{\varepsilon}}_\mu$. With this parametrized functional $\Pi_{[\varepsilon,\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda})$, a *saddle-point* problem can be stated as:

PROBLEM II (HF) (*Microscale saddle-point problem)*: Given the macroscale strain, $\boldsymbol{\varepsilon}$, find $\tilde{\boldsymbol{\varepsilon}}_\mu$ and $\boldsymbol{\lambda}$ satisfying:

$$\{\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{\varepsilon}, \boldsymbol{\mu}), \boldsymbol{\lambda}(\boldsymbol{\varepsilon}, \boldsymbol{\mu})\} = \arg\Big\{\min_{\tilde{\boldsymbol{\varepsilon}}_\mu \in \mathcal{E}_\mu} \max_{\boldsymbol{\lambda}\in\mathbb{S}^{n\times n}} \Pi_{[\varepsilon,\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda})\Big\} \tag{3.11}$$

Such that:

$$\dot{\boldsymbol{\mu}} = f(\boldsymbol{\varepsilon}_\mu, \boldsymbol{\mu})$$

where $\mathcal{E}_\mu$ stands for the space of microscale kinematically compatible strain fluctuations and, f stands for the evolution equation of the internal variables. After considering that the microscale stress field $\boldsymbol{\sigma}_\mu$ is given by Eq. (3.9), the following optimality conditions emerge:

$$\int_{\mathcal{B}_\mu} [\boldsymbol{\sigma}_\mu(\tilde{\boldsymbol{\varepsilon}}_\mu)(\boldsymbol{y}) + \boldsymbol{\lambda}] : \hat{\boldsymbol{\varepsilon}}_\mu\, d\mathcal{B}_\mu = 0; \qquad \forall \hat{\boldsymbol{\varepsilon}}_\mu \in \mathcal{E}_\mu \tag{3.12}$$

$$\hat{\boldsymbol{\lambda}} : \int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu\, d\mathcal{B}_\mu = \mathbf{0}; \qquad \forall \hat{\boldsymbol{\lambda}} \in \mathbb{S}^{n\times n} \tag{3.13}$$

Eqs. (3.12) and (3.13) provide the solution of the saddle-point problem stated in Eq. (3.11). It can be proven that Eqs. (3.12) and (3.13) make *PROBLEM II* equivalent to the original problem in Eq. (2.10), but now rephrased in terms of the microscale strain fluctuations $\tilde{\boldsymbol{\varepsilon}}_\mu$ (see *PROBLEM I-R* in Paper #3).

3.2.3.1 *Computation of the reduced basis functions*

The transition from the high-dimensional finite element space to the reduced-order space, is accomplished by applying the POD technique, now for non-smooth problems. The standard reduced order model is based on the reduction of the strain fluctuation field $\tilde{\boldsymbol{\varepsilon}}_\mu$. The first step consists of generating a collection of solutions (samples) from different trial loading cases, representatives of all possible loading cases.

In each trial case, the microscale strain fluctuation $\tilde{\boldsymbol{\varepsilon}}_\mu$ at every Gauss point, is collected and stored in the snapshot matrix as a column vector:

$$\mathbf{X}_k = [\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_1), \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_2), \cdots, \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_{N_g})]^T$$

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \cdots, \mathbf{X}_{n_{snp}}] \in \mathbb{R}^{(N_g \cdot n_\sigma) \times n_{snp}} \tag{3.14}$$

See: Sec.4 Paper #3

where n_{snp} is the number of snapshots vectors. Therefore, $\mathbf{X}$ represents a number of sampled solutions obtained with the HF model under different loading conditions. For more details, the reader is addressed to Sec. 4 in (Oliver et al., 2017a).

In order to get a more accurate estimation of the dominant modes of the microscale strain fluctuations, it is convenient to separate the microscale into specific *sub-blocks* in accordance with the type of material response observed during the load history. The procedure is sketched in Fig. 12:

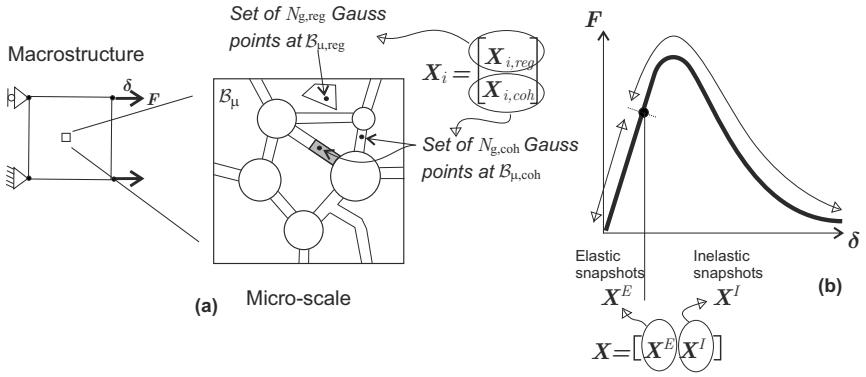


Figure 12: (a) All entries of $\mathbf{X}$ are partitions into two sub-blocks: $\mathbf{X}_{reg}$ and $\mathbf{X}_{coh}$, the strain fluctuations from points at *regular* and *cohesive* domains, respectively. (b) Snapshots taken from the elastic regime of the failure cell correspond to the sub-block $\mathbf{X}^E$. Snapshots taken during the inelastic regime (i. e. at least one Gauss point at $\mathcal{B}_{\mu,\text{coh}}$ is in inelastic state) correspond to the sub-block $\mathbf{X}^I$.

In addition, the *Elastic-Inelastic snapshot decomposition* above explained (Hernández et al., 2014b), is also applied to each sub-block. Thus, without loss of generality, the snapshot entries are organized so that the first, $N_{g,reg}$ entries correspond to Gauss points in $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}}$, while the remaining $N_{g,coh}$ entries correspond to Gauss points in $\mathcal{B}_{\mu,\text{coh}}$.

See: Sec.3.2.2 Paper #3

In accordance with this criterion, the snapshot matrix (3.14) can now be partitioned into sub-matrices as follows:

$$[\mathbf{X}] = \begin{bmatrix} \mathbf{X}_{\text{reg}} \\ \mathbf{X}_{\text{coh}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^E_{\text{reg}} & \mathbf{X}^I_{\text{reg}} \\ \mathbf{X}^E_{\text{coh}} & \mathbf{X}^I_{\text{coh}} \end{bmatrix} \qquad N_g = N_{g,reg} + N_{g,coh} \tag{3.15}$$

where $\mathbf{X}_{\text{reg}}$ and $\mathbf{X}_{\text{coh}}$ collect the strain fluctuations located outside and inside the cohesive bands, respectively. The right hand side matrix in (3.15) emphasizes the double partition performed in accordance with elastic-inelastic regimes.

After some additional manipulations, the corresponding set of orthonormal basis is obtained as:

$$\boldsymbol{\Psi}_G = [\boldsymbol{\Psi}_G^E \;\; \boldsymbol{\Psi}_G^I] \tag{3.16}$$

The number of reduced basis in $\boldsymbol{\Psi}_G$ is: $n_\varepsilon = 6 + n^I_{\varepsilon,reg} + n^I_{\varepsilon,coh}$, where the values $n^I_{\varepsilon,reg}$ and $n^I_{\varepsilon,coh}$ are obtained from the SVD applied to the projected inelastic snapshots. Additional details can be found in Sec. 3.2.2 in Paper #3.

3.2.3.2 *Formulation of the reduced order model*

Once the reduced basis $\boldsymbol{\Psi}_G$ is known, the strain fluctuations are interpolated as a linear combination of the elements of this basis as:

$$\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{\mu}}(\boldsymbol{y}, t) = \sum_{i=1}^{n_\varepsilon} \boldsymbol{\Psi}_i(\boldsymbol{y}) \cdot c_i(t) = \boldsymbol{\Psi}_G(\boldsymbol{y}) \cdot \boldsymbol{c}(t) \tag{3.17}$$

where each element $\boldsymbol{\Psi}_i$, of the basis $\boldsymbol{\Psi}_G$, is a microscale strain fluctuation mode and the vector of time dependent coefficients $\boldsymbol{c}(t) = [c_1, c_2, \cdots, c_{n_\varepsilon}]$ ($\boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}$) represents their corresponding amplitudes (the actual unknowns of the problem). In the same way, the variations of the microscale strain fluctuations are expressed similar to (3.17) as:

See: ***Sec.3.1 Paper #3***

$$\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{\mu}}(\boldsymbol{y}, t) = \boldsymbol{\Psi}_G(\boldsymbol{y}) \cdot \hat{\boldsymbol{c}}(t) \tag{3.18}$$

The problem solved in the online stage is then the following:

PROBLEM III (ROM) (RVE *saddle point problem)*:
Given the macroscale strain, $\boldsymbol{\varepsilon}$, find $\mathbf{c} \in \mathbb{R}^{n_\varepsilon}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{n_\sigma}$ satisfying:

$$\{\boldsymbol{c}(\boldsymbol{\varepsilon}, \boldsymbol{\mu}), \boldsymbol{\lambda}(\boldsymbol{\varepsilon}, \boldsymbol{\mu})\} = \arg\Big\{\min_{\boldsymbol{c}\in\mathbb{R}_{n_\varepsilon}} \max_{\boldsymbol{\lambda}\in\mathbb{R}^{n_\sigma}} \Pi_{[\boldsymbol{\varepsilon},\boldsymbol{\mu}]}(\boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\lambda})\Big\} \tag{3.19}$$

$$\text{with} \quad \Pi_{[\boldsymbol{\varepsilon},\boldsymbol{\mu}]}(\boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\lambda}) = \int_{\mathcal{B}_\mu} \varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c})\, d\mathcal{B}_\mu + \boldsymbol{\lambda}^T \left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}\, d\mathcal{B}_\mu\right) \boldsymbol{c}$$

such that:

$$\dot{\boldsymbol{\mu}} = f(\boldsymbol{\varepsilon}_\mu, \boldsymbol{\mu})$$

The optimality conditions for the problem above yield:

$$\frac{\partial}{\partial \boldsymbol{c}} \Pi_{[\boldsymbol{\varepsilon},\boldsymbol{\mu}]}(\boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\lambda}) = \int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c})\, d\mathcal{B}_\mu + \left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T\, d\mathcal{B}_\mu\right) \boldsymbol{\lambda} = \mathbf{0} \tag{3.20}$$

$$\frac{\partial}{\partial \boldsymbol{\lambda}} \Pi_{[\boldsymbol{\varepsilon},\boldsymbol{\mu}]}(\boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\lambda}) = \left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T\, d\mathcal{B}_\mu\right) \boldsymbol{c} = \mathbf{0} \tag{3.21}$$

which, expressed in matrix notation, yield:

$$\boldsymbol{\Psi}_G^T[\mathbb{W}] \left([\boldsymbol{\sigma}_\mu(\boldsymbol{c})]_G + [\boldsymbol{\lambda}]_G\right) = \mathbf{0} \tag{3.22}$$

$$[W] \cdot [\boldsymbol{\Psi}_G]\boldsymbol{c} = \mathbf{0} \tag{3.23}$$

where $[\boldsymbol{\sigma}_\mu(\boldsymbol{c})]_G \in \mathbb{R}^{(n_\sigma \cdot N_g)}$ is the column vector constituted by piling-up the N_g stress vectors, $\boldsymbol{\sigma}_\mu(\boldsymbol{c}) \in \mathbb{R}^{(n_\sigma)}$, evaluated at the integration Gauss points. The column vector $[\boldsymbol{\lambda}]_G$ is also the pilled-up of N_g repeated values of the same constants vector $\boldsymbol{\lambda} \in \mathbb{R}^{(n_\sigma)}$. The square diagonal matrix $[\mathbb{W}] \in \mathbb{R}^{(N_g n_\sigma \times N_g n_\sigma)}$, and the rectangular matrix $[W] \in \mathbb{R}^{(n_\sigma \times N_g n_\sigma)}$, collect the Gauss weights: $w_1, w_2, ..., w_{N_g}$, which for plane strain cases ($n_\sigma = 4$) are distributed in sub-block matrices $\mathbb{W}_i \in \mathbb{R}^{(4\times 4)} (i = 1, 2, 3, 4, ..., N_g)$, as follows:

See: ***Sec. 3.2 Paper #3***

$$\mathbb{W} = \begin{bmatrix} \mathbb{W}_1 & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & \mathbb{W}_1 & \dots & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & \mathbb{W}_{N_g} \end{bmatrix}; \quad [W] = [\mathbb{W}_1 \dots \mathbb{W}_{N_g}] \tag{3.24}$$

being $\mathbb{W}_i$ a matrix with the corresponding i-th Gauss weight placed at the diagonal.

The unknowns for the reduced order model are, the modal amplitudes $\boldsymbol{c}(t)$, and the Lagrange multiplier $\boldsymbol{\lambda}$.

Considering the system of equations (3.22) and (3.23) for $\boldsymbol{c}$ and $\boldsymbol{\lambda}$, it could be expected that this problem, of $n_\varepsilon + n_\sigma$ equations, should be less computationally costly, than the HF model. However, this is not the case: the numerical simulations with the ROM model do not substantially reduce the computational cost, and little (or even smaller than one) speedups are obtained. This fact highlights that the actual bottleneck for fast online computation is not the solution of the balance equations but, rather, the determination of the stresses, internal forces and stiffness matrices at every integration point of the underlying finite element mesh. Therefore, an additional technique is proposed to reduce the amount of integration points in which the constitutive equation is evaluated.

3.3 Numerical Integration: Reduced Order Quadrature Technique (ROQ)

Attention is then focused on reducing the computational cost arisen by the use of a classical Gauss quadrature for the numerical integration of the optimality conditions (Eqs. (3.20) and (3.21)).

For this purpose, a *reduced integration technique* has been developed by resorting to a nonconventional method, termed Reduced Optimal Quadrature (ROQ), to integrate the term involving the microscale free energy $\varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c})$ in (3.19):

$$\int_{\mathcal{B}_\mu} \varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\mu})\, d\mathcal{B}_\mu \approx \sum_{j=1}^{N_r} \varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\mu})\, w_j := \int_* \varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\mu})\, d\mathcal{B}_\mu \tag{3.25}$$

Where $\int_*(\cdot)\, d\mathcal{B}_\mu$ stands for he ROQ.

The ROQ technique is based on selecting, from the initial set of "Gauss" sampling points, and through an adequate algorithm, an equivalent subset of sampling points $z_j;\, j = 1, 2, \ldots, N_r$, and their new corresponding weights $w_j;\, j = 1, 2, \ldots, N_r$. The success of the reduced integration numerical scheme, in front of the conventional Gauss quadrature, lies on the fact that it is possible to reduce notably the number of involved quadrature points to $N_r \ll N_g$, being N_g the number of integration points for the Gauss quadrature scheme, keeping under strict control, or even reducing to zero, the numerical error introduced by the reduced quadrature rule. Then, the microscale potential energy in Eq. (3.25), is re-expressed as:

$$\Pi^*_{[\boldsymbol{\varepsilon},\boldsymbol{\mu}]}(\boldsymbol{\Psi c}, \boldsymbol{\lambda}) = \int_* \varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi c})\, d\mathcal{B}_\mu + \boldsymbol{\lambda}^T \left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}\, d\mathcal{B}_\mu \right) \boldsymbol{c} \tag{3.26}$$

In consequence, the corresponding optimality conditions (equilibrium equations) to be solved during the online stage are:

PROBLEM IV (HPROM) (*Microscale reduced saddle-point problem)*: Given the macroscale strain, $\boldsymbol{\varepsilon}$, find $\boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{n_\sigma}$ satisfying:

$$\underbrace{\int_* \boldsymbol{\Psi}^T \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi c})\, d\mathcal{B}_\mu}_{\text{Reduced o. quadrature}} + \underbrace{\left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T\, d\mathcal{B}_\mu \right)}_{\text{Gauss quadrature}} \boldsymbol{\lambda} = \mathbf{0} \tag{3.27}$$

$$\underbrace{\left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T\, d\mathcal{B}_\mu \right)}_{\text{Gauss quadrature}} \boldsymbol{c} = \mathbf{0} \tag{3.28}$$

A similar procedure could also be used for the integral terms (underlined as "Gauss quadrature") in Eqs. (3.27) and (3.28). However, this would not produce a substantial computational cost gain due to the fact that those terms are constant (not depending neither on the unknowns of the problem nor on the constitutive internal variables). They are required to be integrated only once, via the standard Gauss quadrature, and the result can be stored, and retrieved when necessary, during the online stage execution.

3.3.1 *Reduced Optimal Quadrature*

In spite that the goal of the ROQ is to develop a reduced cost interpolation scheme as a general framework for both static and dynamic problems, attention is focussed here on the multiscale quasi-static fracture problems. The minimum number of quadrature points providing an admissible integration error in the free energy integral, (3.25) is based on the optimal linear expansion of φ_μ in terms of the free energy modes $\boldsymbol{\Phi}_i\, (i = 1, 2, 3, \ldots, n_\varphi)$ and its corresponding amplitudes

f_i. Thus, a similar expression to Eq. (3.17)) can be adopted for constructing the reduced microscale strain fluctuations, as follows:

$$\varphi_\mu(\boldsymbol{\Psi}(\boldsymbol{y})\boldsymbol{c}) = \sum_{i=1}^{n_\varphi} \boldsymbol{\Phi}_i(\boldsymbol{y}) f_i(\boldsymbol{\varepsilon}, \boldsymbol{c}, \boldsymbol{\mu}) \tag{3.29}$$

With the previous approximation in hand, the total microscale free energy can be expressed as:

$$\int_{\mathcal{B}_\mu} \varphi_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\mu})\, d\mathcal{B}_\mu \approx \sum_{i=1}^{n_\varphi} \underbrace{\left(\int_{\mathcal{B}_\mu} \boldsymbol{\Phi}_i(\boldsymbol{y})\, d\mathcal{B}_\mu \right)}_{\text{Gauss quadrature}} f_i(\boldsymbol{\varepsilon}, \boldsymbol{c}, \boldsymbol{\mu}) \approx \sum_{i=1}^{n_\varphi} \underbrace{\left(\int_{*} \boldsymbol{\Phi}_i(\boldsymbol{y})\, d\mathcal{B}_\mu \right)}_{\text{Reduced o. quadrature}} f_i(\boldsymbol{\varepsilon}, \boldsymbol{c}, \boldsymbol{\mu}) \tag{3.30}$$

3.3.2 *A Greedy algorithm for obtaining a reduced quadrature rule*

In order to obtain the reduced optimal numerical quadrature rule, the following optimization problem is considered:

OPTIMIZATION PROBLEM : Given the expanded reduced basis $\boldsymbol{\Phi}$, and the set of sampling points $\mathcal{S} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{N_g}\}$, find $\boldsymbol{\omega} \in \mathbb{R}_+^{N_r}$ and $\mathcal{Z} \in \mathbb{N}^{N_r}$ satisfying:

$$(\boldsymbol{\omega}, \mathcal{Z}) = \arg\min_{(\mathbf{w}\in\mathbb{R}_+^m, \bar{\mathcal{Z}}_g \in \mathcal{B}_\mu)} \sqrt{\sum_{i=1}^{n_\varphi} (e_i)^2 + (e_{\text{vol}})^2} \tag{3.31}$$

being:

$$e_i := \sum_{j=1}^{N_r} \omega_j \boldsymbol{\Phi}_i(\bar{\boldsymbol{y}}_j) - \int_{\mathcal{B}_\mu} \boldsymbol{\Phi}_i(\boldsymbol{y})\, d\mathcal{B}_\mu \qquad e_{\text{vol}} := \sum_{j=1}^{N_r} \omega_j - \Omega_\mu \tag{3.32}$$

Where, e_i and e_{vol} stand for the error committed through the reduced integration of every free energy reduced basis function, and the error in the integration of the volume Ω_μ, respectively. The resulting algorithm (described in the flowchart of **Box IV** in Sec. 5 of Paper #4 (Hernández et al., 2017)) returns a sub-set of optimal Gauss points, and the corresponding weights, that integrate exactly the basis $\boldsymbol{\Phi}(\boldsymbol{y})$ and, therefore, the free energy in Eq. (3.29).

See: Sec. 5 Paper #4

3.3.2.1 *Computation of the reduced basis functions*

Regarding the computation of the microscale energy reduced basis $\boldsymbol{\Phi}$, a SVD-based strategy is used in the off-line stage, similar to that described in Sec. 3.2.1.

The method is again based on the construction of a snapshots matrix, in this case, for the free energy, and the computation of its corresponding reduced basis via SVD. For this purpose, two options appear:

- To construct the microscale energy snapshots by collecting solutions of the ROM in (3.19). This strategy implies the following actions: (a) the ROM is constructed considering a number of microscale strain fluctuation modes n_ε. In consequence, the obtained energy snapshots matrix $\mathbf{X}^\varphi$ provides a reduced basis, which spans a space determined by the selected set of strain fluctuation modes.

- To construct the microscale energy snapshots by collecting solutions of the HF model in (3.11). This strategy computes *simultaneously* the microscale energy and the strain fluctuation snapshots. The resulting reduced bases are consequently independent from each other, but converging to the HF solution as the number of strain basis functions n_ε, and energy basis functions n_ψ, increase.

The first method is considered as the algorithmically *consistent* strategy. However, it is also more expensive than the second one. The reason relies on the fact that, in order to get the reduced basis for the microscale energy $\mathbf{\Phi}$, training trajectories have to be computed twice: a) First using the HF model to obtain the strain modes $\mathbf{\Psi}$, b) Second, using the ROM model to obtain the corresponding free-energy modes $\mathbf{\Phi}$.

Both strategies have been tested and both provide accurate results. However, the later, being the cheaper and simpler one, was adopted as the most convenient.

In summary, both the strains and the free energies of the microscale are sampled simultaneously at the off-line stage, for different sampling trajectories with the HF model, and a series of q_{snp} snapshots of energy, φ_μ, are evaluated and collected for each Gauss point. Then, the microscale energy snapshot matrix $\mathbf{X}^\varphi$ is built as:

$$\mathbf{X}^\varphi = [\mathbf{X}^\varphi_1, \mathbf{X}^\varphi_2, \cdots, \mathbf{X}^\varphi_{p_{snp}}] \in \mathbb{R}^{(N_g \cdot n_\sigma)\times p_{snp}}$$

$$\mathbf{X}^\varphi_k = [\varphi_\mu(\boldsymbol{y}_1), \varphi_\mu(\boldsymbol{y}_2), \ldots, \varphi_\mu(\boldsymbol{y}_{N_g})]^T_k \tag{3.33}$$

In accordance with the position of the Gauss point, in the finite element mesh, and following a similar procedure to that adopted in Eq. (3.15), this snapshot matrix is also partitioned into components associated to the domains $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,\text{coh}}$ and $\mathcal{B}_{\mu,\text{coh}}$ as:

*See: **Sec. 3.4** **Paper #3***

$$[\mathbf{X}^\varphi] = \begin{bmatrix} \mathbf{X}^{\text{reg}}_\varphi \\ \mathbf{X}^{\text{coh}}_\varphi \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\text{reg},E}_\varphi & \mathbf{X}^{\text{reg},I}_\varphi \\ \mathbf{X}^{\text{coh},E}_\varphi & \mathbf{X}^{\text{coh},I}_\varphi \end{bmatrix} \qquad N_g = N_{g,reg} + N_{g,coh} \tag{3.34}$$

and the SVD technique is then separately applied to both partitions of $\mathbf{X}^E_\varphi$ to obtain two distinct (orthogonal) bases, for the elastic regime of both subdomains:

$$\mathbf{\Phi}^E = \begin{bmatrix} \mathbf{\Phi}^{\text{reg},E} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}^{\text{coh},E} \end{bmatrix}; \quad \mathbf{\Phi}^{\text{reg},E} \in \mathbb{R}^{(N_{g,reg}\times \text{n}_{elas})}; \; \mathbf{\Phi}^{\text{coh},E} \in \mathbb{R}^{(N_{g,coh}\times \text{n}_{elas})} \tag{3.35}$$

The corresponding inelastic reduced basis functions are also computed via SVD, following a procedure similar to the one described in Sec. 3.2.1. The complete reduced basis for the energy field $\boldsymbol{\Phi}$, is made of the union of $\boldsymbol{\Phi}^E$ and $\boldsymbol{\Phi}^I$:

$$\boldsymbol{\Phi} = [\boldsymbol{\Phi}^E \;\; \boldsymbol{\Phi}^I] \tag{3.36}$$

The number of basis vectors in $\boldsymbol{\Phi}$ is: $n_\varphi = 12 + n^I_{\varphi,reg} + n^I_{\varphi,coh}$, where the values of $n^I_{\varphi,reg}$ and $n^I_{\varphi,coh}$ are obtained from the solution of the SVD applied to the inelastic projected snapshots.

3.4 Numerical assessment and approximation errors

The accuracy of the reduced models, ROM and HPROM, depends on several aspects. In order to assess it, three different sets of tests are done:

- *Consistency tests:* A set of trajectories already sampled with the HF model during the microscale sampling process in the off-line stage, are re-evaluated using the ROM and HPROM strategies. This kind of assessment provides an estimation of the quality and richness of the basis to reproduce the stored snapshots, and the accuracy of the ROQ scheme. It is expected that the error with respect to the HF solutions (*consistency error*) tends to zero as the number of considered modes, for each reduction strategies, are increased.
- *Accuracy tests:* Similarly to the aforementioned consistency tests, the representative cell is subjected to a *unsampled loading trajectory*. In these cases, the quality of the reduced bases and the ROQ scheme is also tested. In contrast with the previous case, unsampled trajectories during the off-line stage, are not supposed to be exactly captured, due to the underlying *sampling error*.
- *Multiscale structural tests:* These kind of tests are based on multiscale benchmarks. The aim is to evaluate the accuracy of the solutions when the proposed overall HPROM strategy is applied, and to obtain the corresponding speed-ups.

Details on this issue can be found in Sec. 5 in Paper #3.

3.5 Representative example

A squared microscale model, made of a matrix and randomly distributed aggregates, is devised and tested (see figure 14) to simulate the microstructure of a cementitious-like material (concrete). Relevant details about the finite element model are presented in Table 1. To mimic the concrete material response, the failure cell is modeled with three components: *aggregates*, which are assumed to be elastic, *bulk matrix*, also assumed elastic, and *interfaces* (matrix-matrix and matrix-aggregates), which are modeled with cohesive-band equipped with an

isotropic damage constitutive law. The properties of the components in the microscale are defined in Table 2.

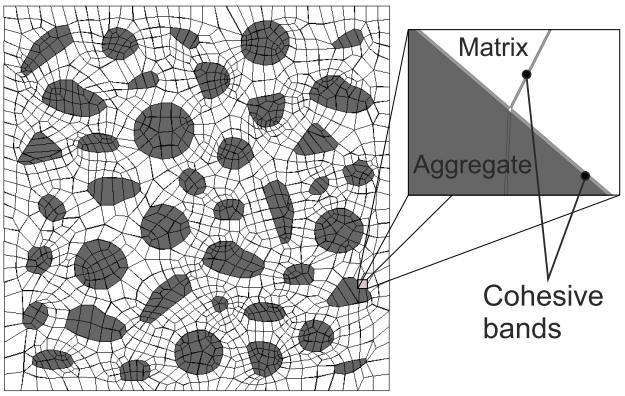


Figure 13: Failure cell

Number of elements	Number of D.o.f.	Number of Cohesive Bands	Total number of Gauss points (N_g)
5409	14256	2189	21636

Table 1: Discretization of the Microscale

3.5.1 *Design of the HPROM Strategy*

Figure 14 shows the summary of a number of results obtained by running the HPROM strategy in a number of cases for the microstructure in Fig. 13.

In general terms, Figure 14 can be used as an "abacus" for a-priori selection by the user of the HPROM strategy in a multiscale problem (for a given microstructure at the RVE). For instance, by selecting the admissible error (3,5%) in the top figure, the number of strain modes $n_\varepsilon = 80$) is obtained. Entering in the lower plot, with this result ($n_\varepsilon = 80$), one obtains the suitable number of integration points $OQN \simeq 200$ and the expected speedup $\simeq 110$.

The availability of a *catalog* (constructed off-line) for a specific RVE microstructure, allows the user's *a-priori* selection of the appropriated HPROM strategy, by balancing the admissible error vs. the desired speedup.

	E_μ [MPa]	ν_μ	$\sigma_{\mu u}$ [MPa]	$G_{\mu f}$ [N/m]
Elastic matrix	1.85×10^4	0.18	—	—
Elastic aggregate	3.70×10^4	0.18	—	—
Cohesive bands of matrix-matrix interface	1.85×10^4	0.18	2.60	140
Cohesive bands of matrix-aggregate interface	1.85×10^4	0.18	—	—

Table 2: Material properties of the sampled microcell. Properties are: E_μ (Young's modulus), ν_μ (Poisson ratio), $\sigma_{\mu u}$ (ultimate tensile stress) and $G_{\mu,f}$ (fracture energy).

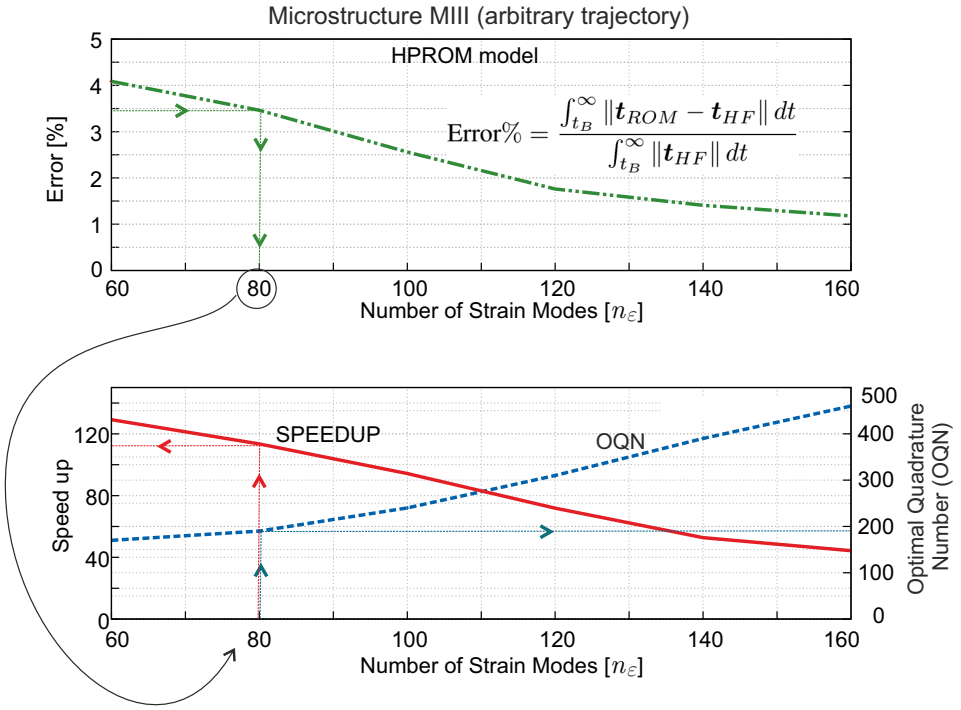


Figure 14: HPROM design diagrams. *Top*: HPROM error in terms of the number of strain modes. *Bottom:* OQN and obtained speedup in terms of the number of strain modes. By selecting the admissible error (say 3,5%) in the upper diagram, one obtains the requested number of strain modes, $n_\varepsilon = 80$. Entering with this result in the lower diagram one obtains the suitable number of integration points ($OQN = 200$) and the resulting speedup ($speedup = 110$).

3.5.2 *Multiscale crack propagation problem: L-shaped panel*

The test shown in Figure 15 is a benchmark commonly used for testing macroscale propagating fracture models. This concrete-like specimen is con-

sidered here to test the qualitative results and convergence properties of the proposed HPROM approach, when utilized in real FE^2 multiscale crack propagation problems.

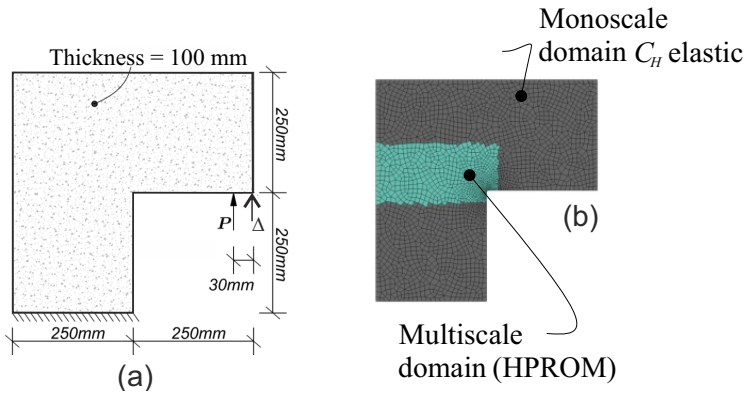


Figure 15: L-shaped panel: a) Specimen geometry; b) Finite element mesh

The geometry of the simulated specimen is depicted in Figure 15-a. As shown in Figure 15-b, the domain of the L-shaped panel is split into two domains: 1) the multiscale domain (with 721 elements) corresponding to the region where the crack is expected to propagate, modeled with the HPROM of the microstructure depicted in Figure 13, and 2) the remaining part of the panel, which is modeled with an elastic monoscale approach (using 1709 elements), where the elasticity tensor is obtained through an *elastic homogenization* of the micro-structure elastic properties. Even for this (rather coarse) multiscale problem, the high fidelity HF computational solution is extremely costly to handle, until the point that, with the available computational resources[5], it was not possible to display the complete action-response curve (in Fig. 16).

However, the remaining structural responses in Figure 16, obtained through a number of HPROM strategies, involve very reasonable computational costs, and they were obtained in advance *with no previous knowledge of the HF* results. The accuracy is very good, and a response indistinguishable from the HF can be obtained 60 times faster ($speedup = 60$). A less accurate response, but with a fairly good agreement with the HF can be obtained with $speedup = 130$.

In Fig. 17, the evolution of the microscale crack opening is shown. It is worth noting that, both, the microscale failure mechanism and displacement jump vary along the macroscale in agreement with the crack propagation direction observed at the macroscale.

5 A cluster of 500 cores, is used. The multiscale finite element code is written in Matlab©environment.

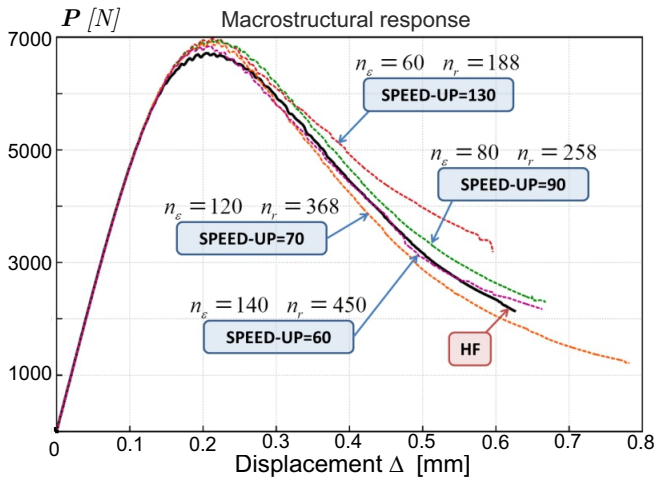


Figure 16: L-shaped panel: Structural responses in terms of force P vs. vertical displacement Δ, for different RVE HPROM strategies, and obtained speed-ups.

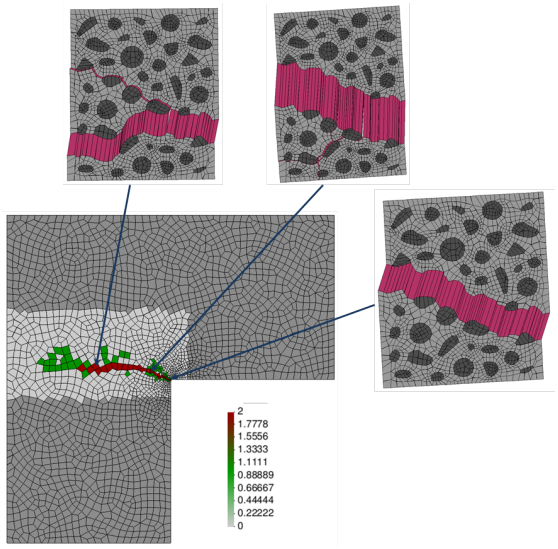


Figure 17: L-shaped panel: microscale crack activation along the crack-path field, using $n_\varepsilon = 80$ and $n_r = 258$.

This illustrates the new paradigm that is set and the computational possibilities open by the HPROM strategies in computational multiscale modeling explored in this Thesis.

4

DISCUSSION, CONCLUSIONS AND FUTURE WORK

4.1 Discussion and Conclusions

4.1.1 *Overview of the work*

Multiscale modeling is foreseen to become a key approach to enable the next wave of design paradigms for engineering materials and structures. Indeed, it has an excellent potential to account for the physical links between different scales, involving the diverse phenomenologies intervening in the mechanical response of materials (grains, particles, defects, inclusions, etc.).

Quoting from a report by a group of experts to the US National Science Foundation (Oden et al., 2006):

". . . . In recent years, a large and growing body of literature in physics, chemistry, biology, and engineering has focused on various methods to fit together simulation models of two or more scales, and this has led to the development of various multi-level modeling approaches. . . .. To date, however, progress on multiscale modeling has been agonizingly slow. Only a series of major breakthroughs will help us establish a general mathematical and computational framework for handling multiscale events and reveal to us the commonalities and limitations of existing methods . . . .".

In this sense, the effort invested in developing and using multiscale models, has been, in many cases, fruitless, due to the involved computational cost in this kind of methodologies. This limitation becomes a bottleneck for multiscale modeling, usually discarded, or, relegated to the availability of supercomputers, and, therefore, not always accessible to the whole computational mechanics community.

In addition, while multiscale models exhibiting material hardening behavior have widely been studied, multiscale models dealing with material softening behavior are in an early stages of development.

Therefore, the development of a reliable, minimally intrusive multiscale fracture models becomes a crucial task, not only in order to have a robust and consistent multiscale fracture numerical tools, but also for developing their related reduced order models that allow their use in complex cases that can be used for industrial purposes, with an affordable cost. These are the fundamental reasons for the research and development about this issues.

A sketch of the overall work carried out in this thesis is shown in Figure 18. In there, contributions are chronologically numbered and highlighted with a blue arrow. Contributed papers are numbered from P1, corresponding to the the first contribution (Paper #1), to P6 (Paper #6) the last one; in this context, CB means *Chapter in Book*.

In what follows, they are specifically commented, and the corresponding conclusions and achievements, are presented.

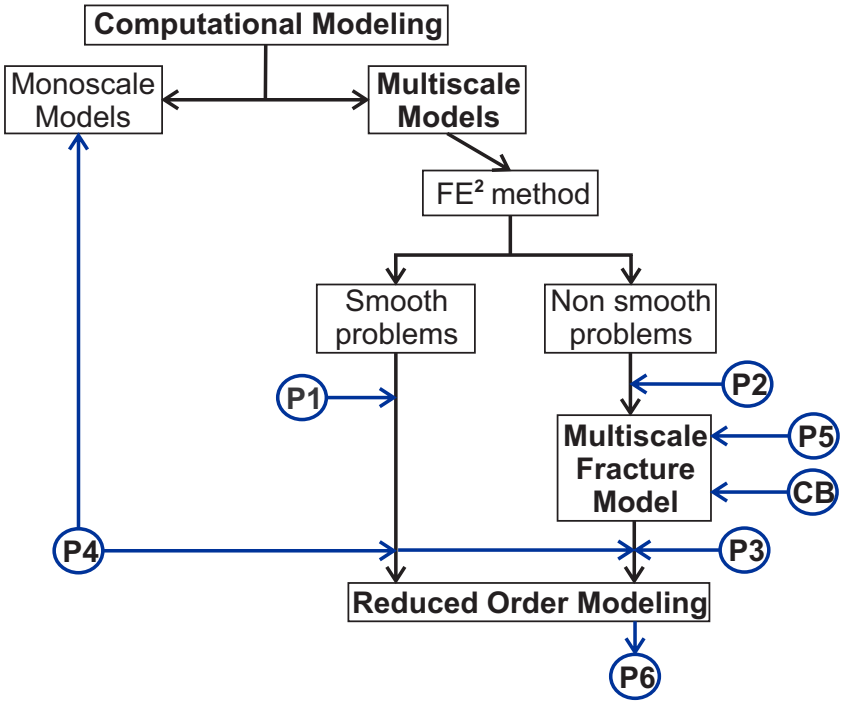


Figure 18: Global Flow Chart of the Thesis

4.1.2 *Paper #1: J. A. Hernández, J. Oliver, A. E. Huespe, M. A. Caicedo, J. C. Cante. High-performance model reduction techniques in computational multiscale homogenization, Computer Methods in Applied Mechanics and Engineering - 2014, Volume 276, Pages 149–189*

This Article presents the first research developments in this Thesis on Model Order Reduction (MOR) techniques applied to multiscale modeling. The scope of this publication is limited to smooth problems and exclude fracture processes. Techniques like *interpolation* methods via HPROM have been studied.

The concept of a two-stage reduction (ROM-HPROM) is presented. The first reduction, denoted as ROM, is performed via POD, taking the displacement fluctuation field as a primal variable. The second reduction, denoted as HPROM, is performed via *interpolation* techniques (DEIM) of the microscale stress field.

It is shown that the interpolation-based HPROM obtained in this way, leads to an ill-posed mathematical formulation when the reduction process involves an interpolant constructed using POD modes provided by the primal variable (microscale stress field). This issue has been studied in the paper, and a robust and consistent solution has been proposed.

An additional aspect in this contribution, is the selection of the interpolation points for the stress field. These interpolation points are chosen guided, not only by accuracy requirements, but also by stability considerations.

The method of selection of the interpolation points (*Greedy Algorithm*) is, at the present, an intensive research field. However, although in the literature there are several alternative algorithms, none of them offers a robust and general treatment to handle this Thesis purposes.

Different measures of error have been presented to test the accuracy and the convergence. The work is assessed by the homogenization of a highly complex porous metal material. The results show that, the speed-up factor is about three orders of magnitude, for an error in stress smaller than 10%.

As conclusions of this work, it can be stated that:

- The hyper-reduced form of the RVE equilibrium equation has a conceptual simplicity, and the corresponding solution scheme is also very simple to implement. Taking as departure point an existing FE code, one only has to replace the typical loop over elements in the FE code by a loop over the pre-selected sampling points.
- Storage of history data (internal variables) is only required at the pre-selected sampling points.
- Consistency with respect to the HF solution is achieved when the amount of reduced order basis functions, for both reductions, is increased.

In consequence, the numerical results suggest that this HPROM provides accurate solutions to problems exhibiting hardening behavior. However, some questions need to be further analyzed. For example:

- *Can the model order reduction techniques capture the RVE solution in problems displaying crack propagation processes?*
- *Will the number of modes necessary to accurately replicate its solution, increase with the number of potential crack paths (i.e., with the geometrical complexity of the RVE)?*

These questions motivated the next research work: i. e.the development of a reduced order model applied to problems exhibiting discontinuous fields, and in particular, the case of the quasi-brittle fracture.

4.1.3 *Paper #2: J. Oliver, M. Caicedo, E. Roubin, A. E. Huespe, J. A. Hernández. Continuum approach to computational multiscale modeling of propagating fracture, Computer Methods in Applied Mechanics and Engineering - 2015, Volume 294, Pages 384–427*

This work presents a novel approach to two-scale modeling of propagating fracture, based on computational homogenization FE^2. The specific features of this contribution are:

- Extends the homogenization paradigms for smooth problems presented in Paper #1 – typically the Hill–Mandel principle and the stress-strain homogenization procedures – to non-smooth problems.
- In both scales of analysis, a continuum (stress–strain) constitutive relationship is considered, instead of making use of the most common discrete traction/separation-law. This contributes to provide a unified setting for smooth and non-smooth, problems. This is achieved by resorting to the well-established Continuum Strong Discontinuity Approach (CSDA).
- As for the multiscale modeling issue, it involves a new and crucial additional entity: a *characteristic length*, which is point-wise obtained from the geometrical features of the failure mechanism developed at the low scale. As a specific feature of the presented approach, this characteristic length is exported, in addition to the homogenized stresses and the tangent constitutive operator, to the macroscale, and considered as the bandwidth of a propagating strain localization band, at that scale.
- Consistently with the characteristic length, a specific computational procedure is used for modeling the onset and propagation of this localization band at the macro-scale. It is based on the crack-path-field and strain injection techniques, developed (Oliver et al., 2014). This computational procedure ensures the macroscale mesh-size and microscale RVE-size objectivity of the results, and a consistent energy dissipation at both scales.

The approach has been validated and tested using classical benchmarks in fracture mechanics. After validation, some aspects of the proposed approach can be emphasized:

- From the computational point of view, the proposed technique is minimally invasive with regards to procedures well established in the literature on multiscale modeling of materials. In fact, in terms of the computational homogenization, the proposed approach displays no substantial difference with respect to the ones used for smooth (continuous) problems. In terms of material failure propagation, existing algorithms for monoscale crack propagation modeling can be easily extended to this multiscale case. In addition, this multiscale approach is extensible to other families of propagation schemes.
- Consistency has been assessed by comparison, with a number of representative cases, through results obtained with the proposed FE^2 and the ones obtained by Direct Numerical Simulation (DNS). In the same way, objectivity have been also checked in terms of finite element mesh size and bias, at the macro-scale, and the failure-cell (size and shape) at the micro/meso scale.

As mentioned, multiscale computational fracture problems and their extension to 3D cases, face a great challenge: the enormous involved computational cost. In consequence, next step is the development of a reduced order model aiming at diminishing the computational burden of the developed multiscale fracture model.

4.1.3.1 *HPROM for hardening processes applied to quasi-brittle fracture*

The reduced order model described in Paper #1 was used as a first attempt. However, the results were very unsatisfactory. The conclusions of this interpolation-based approach to multiscale reduced order modeling in fracture cases were:

- The reduced basis for the microscale displacement fluctuations obtained via SVD does not make a clear distinction between smooth and non-smooth domains. Hence, a large set of displacement modes (considerably larger than the one requested in hardening problems) has to be used to retrieve accurate solutions.
- The stress snapshots, taken from high localized strain stages with released near-to-zero stresses, are numerically neglected by the SVD[1], this taking interpolation-based HPROM methods to fail in reproducing the post-critical stages.
- To obtain a good approximation with the HPROM, it is necessary to largely increase the number of displacement and stress modes, but, in this scenario, the interpolation method is not longer robust.

1 The SVD strategy, gives importance to repeated snapshots, and mainly, snapshots which euclidean norm is considerably high.

This suggests additional research and exploration of specific model order reduction techniques for multiscale fracture problems.

4.1.4 *Paper #3: J. Oliver, M. Caicedo, A. E. Huespe, J. A. Hernández, E. Roubin. Reduced Order Modeling strategies for Computational Multiscale Fracture, Computer Methods in Applied Mechanics and Engineering - 2016, Volume 313, Pages 560–595*

This article proposes a set of new computational techniques to solve multiscale problems via HPROM techniques. These techniques have been applied to the multiscale model described in 4.1.3, and they are summarized next:

- A *domain separation* strategy. The RVE is split into the *regular domain* (made of the elastic matrix and possible inclusions) and the *singular domain* (the cohesive bands exhibiting a softening cohesive behavior). These are designed to provide a sufficiently good representation of the microscopic fracture and of its effects on the homogenized material behavior (Oliver et al., 2015). The distinct constitutive behavior of both domains suggests a specific ROM strategy for each of them, in order to obtain a reduction strategy with information on the mechanical variables in every specific subdomain. Therefore, selection of the ROM low-dimensional projection space is made independently for each of these domains.

- In combination with the previous strategy, the ROM for the RVE is formulated in an *unconventional manner* i. e.: in terms of the strain fluctuations rather than in terms of the conventional displacement fluctuations. The reduced strain fluctuation space is spanned by basis functions satisfying, by construction, the strain compatibility conditions, this guaranteeing that, after reduction, the solution in the strain fluctuation space also satisfies the strain compatibility.

- A specific Reduced Optimal Quadrature (ROQ) is used as a key technique to obtain relevant computational cost reduction from the ROM. This technique consists of replacing the standard Gauss integration rule by an optimal quadrature, involving much less sampling points, has been proposed in other works (Farhat et al., 2015; Hernández et al., 2017) as an ingredient of HyPer-Reduced Order Modeling (HPROM) strategies. In these works, the reduced numerical integration technique is applied to the variational equations of the problem (i.e. internal forces, involving n-dimensional vector entities) whereas, in the herein proposed approach, a similar reduced integration technique is applied, again unconventionally, to the primitive problem, i.e: the functional (a scalar entity) in the micro-scale saddle-point problem that supplies the RVE variational equations. In the present proposal, this functional turns out to be the stored energy (free energy) at the RVE, which, being a scalar entity, is much less demanding in terms of the integration rule.

In a first validation stage, in order to test the sensibility of the reduction techniques, a set of three different failure cells have been tested, by increasing the complexity and, consequently, the amount of cohesive bands. A-priori and a-posteriori errors analysis are performed, showing that, increasing the complexity (number of involved operations) at the microscale, the amount of required strain and free energy modes increases only slightly for a given error. This is a clearly promising scenario.

Finally, this reduced multiscale model was also validated and tested with the *L-Shape Panel* test, comparing the solution with the one given by the HF (obtained with the approach described in the Paper #2), and analyzing the impact on the use of different amounts of reduced order basis functions of both, the strain fluctuations and the free energy.

Several aspects of the proposed methodology can be highlighted as new contributions:

- The RVE domain separation technique: to account for distinct constitutive models used at the RVE and take the maximum advantage of this distinction.
- A strain-based formulation of the variational RVE problem allowing a simpler application of the previous technique, without the need of introducing compatibility constraints.
- A specific sampling program, for the construction of the sets of snapshots in the off-line stage of the HPROM procedure, in accordance with the rest of elements of the proposed strategy.
- The Reduced Optimal Quadrature (ROQ) technique, which resorts to the primitive formulation of the RVE problem as a saddle-point problem.

At this point it can be argued that only idealized, two-dimensional, problems have been considered. The real interest of many multiscale modeling problems residing on actual three-dimensional problems, the following question arises:

to what extent these techniques can be extended to three-dimensional problems, where the involved RVE complexity and the associated computational cost can be two or three orders of magnitudes larger than in 2D problems?

In Fig. 19, the results obtained from different kind of 2D microscale morphologies are presented. They show a very relevant property: **the obtained speedup "scales" linearly with the problem-complexity**. Therefore one could think of achievable values of $\approx 10^4$–10^5 for the speedup in 3D problems. This fact (in conjunction with, the additional usage of HPC procedures), could turn affordable 3D multiscale fracture modelling.

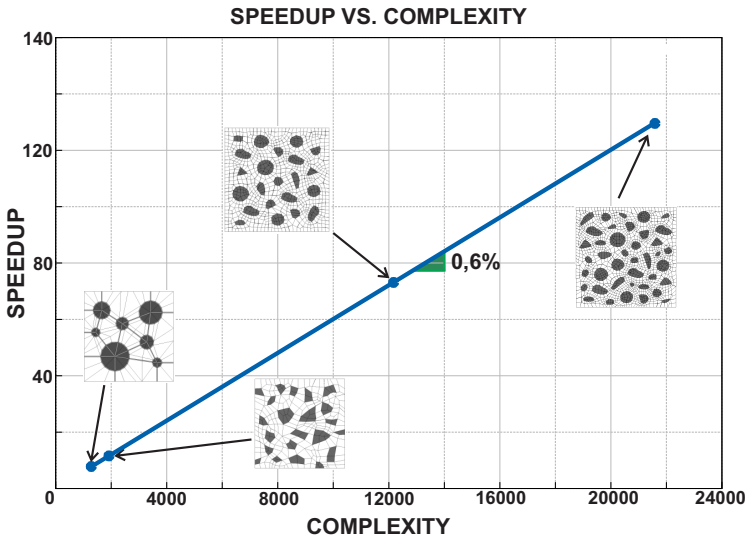


Figure 19: Speedup scalability.

4.1.5 *Paper #4: J. A. Hernández, M. A. Caicedo, A. Ferrer. Dimensional hyper-reduction of nonlinear finite element models via empirical cubature, Journal of Computer Methods in Applied Mechanics and Engineering - 2016, Volume 313, Pages 560–595*

This work has been developed in combination with the reduced order model for non-smooth problems (see 4.1.4). The main objective is to develop the algorithmic procedure in a general setting to be applied to different problems involving integral operators that can be sampled. Not only problems involving multiple scales can be analyzed, but also monoscale (static and dynamic) problems based on the Finite Element method.

It is presented a general framework for the dimensional reduction in terms of numbers of degrees of freedom as well as number of integration points of nonlinear parametrized finite element models.

As in previous cases (see 4.1.2 and 4.1.4), the reduction process is divided into two sequential stages, the first consists of a Galerkin projection of the strain fluctuations, via POD, and the second consists of a novel cubature rule also used in 4.1.4. In this case, this method is deeply studied and analyzed. The distinguish features of the proposed method to be highlighted are:

- The minimization method is set in terms of orthogonal basis vectors (obtained via Singular Value Decomposition SVD) rater than in terms of snapshots taken from the integrand.
- The volume of the domain is exactly integrated.

- The selection algorithm does not require solve, in all iterations, a non-negative least-squares problem to obtain positive weights.

This model is tested through two structural examples, (quasi-static bending, and resonant vibration of elasto-plastic composite plates). The total amount of integration points is reduced three order of magnitudes, this methodology can be applied to different primary variables, in 4.1.4, attention was focused on use the free energy to determine the reduced integration rule.

Several issues have been improved by this research: firstly, the *robustness*, one of the most attractive features of the proposed hyper-reduced order model (and in general, of all cubature-based ROMs) is that it preserves the spectral properties of the Jacobian matrix of the finite element motion equations. Secondly, the improved version of the *Empirical cubature method*, in contrast with other similar techniques proposed in the literature, in which the weights at almost all iterations of the greedy algorithm are calculated with a standard, unconstrained least-squares. In fact, the nonnegative least squares problem is included to filter out small negative weights caused by roundoff errors. And finally, for implementation purposes, the "format" of the finite element method is conserved.

4.1.6 *Paper #5: J. Oliver, M. Caicedo, E. Roubin, A. E. Huespe. Continuum Approach to Computational Multi-Scale Modeling of Fracture, Key Engineering Materials - 2014, Volume 627, Pages 349–352*

This work presents a brief summary of the two-scale approach for modeling failure propagation, providing details about propagation at the macro and micro levels. This publication is centered in exploring the applicability of the method to structural problems. The four-point bending and the *Nooru-Mohamed* problems have been chosen as benchmarks, taking the material properties form experimental tests.

In the case of the Nooru-Mohamed test, it has been shown, the influence of the horizontal load (shear force) in the microscale behavior, and the activation of different crack patterns, representing the macroscale changes in the crack propagation scheme. In the four-point bending test, it is displayed the influence on the macroscale propagation scheme, when critical failure mechanisms at the microscale are precluded.

4.1.7 *Paper #6: M. Caicedo, J. Oliver, A. E. Huespe, O. Lloberas-Valls. Model Order Reduction in computational multiscale fracture mechanics, Key Engineering Materials - 2016, Volume 713, Pages 248–253*

This work has a similar objective than the previous one. A brief summary about the reduced order model based on the two-scale approach for modeling failure propagation, has been presented. This work also presents a summary about the results obtained in the *L-Shaped Panel*, and the influence of the size of reduced

order basis functions (for strain fluctuations and free energy) is presented and analyzed.

4.1.8 *Chapter In Book (CB): J. Oliver, M. Caicedo, E. Roubin, J. A. Hernández, A. E. Huespe. Multi-scale (FE^2) analysis of material failure in cement-aggregate-type composite structures, Computational Modeling of Concrete Structures - 2014, Pages 39–49*

This work focuses on exploring different issues of the two-scale approach for modeling failure propagation. Particularly, the total energy dissipation and its relation at both scales is analyzed in some specific fracture problems.

4.2 Ongoing work and future research lines

4.2.1 *Ongoing work*

- Most of the real industrial problems require 3D modelling. This is the reason because an immediate task is to extend all procedures developed in this Thesis to 3D problems. This extension is carried out in Kratos Multi-Physics, an open-source code developed at CIMNE (International Center for Numerical Methods in Engineering).
- The approach developed in Paper #3 is being extended to nonlinear geometrical multiscale problems. Considering an elasto-plastic constitutive model endowed, with hardening behavior, the goal is to study and analyze geometric bifurcation at the macroscale. Some early results have been obtained, exhibiting the potential uses of this methods in nonlinear geometric multiscale problems.
- Fracture processes of composite materials via multiscale modeling, are being studied by using the formulations developed in Paper #2 and Paper #3. The main goals are: to reproduce the experimental behavior of composite sublaminates of ultra-thin plies (Arteiro et al., 2014) and, to devise some design alternatives based on the material failure analysis, taking advantage of the reduced order model techniques.
- In order to obtain an optimum performance of the reduced order models developed in this Thesis, it is convenient to optimize the tasks performed in the *offline stage*, particularly, the Singular Value Decomposition performed after sampling the training trajectories. This method can be highly computational demanding in very fine meshes. Therefore, the study of SVD partitioned procedures, and iterative strategies are presently explored.

4.2.2 *Future research lines*

- Extension of the developed multiscale model to propagating fracture in non-linear dynamic cases. This includes modeling more complex phenomena like branching and multiscale dynamic processes. This field was also studied via monoscale phenomenological modeling in (Belytschko et al., 2003; Prabel et al., 2007; Linder and Armero, 2009; Lloberas-Valls et al., 2016). In addition, inclusion of non-linear hardening behavior, before the onset of material failure in the considered damage model, and consideration of other families of constitutive behavior, like plasticity, rate dependence etc., should be studied.

- The use, in the developed multiscale model for propagating fracture, of other crack propagation models at the microscale, either based on continuum methods (CSDA, non-local models or gradient-regularized models), or discrete methods (cohesive interfaces equipped with traction-separation laws).

- Extension of the reduced order model described in Paper #3 to other microscale failure methodologies, i. e., gradient damage models, Enhanced finite element methods (EFEM), etc. In the same way, the extension of the reduced order model to other multiscale strategies involving fracture processes, i. e., concurrent models (Lloberas-Valls et al., 2012).

CONTRIBUTED PUBLICATIONS

	NAME	JOURNAL	QUARTIL INDEX in JCR	IMPACT FACTOR
Paper #1	High-performance model reduction techniques in computational multiscale homogenization	Computer M. in Applied Mechanics and Engineering	Q1	3.467
Paper #2	Continuum approach to computational multiscale modeling of propagting fracture	Computer M. in Applied Mechanics and Engineering	Q1	3.467
Paper #3	Reduced Order Modeling strategies for Computational Multiscale Fracture	Computer M. in Applied Mechanics and Engineering	Q1	3.467
Paper #4	Dimensional hyper-reduction of nonlinear finite element models via empirical cubature	Computer M. in Applied Mechanics and Engineering	Q1	3.467
Paper #5	Continuum Approach to Computational Multi-Scale Modeling of Fracture	Key Engineering Materials	–	0.39
Paper #6	Model Order Reduction in computational multiscale fracture mechanics	Key Engineering Materials	–	0.39

Table 3: Scientific Contributions in specialized journals

	NAME	BOOK	CONGRESS
CB	Multi-scale (FE^2) analysis of material failure in cement-aggregate-type composite structures	Computational Modeling of Concrete Structures	EURO-C 2014

Table 4: Book Chapters

BIBLIOGRAPHY

E. Aifantis. On the microestructural origin of certain inelastic models. *ASME J. Engrg Mat. Tech.*, 106:326–334, 1984.

David Amsallem, Julien Cortial, Kevin Carlberg, and Charbel Farhat. A method for interpolating on manifolds structural dynamics reduced-order models. *International Journal for Numerical Methods in Engineering*, 80(9):1241–1258, 2009. ISSN 1097-0207. doi: 10.1002/nme.2681. URL http://dx.doi.org/10.1002/nme.2681.

S. An, T. Kim, and D. James. Optimizing cubature for efficient integration of subspace deformations. *ACM transactions on graphics*, 27(5):165, 2009.

A. Arteiro, G. Catalanotti, A.R. Melro, P. Linde, and P.P. Camanho. Micro-mechanical analysis of the in situ effect in polymer composite laminates. *Composite Structures*, 116:827–840, 2014. ISSN 0263–8223. doi: http://dx.doi.org/10.1016/j.compstruct.2014.06.014. URL http://www.sciencedirect.com/science/article/pii/S0263822314002839.

Maxime Barrault, Yvon Maday, Ngoc Cuong Nguyen, and Anthony T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667 – 672, 2004. ISSN 1631-073X. doi: http://dx.doi.org/10.1016/j.crma.2004.08.006. URL http://www.sciencedirect.com/science/article/pii/S1631073X04004248.

Z. Bazant. Can Multiscale-Multiphysics Methods Predict Softening Damage and Structural Failure? *International Journal for Multiscale Computational Engineering*, 8(1):61–67, 2010. ISSN 1543-1649.

Z.P. Bazant and J. Planas. *Fracture and size efect in concrete and other quasibrittle materials*. CRC Press, Boca Raton, FL, 1998.

T. Belytschko, H. Chen, J.X. Xu, and G. Zi. Dynamic crack propagation based on loss of hyperbolicity and a new discontinuous enrichment. *Int. J. Numer. Methods Engrg.*, 58:1873–1905, 2003.

T. Belytschko, S. Loehnert, and J.H. Song. Multiscale aggregating discontinuities: A method for circumventing loss of material stability. *Int. J. Numer. Meth. Engng.*, 73:869–894, 2008.

Pablo J. Blanco, Pablo J. Sánchez, Eduardo A. de Souza Neto, and Raúl A. Feijóo. Variational foundations and generalized unified theory of rve-based multiscale models. *Archives of Computational Methods in Engineering*, 23(2):191–253, 2016.

ISSN 1886-1784. doi: 10.1007/s11831-014-9137-5. URL http://dx.doi.org/10.1007/s11831-014-9137-5.

H. Böhm. Short introduction to basic aspects of continuum micromechanics. *Technical Report. Institute of Lightweight Design and Structural Biomechanics (ILSB). Vienna University of Technology*, 27(206), 2013.

M. Bornert. *Morphologie microstructurale et comportement mécanique; charactérisations expérimentales, approches par bornes et estimations autocohérentes généralisées*. PhD. Thesis, Ecole Nationale des Ponts et Chaussées, 1999.

S. Boyaval. Reduced-basis approach for homogenization beyond the periodic setting. *Arxiv preprint math/0702674*, 2007.

E.W.C. Coenen, V.G. Kouznetsova, E. Bosco, and M.G.D. Geers. A multi-scale approach to bridge microscale damage and macroscale failure: a nested computational homogenization-localization framework. *Int. J. Fracture*, 178(1-2):157–178, 2012.

R. de Borst and H. Mühlhaus. Gradient-dependent plasticity: formulation and algorithmic aspects. *Int J. Numer. Meth. Engrg.*, 35(3):521–539, 1992.

A. Drago and M. J. Pindera. Micro-macromechanical analysis of heterogeneous materials: Macroscopically homogeneous vs periodic microstructures. *Composites science and technology*, 67(6):1243–1263, 2007.

J.D. Eshelby. The determination of the field of an ellipsoidal inclusion and related problems. *Proc. R. Soc. Lond A*, 241:376–396, 1957.

C. Farhat, T. Chapman, and P. Avery. Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering*, 2015.

R. Faria, J. Oliver, and M. Cervera. A strain-based plastic viscous-damage model for massive concrete structures. *International Journal of Solids and Structures*, 35(14):1533–1558, 1998. ISSN 0020-7683. doi: 10.1016/S0020-7683(97)00119-4. URL http://www.sciencedirect.com/science/article/pii/S0020768397001194.

F. Feyel and J.L. Chaboche. FE^2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. *Comput. Meth. App. Mech. Eng.*, 183:309–330, 2000.

Shankar Ganapathysubramanian and Nicholas Zabaras. Design across length scales: a reduced-order model of polycrystal plasticity for the control of microstructure-sensitive material properties. *Computer Methods in Applied Mechanics and Engineering*, 193(45-47):5017 – 5034, 2004. ISSN 0045-7825. doi: http://dx.doi.org/10.1016/j.cma.2004.04.004. URL http://www.sciencedirect.com/science/article/pii/S0045782504002567.

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

I.M. Gitman, H. Askes, and L.J. Sluys. Representative volume: Existence and size determination. *Engineering Fracture Mechanics*, 74:2518–2534, 2007.

D. Gross and T. Seelig. *Fracture mechanics: with an introduction to micromechanics*. Springer–Verlag, Berlin Heidelberg, 2nd edition, 2011.

J.A. Hernández, J. Oliver, A.E. Huespe, M.A. Caicedo, and J.C. Cante. *Computational Homogenization of Inelastic Materials using Model Order Reduction*, volume 141. International Center for Numerical Methods in Engineering (CIMNE) Monograph, 2014a.

J.A. Hernández, J. Oliver, A.E. Huespe, M.A. Caicedo, and J.C. Cante. High-performance model reduction techniques in computational multiscale homogenization. *Computer Methods in Applied Mechanics and Engineering*, 276: 149–189, 2014b.

J.A. Hernández, M.A. Caicedo, and Ferrer A. Dimensional hyper-reduction of nonlinear finite element models via empirical cubature. *Computer Methods in Applied Mechanics and Engineering*, 313:687–722, 2017. ISSN 0045–7825. doi: http://dx.doi.org/10.1016/j.cma.2016.10.022. URL `http://www.sciencedirect.com/science/article/pii/S004578251631355X`.

R. Hill. The elastic behavior of a crystalline aggregate. *Proc. Phys. Soc. London*, A65:349–354, 1952.

R. Hill. A self-consistent mechanics of composite materials. *J. Mech. Phys. Solids*, 13:213–222, 1965.

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

V.G. Kouznetsova. *Computational homogenization for the multi-scale analysis of multi-phase materials*. PhD thesis, Technische Universiteit Eindhoven, Netherlands, 2002.

T. Lesicar, Z. Tonković, and J. Sorić. Multi-Scale Modeling of Heterogeneous Materials and the Validation Challenge. *Key Engineering Materials*, 665:181–184, 2015. doi: 10.4028/www.scientific.net/AMM.70.345. URL `www.scientific.net/KEM.665.181`.

C. Linder and F. Armero. Finite elements with embedded branching. *Finite Elements in Analysis and Design*, 45(4):280–293, 2009. ISSN 0168-874X. doi: http://dx.doi.org/10.1016/j.finel.2008.10.012. URL `http://www.sciencedirect.com/science/article/pii/S0168874X08001595`. The Twentieth Annual Robert J. Melosh Competition.

O. Lloberas-Valls, D.J. Rixen, A. Simone, and L.J. Sluys. Multiscale domain decomposition analysis of quasi-brittle heterogeneous materials. *Int. J. Num. Meth. Eng.*, 89(11):1337–1366, 2012.

O. Lloberas-Valls, A.E. Huespe, J. Oliver, and I.F. Dias. Strain injection techniques in dynamic fracture modeling. *Computer Methods in Applied Mechanics and Engineering*, 308:499–534, 2016. ISSN 0045–7825. doi: http://dx.doi.org/10.1016/j.cma.2016.05.023. URL http://www.sciencedirect.com/science/article/pii/S0045782516304248.

J. Mandel. *Plasticité Classique et Viscoplasticité*. Springer–Verlag, Udine, Italy, 1971.

K. Markov. Elementary micromechanics of heterogeneous media. *Heterogeneous Media: Micromechanics Modeling Methods and Simulations*, pages 1–162, 2000.

K. Matous, M.G. Kulkarni, and P.H. Geubelle. Multiscale cohesive failure modeling of heterogeneous adhesives. *Journal of the Mechanics and Physics of Solids*, 56:1511–1533, 2008.

J.C. Michel and P. Suquet. Nonuniform transformation field analysis. *International Journal of Solids and Structures*, 40(25):6937–6955, 2003. ISSN 0020-7683. doi: 10.1016/S0020-7683(03)00346-9. URL http://www.sciencedirect.com/science/article/pii/S0020768303003469. Special issue in Honor of George J. Dvorak.

J.C. Michel and P. Suquet. Computational analysis of nonlinear composite structures using the nonuniform transformation field analysis. *Computer Methods in Applied Mechanics and Engineering*, 193(48–51):5477–5502, 2004. ISSN 0045-7825. doi: 10.1016/j.cma.2003.12.071. URL http://www.sciencedirect.com/science/article/pii/S004578250400283X. Advances in Computational Plasticity.

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

G. W. Milton. *The Theory of Composites*. Cambridge University Press, Cambridge, UK, 2002.

J.F. Molinari, G. Gazonas, R. Raghupathy, A. Rusinek, and F. Zhou. The cohesive element approach to dynamic fragmentation: the question of energy convergence. *Int. J. Num. Meth. Eng.*, 69(3):484–503, 2007.

E. Monteiro, J. Yvonnet, and Q.C. He. Computational homogenization for nonlinear conduction in heterogeneous materials using model reduction. *Computational Materials Science*, 42(4):704 – 712, 2008. ISSN 0927-0256. doi: http://dx.doi.org/10.1016/j.commatsci.2007.11.001. URL http://www.sciencedirect.com/science/article/pii/S0927025607003242.

T. Mori and K. Tanaka. Average stress in the matrix and average energy of materials with misfitting inclusions. *Acta Metall.*, 21:571–574, 1973.

H. Muhlhaus and I. Vardoulakis. The thickness of shear bands in granulars materials. *Geotechnic*, 37:271–283, 1987.

S. Nemat-Nasser and M. Hori. *Micromechanics: overall properties of heterogeneous materials*. Elsevier, 1999.

N.C. Nguyen. A multiscale reduced-basis method for parametrized elliptic partial differential equations with multiple scales. *Journal of Computational Physics*, 227(23):9807 – 9822, 2008. ISSN 0021-9991. doi: http://dx.doi.org/10.1016/j.jcp.2008.07.025. URL http://www.sciencedirect.com/science/article/pii/S0021999108004087.

V.P. Nguyen, O. Lloberas-Valls, M. Stroeven, and L.J. Sluys. On the existence of representative volumes for softening quasi-brittle materials - A failure zone averaging scheme. *Comput. Meth. App. Mech. Eng.*, 199:3028–3038, 2010a.

V.P. Nguyen, O.Lloberas-Valls, and L.J. Sluys M. Stroeven. Homogenization-based multiscale crack modelling: from micro diffusive damage to macro cracks. *Comput. Meth. App. Mech. Eng.*, 200:1220–1236, 2010b.

J. T. Oden, T. Fish J. Belytschko, T. J. R. Hughes, C. Johnson, L. A Keyes, L Petzold, L. Srolovitz, and S. Yip. Simulation-based engineering science. Report of the nsf sbes panel to the nsf engineering advisory committee, National Science Foundation (USA), 2006.

J. Oliver. A consistent characteristic length for smeared cracking models. *International Journal for Numerical Methods in Engineering*, 28(2):461–474, 1989. ISSN 1097-0207. doi: 10.1002/nme.1620280214. URL http://dx.doi.org/10.1002/nme.1620280214.

J. Oliver. Continuum modelling of strong discontinuities in solids mechanics using damage models. *Comput. Mech.*, 17(1-2):49–61, 1995b.

J. Oliver. Modelling strong discontinuities in solids mechanics via strain softening constitutive equations. Part 1: Fundamentals. *Int. j. numer. methods eng.*, 39(21):3575–3600, 1996a.

J. Oliver. On the discrete constitutive models induced by strong discontinuity kinematics and continuum constitutive equations. *Int. J. Solids Struct.*, 37:7207–7229, 2000.

J. Oliver and A. E. Huespe. Continuum approach to material failure in strong discontinuity settings. *Comp. Meth. Appl. Mech. in Engrg.*, 193:3195–3220, 2004a.

J. Oliver and A.E. Huespe. Theoretical and computational issues in modelling material failure in strong discontinuity scenarios. *Comput. Meth. App. Mech. Eng.*, 193:2987–3014, 2004b.

J. Oliver, A. E. Huespe, M. D. G. Pulido, and E. Chaves. From continuum mechanics to fracture mechanics: the strong discontinuity approach. *Engineering Fracture Mechanics*, 69:113–136, 2002.

J. Oliver, A.E. Huespe, S. Blanco, and D.L. Linero. Stability and robustness issues in numerical modeling of material failure with the strong discontinuity approach. *Comput. Meth. App. Mech. Eng.*, 195(52):7093–7114, 2005.

J. Oliver, I.F. Dias, and A.E. Huespe. Crack-path field and strain-injection techniques in computational modeling of propagating material failure. *Computer Methods in Applied Mechanics and Engineering*, 274:289–348, 2014. ISSN 0045-7825. doi: 10.1016/j.cma.2014.01.008. URL http://www.sciencedirect.com/science/article/pii/S0045782514000139.

J. Oliver, M. Caicedo, E. Roubin, A.E. Huespe, and J.A. HernÃ¡ndez. Continuum approach to computational multiscale modeling of propagating fracture. *Computer Methods in Applied Mechanics and Engineering*, 294:384–427, 2015. ISSN 0045-7825. doi: 10.1016/j.cma.2015.05.012. URL http://www.sciencedirect.com/science/article/pii/S0045782515001851.

J. Oliver, M. Caicedo, A.E. Huespe, J.A. Hernández, and E. Roubin. Reduced order modeling strategies for computational multiscale fracture. *Computer Methods in Applied Mechanics and Engineering*, 313:560–595, 2017a. ISSN 0045-7825. doi: 10.1016/j.cma.2016.09.039. URL http://www.sciencedirect.com/science/article/pii/S0045782516303322.

J. Oliver, M. Caicedo, A.E. Huespe, J.A. HernÃ¡ndez, and E. Roubin. Reduced order modeling strategies for computational multiscale fracture. *Computer Methods in Applied Mechanics and Engineering*, 313:560–595, 2017b. ISSN 0045-7825. doi: 10.1016/j.cma.2016.09.039. URL http://www.sciencedirect.com/science/article/pii/S0045782516303322.

C. Oskay and J. Fish. Eigendeformation-based reduced order homogenization for failure analysis of heterogeneous materials. *Comput. Meth. App. Mech. Eng.*, 196 (7):1216–1243, 2007.

F. Otero, X. Martínez, S. Oller, and O. Salomón. An efficient multi-scale method for non-linear analysis of composite structures. *Composite Structures*, 131:707–719, 2015. ISSN 0263-8223. doi: 10.1016/j.compstruct.2015.06.006. URL http://www.sciencedirect.com/science/article/pii/S0263822315004699.

A. Pandolfi, P. Krysl, and M. Ortiz. Finite element simulation of ring expansion and fragmentation: the capturing of length and time scales through cohesive models of fracture. *Int. J. Fracture*, 95(1-4):279–297, 1999.

R. Peerlings, R. de Borst, W. Brekelmans, and J. de Vree. Gradient Enhanced Damage for Quasi-Brittle Materials. *Int J. Numer. Meth. Engrg.*, 39:3391–3403, 1996.

R. Peerlings, M. Geers, R. de Borst, and W. Brekelmans. A critical comparison of nonlocal and gradient-enhanced softening continua. *Int J. Solids Struct.*, 38: 7723–7746, 2001.

R. Peerlings, R. de Borst, W. Brekelmans, and M. Geers. Localization issues in local and nonlocal continuum approaches to fracture. *European Journal of Mechanics A/Solids*, 21:175–189, 2002.

G. Pijaudier-Cabot and Z. Bazant. Nonlocal damage theory. *ASCE J. Engrg. Mech.*, 113(10):1512–1533, 1987.

P. Ponte Castañeda and P. Suquet. Nonlinear composites. *Advances in Applied Mechanics*, 34:171–302, 1998.

B. Prabel, A. Combescure, A. Gravouil, and S. Marie. Level set x-fem non-matching meshes: application to dynamic crack propagation in elastic-plastic media. *International Journal for Numerical Methods in Engineering*, 69(8):1553–1569, 2007. ISSN 1097-0207. doi: 10.1002/nme.1819. URL `http://dx.doi.org/10.1002/nme.1819`.

B. D. Reddy and J. Simo. Stability and convergence of a class of enhanced strain methods. *SIAM, Jornal on Numerical Analysis*, 32:1705–1728, 1995.

A. Reuss. Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbe–dingung für Einkristalle. *ZAMM*, 9:49–58, 1929.

E.A. Rodrigues, O.L. Manzoli, L.A.G. Bitencourt Jr, and T. Bittencourt. 2D mesoscale model for concrete based on the use of interface element with a high aspect ratio. *International Journal of Solids and Structures*, 94:112–124, 2016.

J. Rots. *Computational modelling of concrete fracture*. PhD thesis, Delft University, Netherlands, 1988.

D. Ryckelynck. Hyper-reduction of mechanical models involving internal variables. *International Journal for Numerical Methods in Engineering*, 77(1):75–89, 2009.

J. Simo and T. Hughes. *Computational inelasticity*. Springer-Verlag, 1998.

J. Simo and T. J. R. Hughes. On the variational foundations of assumed strain methods. *Journal of Applied Mechanics, ASME*, 53:51–54, 1986.

J. Simo and M. Riffai. A class of mixed assumed strain methods and the method of incompatible modes. *Int. J. Numer. Methods eng.*, 29:1595–1638, 1990.

J. Simo, J. Oliver, and F. Armero. An analysis of strong discontinuities induced by strain-softening in rate-independent inelastic solids. *Comput. Mech.*, 12:277–296, 1993.

J.H. Song and T. Belytschko. Multiscale aggregating discontinuities method for micro-macro failure of composites. *Composites*, Part B 40:417–426, 2009.

P. Steinmmann and K. Willam. Localization within the framework of micropolar elasto-plasticity. In V. Mannl et al., editor, *Advances in continuum mechanics*, pages 296–313, Berlin, 1991. Springer Verlag.

Kenjiro Terada, Muneo Hori, Takashi Kyoya, and Noboru Kikuchi. Simulation of the multi-scale convergence in computational homogenization approaches. *International Journal of Solids and Structures*, 37(16):2285–2311, 2000. ISSN 0020-7683. doi: 10.1016/S0020-7683(98)00341-2. URL http://www.sciencedirect.com/science/article/pii/S0020768398003412.

S. Toro, P.J. Sánchez, A.E. Huespe, S.M. Giusti, P.J. Blanco, and R.A. Feijóo. A two-scale failure model for heterogeneous materials: numerical implementation based on the finite element method. *Int. J. Num. Meth. Eng.*, 97(5):313–351, 2014.

S. Toro, P.J. Sánchez, P.J. Blanco, E.A. de Souza Neto, A.E. Huespe, and R.A. Feijóo. Multiscale formulation for material failure accounting for cohesive cracks at the macro and micro scales. *Int. J. of Plasticity*, 76:75–110, 2016.

S. Torquato. *Random Heterogeneous Media*. Springer–Verlag, New York, NY, 2002.

V. Tvergaard and A. Needleman. Efects of nonlocal damage in porous plastic solids. *Int J. Solids Struct.*, 32(8/9):1063–1077, 1995.

J. F. Unger. An FE2-X1 approach for multiscale localization phenomena. *Journal of the Mechanics and Physics of Solids*, 61(4):928–948, 2013. ISSN 0022-5096. doi: 10.1016/j.jmps.2012.12.010. URL http://www.sciencedirect.com/science/article/pii/S0022509613000069.

W. Voigt. Über die Beziehung zwischen den beiden Elasticitäts-Constanten isotroper Körper. *Ann.Phys.*, 38:573–587, 1889.

J. Yvonnet and Q-C. He. The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains. *Journal of Computational Physics*, 223(1):341–368, 2007.

S. Zhang and C. Oskay. Reduced order variational multiscale enrichment method for elasto-viscoplastic problems. *Computer Methods in Applied Mechanics and Engineering*, 300:199–224, 2016.

O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method*. Butterworth-Heinemann, Oxford, UK, 2000.

Part II

APPENDIX

List of participations in national and international conferences.

A

CONTRIBUTIONS TO CONFERENCES AND WORKSHOP PROCEEDINGS

During the development of the Thesis, the partial advances and results have been presented in several national and international specialized conferences and whorshops. The list of attended conferences are:

- **Caicedo M., Oliver J., Huespe A. E., Hernández J., Ferdinand C.** Computational Multiscale Modelling of materials using Reduced-order modeling techniques, Whorkshop CIMNE-Lulea, Barcelona, Spain. May 2011.
- **Oliver J., Hernández J.A., Huespe A., Caicedo M.** High performance Model-Order-Reduction methods in computational multi-scale simulations of non-linear solids, European Congress on Computational Methods in Applied Sciences and Engineering (**ECCOMAS 2012**), Vienna, Austria, September 2012.
- **Oliver J., Hernández J.A., Huespe A., Caicedo M.** High performance Model-Order-Reduction methods in computational multi-scale simulations of non-linear solids, 1st Spain - Japan Whorkshop on Computational Mechanics, Barcelona, Spain, September 2012. **Link to Publisher**
- **Oliver J., Huespe A., Caicedo M., Hernández J.A.** Two-scale modeling of material failure based on the Continuum Strong Discontinuity Approach, The Third International Conference on Computational Modeling of Fracture and Failure of Materials and Structures (**CFRAC 2013**), Prague, Czech Republic, 5–7 June 2013.
- **Hernández J.A., Oliver J., Huespe A., Caicedo M., Cante J.C** Application of High-Performance Reduced-Order Modeling (HP-ROM) to two-scale homogenization problems, Congress of Numerical Methods in Engineering 2013, SEMNI, Bilbao, Spain, 25–28 June 2013.
- **Oliver J., Huespe A., Caicedo M., Hernández J.A.** Continuum Approach to Computational Multi-scale Modeling of Material Failure, XII International Conference on Computational Plasticity: Fundamentals and Applications (**COMPLAS 2014**), Barcelona, Spain, 3–5 September 2013.

- **Hernández J.A., Oliver J., Huespe A., Caicedo M., Cante J.C** On the application of high-performance model reduction techniques to homogenization of heterogeneous materials, XII International Conference on Computational Plasticity: Fundamentals and Applications (**COMPLAS 2014**), Barcelona, Spain, 3–5 September 2013.
- **Oliver J., Huespe A., Caicedo M., Roubin E.,Hernández J.A.** Multi-scale analysis of material failure in cement/aggregate composite structures, Computational modelling of Concrete and Concrete Structures (**EURO–C 2014**), St. Anton am Arlberg, Austria, 24–27 March 2014.
- **Hernández J.A., Oliver J., Huespe A., Caicedo M.** Computational homogenization using high-performance, reduced-order modeling, XI World Congress on Computational Mecnahics (**XI WCCM 2014**), Barcelona, Spain, 20–25 July 2014.
- **Caicedo M, Oliver J., Huespe A., Roubin E., Hernández J.A.** Continuum multi-scale (FE^2) Modeling of Material Failure, XI World Congress on Computational Mecnahics (**XI WCCM 2014**), Barcelona, Spain, 20–25 July 2014. **Link to Publisher**
- **Oliver J., Caicedo M., Roubin E., Huespe A.** Continuum approach to computational multi-scale modeling of fracture, 13th International Conference on Fracture and Damage Mechanics (**FDM-2014**), Sao Miguel Island, Azores, Portugal, September 23–25 2014.
- **Oliver J., Caicedo M., Roubin E., Huespe A.** Continuum Multi-scale modeling of propagating fracture in quasi-brittle materials, European Mechanics Congress (**EUROMECH-2015**), Eindhoven, The Netherlands, February 2015.
- **Oliver J., Huespe A. and Caicedo M.** Multiscale modelling of propagating fracture in quasi brittle materials: a continuum approach, IV International Conference on Computational Modeling of Fracture and Failure of Materials and Structures (**CFRAC 2015**), École Normale Superiéure de Cachan, France, 3–5 June 2015.
- **Caicedo M., Roubin E., Huespe A. and Oliver J.** Continuum multi-scale modeling of fracture in cementitious-like materials, Congress on Numerical Methods in Engineering (**CMN2015**), Lisbon, Portugal, Juny 29 - July 2, 2015.
- **Roubin E., Caicedo M., Hernández J., Huespe A. and Oliver J.** On POD based reduction-order modeling in multi-scale material failure simulation, Congress on Numerical Methods in Engineering (**CMN2015**), Lisbon, Portugal, Juny 29 - July 2, 2015.
- **Caicedo M., Oliver J., Roubin E., Huespe A. and Hernández J.** Continuum multi-scale (FE^2) modeling of material failure in concrete-like materials, XIII

International Conference on Computational Plasticity: Fundamentals and Applications (**COMPLAS 2015**), Barcelona, Spain, 1–3 September 2015.

- **Oliver J., Huespe A. and Caicedo M.** A continuum approach for multi-scale propagating material fracture modeling, VII European Congress on Computational Methods in Applied Sciences and Engineering (**ECCOMAS 2016**), Crete, Greece, June 5–10, 2016.

- **Caicedo M., Oliver J., Huespe A. and O. Lloberas-Valls.** Model Order Reduction in Computational Multiscale Fracture Mechanics, 15th International Conference on Fracture and Damage Mechanics (**FDM-2016**), Alicante, Spain, September 14–16, 2016.

- **Oliver J., Huespe A. and Caicedo M.** Hyper-reduced order modeling techniques in computational multiscale fracture, Colloquium 584 Multi-uncertainty and multi-scale methods and related applications (**EuroMesh 2016**), Porto, Portugal, September 14–16, 2016.

- **Oliver J., Huespe A. and Caicedo M.** High-performance model order reduction in computational multiscale fracture, XXII Congreso sobre Métodos Numéricos y sus aplicaciones (**ENIEF 2016**), Córdoba, Argentina, Noviembre 8–11, 2016.

- **Caicedo M., Toro S., Mroginski J. L., Oliver J. and Huespe A. E.** Model Order Reduction applied to Geometric Non-Linear Multiscale Modeling, XXII Congreso sobre Métodos Numéricos y sus aplicaciones (**ENIEF 2016**), Córdoba, Argentina, Noviembre 8–11, 2016.

B

PUBLISHED ARTICLES

B.1 Paper #1

Title: High-performance model reduction techniques in computational multiscale homogenization.
Authors:

- **J. A. Hernández**: Assistant Professor of Structural Engineering and Strength of Materials at the School of Industrial and Aeronautic Engineering of Terrassa, of the Technical University of Catalonia. Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **J. Oliver**: Professor of Continuum Mechanics and Structural analysis at the Escola Tecnica Superior d'Enginyers de Camins, Canals i Ports (Civil Engineering School) of the Universitat Politècnica de Catalunya (Technical University of Catalonia BarcelonaTech). Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **A. E. Huespe**: Professor of Mechanics at the Faculty of Chemical Engineering, Dept. of Materials, National University of Litoral, Santa Fe, Argentina. Independent researcher of Conicet at CIMEC (Centro de Investigaciones en Mecánica Computacional), National University of Litoral (UNL).
- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **J. C. Cante**: Associate Professor of Computational Engineering at the Escola Tècnica Superior d'Enginyeries Industrial i Aeronàutica de Terrassa – Universitat Politècnica de Catalunya (Technical University of Catalonia, BarcelonaTech). Associate researcher at the International Center for Numerical Methods in Engineering (CIMNE).

Journal of Computer Methods in Applied Mechanics and Engineering
Editors: Thomas J.R. Hughes, J. Tinsley Oden, Manolis Papadrakakis
ISSN: 0045-7825
Elsevier Editors
http://dx.doi.org/10.1016/j.cma.2014.03.011
Link to Publisher

High-performance model reduction techniques in computational multiscale homogenization

J.A. Hernández[a,c,*], J. Oliver[a,b], A.E. Huespe[b,d], M.A. Caicedo[a], J.C. Cante[a,c]

[a] *Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), Technical University of Catalonia, Edificio C1, Campus Norte, Jordi Girona 1-3, Barcelona 08034, Spain*

[b] *E.T.S. d'Enginyers de Camins, Canals i Ports, Technical University of Catalonia, Edificio C1, Campus Norte, Jordi Girona 1-3, Barcelona 08034, Spain*

[c] *E.T.S. d'Enginyeries Industrial i Aeronàutica de Terrassa, Technical University of Catalonia, C/ Colom, 11, Terrassa 08222, Spain*

[d] *CIMEC, CONICET, Güemes 3450, 3000 Santa Fe, Argentina*

Abstract

A novel model-order reduction technique for the solution of the fine-scale equilibrium problem appearing in computational homogenization is presented. The reduced set of empirical shape functions is obtained using a partitioned version —that accounts for the elastic/inelastic character of the solution— of the Proper Orthogonal Decomposition (POD). On the other hand, it is shown that the standard approach of replacing the nonaffine term by an interpolant constructed using only POD modes leads to ill-posed formulations. We demonstrate that this ill-posedness can be avoided by enriching the approximation space with the span of the gradient of the empirical shape functions. Furthermore, interpolation points are chosen guided, not only by accuracy requirements, but also by stability considerations. The approach is assessed in the homogenization of a highly complex porous metal material. Computed results show that computational complexity is independent of the size and geometrical complexity of the representative volume element. The speedup factor is over three orders of magnitude —as compared with finite element analysis— whereas the maximum error in stresses is less than 10%.

*Corresponding author

Email address: jhortega@cimne.upc.edu (J.A. Hernández)

Keywords: Multiscale, homogenization, model reduction,
High-Performance reduced-order model, hyperreduction, POD

1. Introduction

1.1. Motivation and goal

The major challenge in the macro-scale continuum description of heterogeneous materials such as composites and polycrystalline metals (that exhibit a clear *heterogeneous* composition at the the micro-, or meso-, scale, but that can be regarded, for practical purposes, as *homogeneous* at the the macro-scale) lies in the determination of a constitutive connection, between macro-stresses and macro-strains, that accurately reflects the properties and geometrical arrangement of the distinct phases at the finer scale. It is well-known [34] that, under the hypotheses of either *periodicity or statistical homogeneity*, on the one hand; and *scale separation*, on the other hand, this constitutive link can be systematically established by solving, for *each* point at the coarse scale, a boundary value problem (BVP) on a certain representative microscopic subdomain. In a strain-driven formulation of this BVP, the macro-strain at a given point acts as "loading parameter", in the form of appropriate essential boundary conditions, whereas the associated macro-stress is obtained through volume averaging —i.e., *homogenization*— of the corresponding micro-stress field.

Methods dealing with the solution of this BVP range from purely analytical approaches to *direct computational methods*, such as the two-level, Finite Element (FE^2) method [29]. Analytical approaches are computationally inexpensive, but only valid for certain types of geometrically and constitutively simple micro-structures. By contrast, direct computational methods have no other limitation in scope than the imposed by the aforementioned hypotheses of statistical homogeneity and scale separation —in these methods, the microscopic BVP at each coarse-scale point is attacked using no other approximation than the spatial discretization of the pertinent solution strategy, thus, circumventing the need for introducing *ad-hoc*, simplifying assumptions regarding the topological arrangement of the micro-phases and/or their collective constitutive behavior. Needless to say, the versatility of direct computational homogenization comes at a significant price: its enormous computational cost.

Between these two extremes (purely analytical and direct computational methods), there are homogenization strategies that can be termed *semi-analytical*, since they combine analytical results with numerical computations. Such is the case of the *Transformation Field Analysis* (TFA) [25] and variants thereof [49, 50, 56, 30], which are based on the *pre-computation* of certain characteristic operators (strain localization and influence tensors) using the information obtained from solving a carefully chosen battery of fine-scale BPVs. Although these methods have notably widen the scope of classical analytical approaches —while maintaining their low computational cost—, they are still predicated, to a lesser or greater extent, on ad-hoc assumptions connected with the constitutive description of the involved phases. Consideration of new materials with unstudied compositions using semi-analytical approaches, thus, requires additional research efforts by specialists in the field and eventual modifications of the corresponding mathematical and numerical formulations —in contrast to direct computational homogenization approaches, such as the FE^2 method, in which the formulation is "material-independent", and hence more versatile.

The current state of affairs in the field of two-scale homogenization seems to call, thus, for a unified homogenization approach that combines somewhat the advantages of direct computational homogenization and analytical and semi-analytical techniques. It would be desirable to have a homogenization method with a computational cost virtually *independent of the geometric complexity* of the considered representative volume, as in analytical and semi-analytical techniques. At the same time, it would be also interesting to arrive at a method whose mathematical formulation dispenses with *ad-hoc*, simplifying assumptions related with the composition of the heterogeneous material; i.e, one enjoying the versatility, unrestricted applicability and "user-friendliness" —insofar as it would totally relieve the modeler from the often exceedingly difficult task of visualizing such assumptions — of direct computational homogenization methods. The goal of the present paper is to show that these desirable attributes can be achieved, for arbitrarily complex heterogeneous materials *well into the inelastic range*, by using the so-called [47] *Reduced-Basis* (RB) *approximation* in the solution of the fine-scale BVPs.

1.2. The reduced-basis method

Generally speaking, the reduced-basis approximation is a class of *Galerkin* approximation procedure that employs, as opposed to the FE method, but

similarly to classical Rayleigh-Ritz solution techniques [19], *globally supported basis functions*. The main difference with respect to classical Rayleigh-Ritz schemes is that these basis functions *or modes* are not constructed from either polynomials or transcendental functions (sines, cosines ...), but rather are determined from a larger set of *previously* computed —using *the finite element* (FE) method or other classical solution techniques— solutions of the BVP at appropriately selected values of the input of interest. These functions are commonly termed *empirical basis functions* [41], the qualifier empirical meaning "derived from computational experiments".

1.2.1. Dimensionality reduction

As noted earlier, the input of interest or "loading" parameter in the fine-scale problem is the macro-scale strain tensor. Accordingly, the starting point for constructing the basis functions consists in solving, using the FE method, a battery of BVPs for various, representative *macro-strain histories*. The outcome of these FE calculations is a data set comprising an ensemble of hundred or even thousand (depending on the number of time steps into which the strain histories are discretized) displacement field solutions (also called *snapshots*). Were all these snapshots barely correlated with each other, the dimension of the manifold spanned by them would prove overly high, rendering the entire approach impractical —it would no longer qualify as a truly *reduced* basis method. Fortunately, as we show in the present paper, in general, most of these snapshots do display strong linear correlations between each other —i.e., they have redundant information—, and, in addition, contain deformation modes that are *irrelevant* to the quality of coarse-scale predictions. All that is required to obtain a much lower dimensional representation of the solution data set, and therewith the desired reduced basis, is an automatic means to identify and remove this redundant and irrelevant information, while preserving, as much as possible, its *essential* features. The problem of removing unnecessary complexity from huge data sets so as to uncover *dominant patterns* is the central concern of disciplines such as digital image compression [60] and patter recognition [8], to name but a few, and thereby many efficient *dimensionality reduction* (or *data compression*, in more common parlance) algorithms already exist to deal with it. In the present work, we employ one of the simplest and most popular of these di-

mensionality reduction algorithms: the Proper Orthogonal Decomposition[1] (POD).

It may be inferred from the above that the proposed homogenization method, like analytical and semi-analytical strategies, does introduce *simplifications* in solving the fine-scale BVP. However, as opposed to analytical, and to a lesser extent, semi-analytical procedures, these simplifications are not introduced by the modeler, but rather are automatically carried out by the abovementioned dimensionality reduction methods (in an offline stage, prior to the overall multiscale analysis). In other words, in the proposed method, the *task of discerning what is essential and what is not*[2] *is entirely delegated to the computer* itself, and hence, its success does not depend upon the depth of insight, experience, and knowledge base of the modeler —only some discretion is to be exercised in choosing appropriate strain paths for the offline FE analyses. This feature naturally confers the advantages of versatility and "user-friendliness" enjoyed by direct computational methods.

1.2.2. Numerical integration

Once the global shape functions have been determined, the next step is to introduce an *efficient* method for *numerically evaluating the integrals* appearing in the weak form of the cell BVP. Of course one can simply use the same Gauss quadrature formulae and the same sampling points (a total number of $n_g = \mathcal{O}(n)$, n being the number of mesh nodes) as the underlying finite element model. But this would be akin to integrating, say, a third-order polynomial function using thousand of sampling points—a profligate waste of computational resources. Since displacement solutions for the cell BVP are constrained to lie in a reduced-order space of dimension $n_u << n$, it is reasonable to expect that the corresponding stresses, internal forces and Jacobians will also reside in reduced-order spaces of dimensions of order $\mathcal{O}(n_u)$, and consequently, *only* $p = \mathcal{O}(n_u) << n_g$ *sampling points* would suffice in principle to accurately evaluate the corresponding integrals. The challenging questions that have to be confronted are where to locate these p sampling points and, loosely speaking, how to determine their associated

[1] By constraining the cell to deform only into the deformation modes determined by the POD, one automatically obtains a genuine *reduced-order model* (ROM) of the cell.

[2] Discerning what is essential and what is not is, according to M.Ashby [4], the key to any successful computational model (i.e., one that strikes the right balance between accuracy and simplicity)

weighting functions so that maximum accuracy in the integration is attained.

Approaches found in the model reduction literature that, directly or indirectly, deal with these fundamental questions can be broadly classified either as *interpolatory* methods [7, 33, 54, 18, 5] or *Gauss-type quadrature* strategies [3, 39]. In both types of approaches, the integrand or part of the integrand is approximated by a linear combination of a reduced set of empirical modes. In interpolatory approaches, the coefficients in this approximation are obtained by *interpolation* at a set of pre-selected sampling points; the criterion for choosing the location of such points is the *minimization of the interpolation error* over the finite element snapshots. In Gauss-type quadrature procedures, on the other hand, the selection of sampling points and the calculation of the accompanying weighting factors are *simultaneously* carried out, guided by a criterion of *minimum integration* error over the snapshots.

In the BVP under consideration, the *output of interest* is the volume average of the stresses over the cell domain and, therefore, accuracy is required not only in the integration of the equilibrium equation, but also on the approximation of the stresses themselves. This is the reason why attention is focused here on interpolatory integration strategies, the variable subject to spatial interpolation being precisely the stresses.

1.3. Originality of this work

The idea of exploiting the synergistic combination of multiscale modeling and reduced basis approximation is admittedly not new. In the specific context of two-scale homogenization, it has been recently explored by Boyaval [10], Yvonnet et al. [62], and Monteiro et al. [51]. Traces of this idea can also be found in articles dealing with more general hierarchical multiscale techniques —that do not presuppose either scale separation or periodicity/statistical homogeneity, or both—, namely, in the *multiscale finite element method* [53, 26, 27], in the *heterogeneous multiscale method* [2, 1], and in multiscale approaches based on the Proper Generalized Decomposition (PGD)[21]. However, it should be noted that none of the above cited papers confronts the previously described, crucial question of how to *efficiently integrate* the resulting reduced-order equations, simply because, in most of them [10, 53, 26, 27, 2, 1], integration is not an issue — the fine-scale BVPs addressed in these works bear an *affine* relation with the corresponding coarse-scale, input parameter, as in linear elasticity, and, consequently, all integrals can be *pre-computed*, i.e., evaluated *offline*, with no impact in the online computational cost. Thus, the development of reduced-order models

endowed with efficient, mesh-size independent integration schemes —able to handle any material composition— is a research area that, to the best of the authors' knowledge, still remains uncharted.

1.3.1. Main original contributions

The theory underlying reduced-order models (ROMs) that incorporate *efficient* interpolatory integration schemes is still at its embryonic stage of development —the first general proposal for parametrized BVPs dates back to 2004 [7]— and many fundamental issues remain to be addressed. Foremost among these is the crucial question of *well-posedness* of the resulting system of algebraic equations: does the replacement of the integrand, or *non-affine* term in the integrand, by a reduced-order interpolant always lead to a well-posed, discrete problem ? Examination of the reduced basis literature indicates that apparently no researcher has so far been confronted with ill-posed reduced-order equations, a fact that might certainly promote the view that uniqueness of solution can be taken for granted whenever the full-order model is well-posed. Unfortunately, this is not always so: we demonstrate in this paper that the choice of the reduced-order space in which the interpolant of the integrand resides has a profound impact on the well-posedness of the discrete problem. In particular, we show that, in the case of the fine-scale boundary-value problem, the widely adopted [33] approach of determining the basis functions for this space from (converged) FE snapshots leads invariably to *ill-posed*, discrete formulations. The *main original contribution of the present work to the field of reduced-order modeling is the development of an interpolatory integration method that safely overcomes this type of ill-posedness*. The gist of the method is to *expand* the interpolation space so that it embraces, aside from the span of the POD stress basis functions, the space generated —and herein lies the novelty— by the *gradient of the (reduced-order) shape functions*. Furthermore, it is shown that, in contrast to the situation encountered when using standard interpolatory schemes in other parametrized BVPs [33], in the BVP under consideration, the *number and particular placement of sampling points within the integration domain influence notably the spectral properties* (positive definiteness) *of the Jacobian matrix of the governing equation*, and therefore, the *convergence characteristics of the accompanying Newton-Raphson solution algorithm*. Another innovative ingredient of the present paper is a points selection algorithm that does acknowledge this peculiarity and chooses the desired sampling points guided, not only by accuracy requirements (minimization of the interpolation

error over the FE stress snapshot), but also by stability considerations.

2. RVE equilibrium problem

In this section, we present the variational statement and finite element discretization of the *fine-scale equilibrium problem*, which, recall, is the parameterized BVP we wish to efficiently solve using the reduced-basis approximation.

2.1. Preliminaries

Let $\Omega \subset \mathbb{R}^d\,(d = 2, 3)$ be a subvolume of characteristic length $l << l_M$ (l_M is the characteristic length of the macro-continuum Ω_M, see Figure 1) that is *representative* of the heterogeneous material as a whole. In micro-structures that exhibit statistical homogeneity, this domain receives the name of *Representative Volume Element* (RVE), whereas in micro-structures that display periodicity, it is commonly known as repeating unit cell (RUC), or simply *unit cell* [24]. In the sequel, the acronym RVE will be used to refer to Ω.

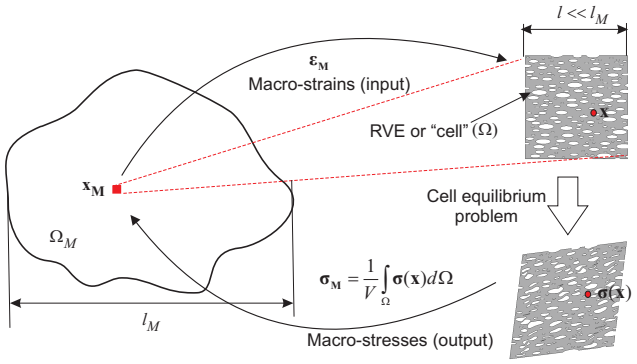


Figure 1: First-order homogenization.

In the homogenization approach adopted in this work —commonly known as *first-order* homogenization [32, 40]—, the strain field $\boldsymbol{\epsilon}(\mathbf{x})$ at any point $\mathbf{x} \in \Omega$ is assumed to be decomposed into macroscopic and fluctuating contributions; under the hypothesis of infinitesimal deformations, this decomposition can be written as:

$$\boldsymbol{\epsilon}(\mathbf{x}) = \boldsymbol{\epsilon_M} + \nabla^s \boldsymbol{u}(\mathbf{x}). \tag{1}$$

Here, $\boldsymbol{\epsilon_M}$ stands[3] for the macroscopic strain tensor (the *input parameter* in the problem) and $\nabla^s \boldsymbol{u}$ denotes the symmetric gradient of the *displacement fluctuation field* (this field is, in turn, the *basic unknown* of the problem).

Implicit in the scale separation assumption is the fact that fine-scale deformations only influence coarse-scale behavior through its volume average over the RVE. It can be shown (see, for instance, Ref. [22]) that this implies that the boundary conditions (BCs) prescribed on the RVE must be homogeneous (i.e., $\boldsymbol{A}_0\boldsymbol{u} = \mathbf{0}$ on $\partial\Omega$, $\boldsymbol{A}_0$ being a certain linear operator). The natural choice for a repeating unit cell is to employ *periodic* boundary conditions (See Refs. [9, 48] for more details on how to prescribe this type of BCs). In statistically homogeneous micro-structures, by contrast, there is a certain latitude in the choice of boundary conditions (vanishing fluctuations, uniform tractions, quasi-periodic conditions ...). In the examples shown later, vanishing boundary conditions are used ($\boldsymbol{u} = \mathbf{0}$ on $\partial\Omega$).

2.2. Variational formulation

2.2.1. Trial and test spaces

The trial space, i.e., the *set of kinematically admissible displacement fluctuation* fields, is defined formally as

$$\mathcal{V}_u = \left\{ \boldsymbol{u} \in H^1(\Omega)^d \mid \boldsymbol{A}_0\boldsymbol{u} = \mathbf{0}, \;\; on \;\; \partial\Omega \right\}, \tag{2}$$

where $H^1(\Omega)^d$ stands for the Sobolev space of functions possessing square integrable derivatives over Ω. Note that this set forms a *vector space*. Since the *test* functions $\boldsymbol{\eta}$ appearing in the variational statement shown in the following are *kinematically admissible variations* ($\boldsymbol{\eta} := \mathbf{u} - \mathbf{v}, \quad \mathbf{u}, \mathbf{v} \in \mathcal{V}_\mathbf{u}$), $\mathcal{V}_u$ having structure of vector space implies that, in the RVE equilibrium problem, the *spaces of trial and test functions coincide*.

2.2.2. Formal statement

Consider a time discretization of the interval of interest $[t_0, t_f] = \bigcup_{n=1}^{n_{stp}} [t_n, t_{n+1}]$. The current value of the microscopic stress tensor $\boldsymbol{\sigma}_{n+1}$ at each $\mathbf{x} \in \Omega$ is presumed to be entirely determined by, on the one hand, the current value of the

[3] Macroscopic variables will be identified by appending a subscript "M", while variables associated to the fine scale will be designated by bare symbols. For instance, we shall write $\boldsymbol{\epsilon_M}$ and $\boldsymbol{\epsilon}(\mathbf{x})$ to denote the macroscopic strain tensor and the fine-scale strain field, respectively.

microscopic strain tensor $\boldsymbol{\epsilon}_{n+1}(\mathbf{x}) = \boldsymbol{\epsilon}_{\boldsymbol{M}n+1} + \nabla^s \boldsymbol{u}_{n+1}(\mathbf{x})$, and, on the other hand, a set of microscopic internal variables $\boldsymbol{\xi}_{n+1}$ —that encapsulate the *history* of microscopic deformations. The relationship between these variables is established by (phenomenological) rate constitutive equations; these equations may vary from point to point within the RVE (multiphase materials). Likewise, the considered RVE may contain also voids distributed all over the domain. The (incremental) RVE equilibrium problem at time t_{n+1} can be stated as follows (see Ref. [22]): given the *initial data* $\{\boldsymbol{u}_n(\mathbf{x}), \boldsymbol{\epsilon}_{\boldsymbol{M}n}, \boldsymbol{\xi}_n(\mathbf{x})\}$ and the *prescribed* macroscopic strain tensor $\boldsymbol{\epsilon}_{\boldsymbol{M}n+1}$, find $\boldsymbol{u}_{n+1} \in \mathcal{V}_u$ such that

$$\int_\Omega \nabla^s \boldsymbol{\eta} : \boldsymbol{\sigma}_{n+1}(\boldsymbol{\epsilon}_{\boldsymbol{M}n+1} + \nabla^s \boldsymbol{u}_{n+1}, \boldsymbol{\xi}_{n+1})\, d\Omega = 0, \tag{3}$$

for all $\boldsymbol{\eta} \in \mathcal{V}_u$. The actual *output of interest* in this fine-scale BVP is not the displacement fluctuation field *per se*, but rather the macroscopic stress tensor $\boldsymbol{\sigma}_{\boldsymbol{M}}|_{n+1}$, which is defined as the volume average over the RVE of the microscopic stresses:

$$\boldsymbol{\sigma}_{\boldsymbol{M}}|_{n+1} := \frac{1}{V} \int_\Omega \boldsymbol{\sigma}_{n+1}\, d\Omega, \tag{4}$$

where V stands for the volume of the RVE. In order to keep the notation uncluttered, the superindex "n+1" will be hereafter dropped out and all quantities will be assumed to be evaluated at time t_{n+1}; only when confusion is apt to show up, the pertinent distinction will be introduced.

2.3. Finite element formulation

Let $\Omega = \bigcup_{n=1}^{n_e} \Omega^e$ be a finite element discretization of the RVE. It will be assumed that this discretization is fine enough to consider the exact and FE approximated solutions indistinguishable at the accuracy level of interest. Let $\{N_1(\mathbf{x}), N_2(\mathbf{x}) \dots N_n(\mathbf{x})\}$ (n denotes the number of nodes of the discretization) be a set of *shape functions* associated to this discretization. Now we approximate $\boldsymbol{u} \in \mathcal{V}_u$ and $\boldsymbol{\eta} \in \mathcal{V}_u$ as

$$\boldsymbol{u}(\mathbf{x}; \boldsymbol{\epsilon}_{\boldsymbol{M}}) \approx \boldsymbol{u}^{(h)}(\mathbf{x}; \boldsymbol{\epsilon}_{\boldsymbol{M}}) = \sum_{I=1}^{n} N_I(\mathbf{x}) \boldsymbol{U}_I(\boldsymbol{\epsilon}_{\boldsymbol{M}}), \tag{5}$$

$$\boldsymbol{\eta}(\mathbf{x}) \approx \boldsymbol{\eta}^{(h)}(\mathbf{x}) = \sum_{I=1}^{n} N_I(\mathbf{x}) \boldsymbol{\eta}_I, \tag{6}$$

where $\boldsymbol{U}_I \in \mathbb{R}^d$ and $\boldsymbol{\eta}_I \in \mathbb{R}^d$ $(I = 1, 2 \ldots n)$ denote the nodal values of the displacement fluctuations and test functions, respectively. Inserting these approximations in Eq.(3), and exploiting the arbitrariness of coefficients $\boldsymbol{\eta}_I$ $(I = 1, 2 \ldots n)$, one arrives at the following set of discrete equilibrium equations (repeated indices implies summation):

$$\int_\Omega \frac{\partial N_I}{\partial x_j} \boldsymbol{\sigma}_{ij}(\boldsymbol{\epsilon_M} + \nabla^s \boldsymbol{u}^{(h)}, \boldsymbol{\xi})\, d\Omega = 0 \;\; (i = 1 \ldots d;\; I = 1 \ldots n). \tag{7}$$

Introducing Voigt's notation[4], the above equation can be expressed in matrix format as:

$$\int_\Omega \boldsymbol{B}^T \boldsymbol{\sigma}(\boldsymbol{\epsilon_M} + \boldsymbol{BU}, \boldsymbol{\xi})\, d\Omega = \mathbf{0}, \tag{8}$$

As usual, numerical evaluation of the integral in Eq.(8) is carried out by Gaussian quadrature:

$$\int_\Omega \boldsymbol{B}^T \boldsymbol{\sigma}\, d\Omega \approx \sum_{g=1}^{n_g} w_g \boldsymbol{B}^T(\mathbf{x}_g) \boldsymbol{\sigma}(\mathbf{x}_g, ;) = \mathbf{0}. \tag{9}$$

Here, $n_g = \mathcal{O}(n)$ stands for the total number of Gauss points of the mesh; w_g denotes the weight associated to the $g - th$ Gauss point $\mathbf{x}_g$ (this weight includes both the *quadrature* weight itself and the corresponding Jacobian determinant.); and $\boldsymbol{B}(\mathbf{x}_g)$ and $\boldsymbol{\sigma}(\mathbf{x}_g, ;)$ stand for the B-matrix and the stress vector at Gauss point $\mathbf{x}_g$, respectively.

3. Computation of reduced basis

A basic, intuitive picture of the strategy for computing the reduced basis onto which to project the RVE equilibrium equation (3) was already given in the introductory section. In the following, we put the idea behind this strategy on a more rigorous footing. We begin by noting that, from a functional analysis standpoint, the term *model reduction* is conceptually akin

[4]Here, it is convenient to use the so-called *modified Voigt's notation* rather than the standard one. In the *modified* Voigt's notation, both stress $\boldsymbol{\sigma}$ and strain $\boldsymbol{\epsilon}$ tensors are represented as column vectors ($\{\boldsymbol{\sigma}\}$ and $\{\boldsymbol{\epsilon}\}$, respectively) in which the shear components are multiplied by $\sqrt{2}$. The advantage of this notation over the conventional, engineering Voigt's notation is the equivalence between norms; viz., $\|\boldsymbol{\sigma}\| = \sqrt{\boldsymbol{\sigma} : \boldsymbol{\sigma}} = \|\{\boldsymbol{\sigma}\}\| = \sqrt{\{\boldsymbol{\sigma}\}^T \{\boldsymbol{\sigma}\}}$. The reader is urged to consult [20] for further details on this notation.

to the more common term *model discretization*, since both connote *transitions* from higher-dimensional to lower-dimensional solution spaces. Whereas *model discretization* is used to refer to the (classical) passage from the infinite dimensional space $\mathcal{V}_u$ to the finite element subspace $\mathcal{V}_u^h \subset \mathcal{V}_u$, *model reduction* denotes a transition from this finite dimensional space $\mathcal{V}_u^h$ to a significantly smaller manifold $\mathcal{V}_u^* \subset \mathcal{V}_u^h$ —the *reduced-order* space. This latter transition is not carried out directly, but in two sequential steps, namely, *sampling of the input parameter space* and *dimensionality reduction*.

3.1. Sampling of the input parameter space

In constructing the finite element space of *kinematically* admissible functions $\mathcal{V}_u^h$, the only restrictions placed on the motion of the mesh nodes are those imposed at the boundaries. The finite element solution space, thus, does not presuppose any constraint on the motion of the *interior* nodes of the mesh.

However, in actuality, interior nodes cannot *fluctuate* freely, independently from each other, but they rather move according to *deformational patterns* dictated by the *constitutive laws* that govern the mechanical behavior of the distinct phases in the RVE[5]. This means that the solution of the finite element equilibrium equation (3) for given values of the macro-strain tensor $\boldsymbol{\epsilon_M}$ actually lives in a smaller subspace $\mathcal{V}_u^\epsilon \subset \mathcal{V}_u^h$ (in the parlance of model reduction [47, 57], $\mathcal{V}_u^\epsilon$ is the manifold induced by the parametric dependence of the BVP on the input variables).

Yet, in general, this subspace cannot be precisely determined; one has to be content to construct an *approximation* of it as the span of the displacement fluctuation solutions obtained for a judiciously chosen set of n_{hst} input strain histories $\{{}^t\boldsymbol{\epsilon_M}^1, {}^t\boldsymbol{\epsilon_M}^2, \dots {}^t\boldsymbol{\epsilon_M}^{n_{hst}}\}$. Suppose, for simplicity, that each of these strain histories is discretized into equal number of steps n_{stp}, and let

$$\boldsymbol{u}^k(\mathbf{x}) = \boldsymbol{u}(\mathbf{x}; {}^t\boldsymbol{\epsilon_M}_j{}^i), \quad k = (i-1)n_{hst} + j \tag{10}$$

denote the displacement fluctuation solution at the $j-th$ time step of the $i-th$ strain history ($i = 1, 2 \dots n_{hst}$, $j = 1, 2 \dots n_{stp}$). The approximating space for $\mathcal{V}_u^\epsilon$, henceforth called the *snapshots space*, is then defined as:

[5] As noted by Lubliner [45], constitutive laws can be regarded as *internal restrictions* on the kinds of deformation a body can suffer

$$\mathcal{V}_u^{snp} = \text{span}\left\{\boldsymbol{u}^1(\mathbf{x}), \boldsymbol{u}^2(\mathbf{x}), \dots \boldsymbol{u}^{n_{snp}}(\mathbf{x})\right\} \subseteq \mathcal{V}_u^{\epsilon}, \tag{11}$$

$n_{snp} = n_{stp} n_{hst}$ being the total number of snapshots. The matrix containing, in columns, the nodal values of these displacement fluctuations solutions:

$$\boldsymbol{X}_u = \begin{bmatrix} \boldsymbol{U}^1 & \boldsymbol{U}^2 & \cdots & \boldsymbol{U}^{n_{snp}} \end{bmatrix} \in \mathbb{R}^{n \cdot d \times n_{snp}} \tag{12}$$

will correspondingly be termed the (displacement fluctuations) *snapshot matrix*.

3.2. Dimensionality reduction

The next and definitive step in the transition from the high-dimensional finite element space $\mathcal{V}_u^h$ to the desired reduced-order space $\mathcal{V}_u^*$ —in which the fine-scale BVP is to be finally posed— is the *dimensionality reduction* process, in which, as pointed out in the introductory section, the dominant deformational patterns of the RVE response are identified and unveiled by washing out the "inessentials". To accomplish this central task, we employ here a partitioned version of the *Proper Orthogonal Decomposition*[6].

3.2.1. Elastic/Inelastic reduced basis functions

The Proper Orthogonal Decomposition is nothing but a multidimensional data fitting procedure intended to obtain a sequence of orthogonal basis functions whose span best approximate the space of snapshots. As such, the POD is a *purely data-driven* process —it is "agnostic" to the physical origin of the data. For instance, for POD basis construction purposes, it is completely immaterial whether a given snapshot corresponds to a purely linear elastic solution or to a solution well into the inelastic regime. The task of discriminating which features of the RVE response are *essential* and which are not is exclusively guided by statistical considerations: if the elastic response happens to be poorly represented within the snapshot ensemble, the POD may regard as unimportant the contribution of these snapshots, and, as a consequence, the basis functions with largest associated singular values —i.e., the *essential* modes— would hardly contain any information of this range. To accurately replicate the apparently trivial linear elastic behavior, thus, one may be forced to take a relatively large number of basis functions, and this

[6]See Appendix A for a brief description of the POD.

may translate into a significant increase in the overall *online* computational cost. This fact certainly places the POD-based reduced basis approach at a competitive disadvantage compared with semi-analytical homogenization approaches such as the Nonlinear Transformation Field Analysis [50], which do capture exactly (and effortlessly) the linear elastic response of the RVE.

To eliminate this shortcoming, we propose here a slightly different strategy for constructing the reduced basis. The essence of the proposal is to partition the space of snapshots $\mathcal{V}_u^{snp}$ into *elastic* ($\mathcal{V}_{u,el}^{snp}$) and *inelastic* ($\mathcal{V}_{u,inel}^{snp}$) subspaces:

$$\mathcal{V}_u^{snp} = \mathcal{V}_{u,el}^{snp} \oplus \mathcal{V}_{u,inel}^{snp}, \tag{13}$$

($\oplus$ symbolizes direct sum of subspaces [55]) and then obtain the reduced basis as the *union* of the bases for both subspaces. Below, we describe this strategy more in detail.

The first step is to determine an orthogonal basis for $\mathcal{V}_{u,el}^{snp}$. One can do this by simply performing m_e independent, linear elastic finite element analysis of the RVE ($m_e = 6$ for 3D problems[7], and $m_e = 3$ for plane strain), and then orthonormalizing the resulting displacement fluctuation fields. These m_e elastic modes will be considered as the first m_e basis functions of the reduced basis:

$$\text{span}\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_{m_e}\} = \mathcal{V}_{u,el}^{snp}. \tag{14}$$

Once we have at our disposal this set of elastic basis functions, we compute the (orthogonal) projection of each snapshot $\boldsymbol{u}^k$ onto the orthogonal complement of $\mathcal{V}_{u,el}^{snp}$ (which is precisely the aforementioned *inelastic space* $\mathcal{V}_{u,inel}^{snp}$):

$$\boldsymbol{u}_{inel}^k := \boldsymbol{u}^k - \sum_{i=1}^{m_e} \left\langle \boldsymbol{\Phi}_i, \boldsymbol{u}^k \right\rangle_{L_2(\Omega)} \boldsymbol{\Phi}_i, \quad k = 1, 2 \ldots n_{snp}. \tag{15}$$

It is now on this ensemble of *inelastic* snapshots $\{\boldsymbol{u}_{inel}^k\}_{k=1}^{n_{snp}}$ that the previously described POD is applied to obtain the remaining $n_u - m_e$ basis functions. Thus, we finally have:

$$\mathcal{V}_u^* = \mathcal{V}_{u,el}^{snp} \oplus \mathcal{V}_{u,inel}^{snp} = \text{span}\{\overbrace{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_6}^{\text{Elastic modes}}, \quad \overbrace{\boldsymbol{\Phi}_7, \ldots, \boldsymbol{\Phi}_{n_u}}^{\text{“Essential” Inelastic modes}} \quad \}. \tag{16}$$

[7]Strictly speaking, the proposed decomposition is only valid for materials governed by rate-independent constitutive equations.

for 3D problems, and

$$\mathcal{V}_u^* = \text{span}\{\overbrace{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \boldsymbol{\Phi}_3}^{\text{Elastic modes}}, \quad \overbrace{\boldsymbol{\Phi}_4, \ldots, \boldsymbol{\Phi}_{n_u}}^{\text{“Essential” inelastic modes}} \quad \}. \tag{17}$$

for plane strain. In placing the m_e elastic modes within the first m_e positions, the reduced-order model is guaranteed to deliver linear elastic solutions with the same accuracy as the underlying (full-order) finite element model (obviously, provided that $n_u \geq m_e$).

Further details concerning the numerical implementation of this apparently novel —to the best of the authors' knowledge— basis construction strategy can be found in Appendix B.

4. Galerkin projection onto the reduced subspace

We now seek to pose the boundary-value problem represented by Eq.(3) in the reduced-order space $\mathcal{V}_u^* \subseteq \mathcal{V}_u^h$ spanned by the basis functions $\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \ldots, \boldsymbol{\Phi}_{n_u}\}$. To this end, we approximate both test $\boldsymbol{\eta} \in \mathcal{V}_u$ and trial $\boldsymbol{u} \in \mathcal{V}_u$ functions by the following linear expansions:

$$\boldsymbol{u}(\mathbf{x}; \boldsymbol{\epsilon_M}) \approx \boldsymbol{u}^*(\mathbf{x}; \boldsymbol{\epsilon_M}) = \sum_{i=1}^{n_u} \boldsymbol{\Phi}_i(\mathbf{x}) U_i^*(\boldsymbol{\epsilon_M}), \tag{18}$$

$$\boldsymbol{\eta}(\mathbf{x}) \approx \boldsymbol{\eta}^*(\mathbf{x}) = \sum_{i=1}^{n_u} \boldsymbol{\Phi}_i(\mathbf{x}) \eta_i^*, \tag{19}$$

$\boldsymbol{u}^*(\mathbf{x})$ and $\boldsymbol{\eta}^*(\mathbf{x})$ being the low-dimensional approximations of trial and test functions, respectively (hereafter, asterisked symbols will be used to denote low-dimensional approximations of the associated variables). Inserting Eqs. (18) and (19) into Eq.(3), and exploiting the arbitrariness of coefficients η_i^* ($i = 1, 2 \ldots n_u$), we arrive at the following set of n_u equilibrium equations:

$$\int_\Omega \nabla^s \boldsymbol{\Phi}_i(\mathbf{x}) : \boldsymbol{\sigma}(\mathbf{x}; \boldsymbol{\epsilon_M} + \nabla^s \boldsymbol{u}^*, \boldsymbol{\xi})\, d\Omega = 0, \quad i = 1, 2 \ldots n_u. \tag{20}$$

Expressing now the reduced basis functions in the above equation in terms of finite element shape functions (through expression $\boldsymbol{\Phi}_i(\mathbf{x}) = \sum_{I=1}^{n} N_I(\mathbf{x}) \boldsymbol{\Phi}_{Ii}$), we get (in Voigt's notation):

$$\int_\Omega \boldsymbol{B}_i^{*T}(\mathbf{x})\, \boldsymbol{\sigma}(\mathbf{x}; \boldsymbol{\epsilon_M} + \boldsymbol{B}^* \boldsymbol{U}^*, \boldsymbol{\xi})\, d\Omega = \mathbf{0}, \quad i = 1, 2 \ldots n_u, \tag{21}$$

or more compactly:

$$\int_{\Omega} \boldsymbol{B}^{*T}(\mathbf{x})\, \boldsymbol{\sigma}(\mathbf{x}; \boldsymbol{\epsilon_M} + \boldsymbol{B}^{*}\boldsymbol{U}^{*}, \boldsymbol{\xi})\, d\Omega = \mathbf{0}. \tag{22}$$

Here, $\boldsymbol{U}^{*} = \begin{bmatrix} U_1^{*} & U_2^{*} & \cdots & U_{n_u}^{*} \end{bmatrix}^{T} \in \mathbb{R}^{n_u}$ denotes the vector containing the *reduced* displacement fluctuations —the basic unknowns of the reduced-order problem— and $\boldsymbol{B}^{*} : \Omega \to \mathbb{R}^{s \times n_u}$ stands for the *reduced* "B-matrix", defined as:

$$\boldsymbol{B}^{*}(\mathbf{x}) := \boldsymbol{B}(\mathbf{x})\boldsymbol{\Phi}. \tag{23}$$

This matrix connects the gradient of the displacement fluctuation field with the vector of reduced displacement fluctuations:

$$\begin{aligned} \nabla^{s}\boldsymbol{u}^{*} = \sum_{i=1}^{n_u} \boldsymbol{B}_i^{*} U_i^{*} &= \overbrace{\begin{bmatrix} \boldsymbol{B}_1^{*} & \boldsymbol{B}_2^{*} & \cdots & \boldsymbol{B}_{n_u}^{*} \end{bmatrix}}^{\boldsymbol{B}^{*}} \overbrace{\begin{bmatrix} U_1^{*} \\ U_2^{*} \\ \vdots \\ U_{n_u}^{*} \end{bmatrix}}^{\boldsymbol{U}^{*}} \\ &= \boldsymbol{B}^{*}\boldsymbol{U}^{*} = \boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{U}^{*}. \end{aligned} \tag{24}$$

For implementational purposes, it is more expedient to express Eq.(23) in terms of elemental $B-$matrices. To this end, we write:

$$\boldsymbol{B}(\mathbf{x}) = \begin{cases} \boldsymbol{B}^{e}(\mathbf{x}), & \text{if } \mathbf{x} \in \Omega^{e} \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{25}$$

where $\boldsymbol{B}^{e} \in \mathbb{R}^{s \times d \cdot \bar{n}_e}$ denotes the local B-matrix of element Ω^{e} ($\bar{n}_e$, in turn, is the number of nodes in Ω^{e}). Thus,

$$\boldsymbol{B}^{*}(\mathbf{x}) = \boldsymbol{B}(\mathbf{x})\boldsymbol{\Phi} = \boldsymbol{B}^{e}(\mathbf{x})\boldsymbol{\Phi}^{e}. \tag{26}$$

In the above equation, $\boldsymbol{\Phi}^{e} \in \mathbb{R}^{d\bar{n}_e \times n_u}$ represents the block matrix of $\boldsymbol{\Phi}$ corresponding to the $\bar{n}_e$ nodes of finite element Ω^{e} ($e = 1, 2 \ldots n_e$).

5. Stress approximation space

To arrive at an efficient, mesh-size independent integration scheme, two crucial questions have to be addressed, namely, the determination of the

vector space (hereafter denoted by $\mathcal{V}_\sigma^{apr}$) in which the low-dimensional approximation of the *stress field*[8] should lie in order to obtain an accurate and at the same time well-posed ROM; and the calculation of the optimal location of the *sampling or integration points*. Attention here and in the next section is confined to the aspect related to the stress approximation space; the issue related to the selection of sampling points, on the other hand, is examined in Section 7.

5.1. The reduced-order subspace of statically admissible stresses ($\mathcal{V}_\sigma^*$)

At first sight, the problem of constructing a $\mathcal{O}(n_u)$-dimensional representation of the stress field seems quite similar to the problem addressed in Section 3 concerning the reduced basis for the displacement fluctuations: we have to find a set of orthogonal basis functions $\{\boldsymbol{\Psi}_1(\mathbf{x}), \boldsymbol{\Psi}_2(\mathbf{x}) \dots \boldsymbol{\Psi}_{n_\sigma}(\mathbf{x})\}$ ($n_\sigma = \mathcal{O}(n_u)$) such that its span accurately approximates the set of all possible *stress solutions* —that is, the set of all *statically admissible stresses*. Accordingly, following the procedure described in Section 3, we first compute finite element, stress distributions over the RVE for representative macrostrain histories[9]. Then, the elastic/inelastic dimensionality reduction process set forth in Section 3.2.1 is applied to the resulting ensemble of stress solutions $\{\boldsymbol{\sigma}^1(\mathbf{x}), \boldsymbol{\sigma}^2(\mathbf{x}) \dots \boldsymbol{\sigma}^{n_{snp}}(\mathbf{x})\}$, in order to identify both the elastic and the *essential* inelastic stress modes. The space spanned by these modes will be denoted hereafter by $\mathcal{V}_\sigma^*$ and termed the *reduced-order subspace of statically admissible stresses*:

$$\mathcal{V}_\sigma^* = \text{span}\{\overbrace{\boldsymbol{\Psi}_1(\mathbf{x}), \boldsymbol{\Psi}_2(\mathbf{x}), \dots, \boldsymbol{\Psi}_{m_e}(\mathbf{x})}^{\text{Elastic stress modes}}, \overbrace{\boldsymbol{\Psi}_{m_e+1}(\mathbf{x}), \boldsymbol{\Psi}_{m_e+2}(\mathbf{x}), \dots, \boldsymbol{\Psi}_{n_\sigma}(\mathbf{x})}^{\text{“Essential”, inelastic stress modes}}\}. \tag{27}$$

[8] It was mentioned in the introductory section that the central idea of efficient interpolatory approaches for numerical integration of reduced-order BVPs is to replace the nonaffine term in the the integrand by low-dimensional interpolants. In our case, a glance at the reduced-order equilibrium equation (21) readily reveals that such “offending”, nonaffine term is the *stress field* —the reduced B-matrix $\boldsymbol{B}^* = \boldsymbol{B}^*(\mathbf{x})$ is independent of the input parameter $\boldsymbol{\epsilon_M}$ and hence need not be subject to approximation.

[9] The most practical and somehow consistent choice regarding these strain trajectories is to use the same as in the computation of the displacement fluctuations snapshots

5.2. Ill-posedness of the HP-ROM

Let us now try to construct the low-dimensional approximation of the stress field, denoted by[10] $\boldsymbol{\sigma}^*$, as a linear combination of the above described stress reduced basis— hence making $\mathcal{V}_\sigma^{apr} = \mathcal{V}_\sigma^*$—; i.e.,

$$\boldsymbol{\sigma}(\mathbf{x}; \boldsymbol{\epsilon_M}, \boldsymbol{U}^*) \approx \boldsymbol{\sigma}^*(\mathbf{x}; \boldsymbol{\epsilon_M}, \boldsymbol{U}^*) = \sum_{i=1}^{n_\sigma} \boldsymbol{\Psi}_i(\mathbf{x}) c_i(\boldsymbol{\epsilon_M}, \boldsymbol{U}^*), \tag{28}$$

where $c_i \in \mathbb{R}$ $(i = 1, 2 \dots n_\sigma)$. This strategy of approximating the *offending*, *nonaffine* term in the BVP by a linear combination of pre-computed basis functions —obtained, in turn, from samples of the nonaffine term evaluated at the solution— has been successfully applied by several authors, with no apparent —or at least not reported— computational pitfalls, to a wide gamut of problems: nonlinear monotonic elliptic and nonlinear parabolic BPVs [46, 33], nonlinear miscible viscous fingering in porous media [17, 18], uncertainty quantification in inverse problems [31], and nonlinear heat conduction problems [5, 6], to cite but a few.

However, a closer examination of the the RVE equilibrium problem reveals that, in this case, this "standard" strategy proves completely fruitless, for it leads to patently *ill-posed* reduced-order equations. To show this, let us first substitute approximation (28) into Eq.(21):

$$\begin{aligned} &\int_\Omega \boldsymbol{B}^{*T}(\mathbf{x})\, \boldsymbol{\sigma}(\mathbf{x}; \boldsymbol{\epsilon_M}, \boldsymbol{U}^*)\, d\Omega \approx \int_\Omega \boldsymbol{B}^{*T}(\mathbf{x})\, \boldsymbol{\sigma}^*(\mathbf{x}; \boldsymbol{\epsilon_M}, \boldsymbol{U}^*) = \\ &\sum_{i=1}^{n_\sigma} \left(\int_\Omega \boldsymbol{B}^{*T}(\mathbf{x}) \boldsymbol{\Psi}_i(\mathbf{x})\, d\Omega \right) c_i(\boldsymbol{\epsilon_M}, \boldsymbol{U}^*) = \mathbf{0}. \end{aligned} \tag{29}$$

By virtue of Eq.(23), the bracketed integral in the preceding equation can be rephrased as:

$$\int_\Omega \boldsymbol{B}^{*T}(\mathbf{x}) \boldsymbol{\Psi}_i(\mathbf{x})\, d\Omega = \boldsymbol{\Phi}^T \left(\int_\Omega \boldsymbol{B}^T(\mathbf{x}) \boldsymbol{\Psi}_i(\mathbf{x})\, d\Omega \right), \quad i = 1, 2 \dots n_\sigma. \tag{30}$$

Each basis function $\boldsymbol{\Psi}_i(\mathbf{x})$ $(i = 1, 2 \dots n_\sigma)$ is, by construction, a linear combination of the stress snapshots collected during the offline, finite element

[10] Notice that, in keeping with the notational convention introduced in Section 4, the low-dimensional approximation of the stress field is represented by attaching an asterisk to the stress symbol.

analysis; thus, we can write $\boldsymbol{\Psi}_i = \sum_{j=1}^{n_{snp}} \beta_{ij} \boldsymbol{\sigma}^j$ $(i = 1, 2 \dots n_\sigma)$, $\beta_{ij} \in \mathbb{R}$ being the corresponding coefficients in the linear combination. Inserting the above equation into Eq.(30) and considering that $\boldsymbol{\sigma}^j$ $(j = 1, 2 \dots n_{snp})$ are finite element stress solutions —and therefore fulfill the finite element equilibrium equation (8)—, we finally arrive at:

$$\boldsymbol{\Phi}^T \sum_{j=1}^{n_{snp}} \beta_{ij} \overbrace{\left(\int_\Omega \boldsymbol{B}^T \boldsymbol{\sigma}^j \, d\Omega \right)}^{=0} = \mathbf{0}, \quad i = 1, 2 \dots n_\sigma, \tag{31}$$

that is, the integral (30) appearing in the equilibrium equation (29), and hence, the left-hand side of the equation itself, vanishes identically regardless of the value of the modal coefficients $c_i \in \mathbb{R}$ $(i = 1, 2 \dots n_\sigma)$, and therefore, regardless of the value of the reduced displacement fluctuations $\boldsymbol{U}^*$—hence the ill-posedness.

5.3. Proposed remedy: the expanded space approach

It is clear from the foregoing discussion that the root cause of the ill-posedness lies in the fact that the set of all admissible stress fields $(\mathcal{V}_\sigma)$ forms a *vector space*, and, consequently, the POD stress modes $\boldsymbol{\Psi}_i \in \mathcal{V}_\sigma$ $(i = 1, 2 \dots n_\sigma)$ —and any linear combination of them— turn out to be self-equilibrated fields. Thus, for the reduced-order problem to be well-posed, the approximation space $\mathcal{V}_\sigma^{apr}$ cannot be only formed by statically admissible stresses, but *it must also include statically inadmissible fields* —i.e. stress functions that do not satisfy the reduced-order equilibrium equation (21).

One plausible route for determining a low-dimensional approximation space that embraces both statically admissible and statically inadmissible stresses might be to collect, during the offline finite element calculations, not only converged stresses, but also the unconverged ones —i.e., those generated during the corresponding iterative algorithm—, and then perform the POD-based dimensionality reduction over the whole ensemble of snapshots. In the present work, however, we pursue an approach that precludes the necessity of undertaking this computationally laborious and in some aspects objectionable —there is no guarantee that the span of selected, unconverged stress snapshots covers the entire space of statically inadmissible stresses— process. The idea behind the employed approach was originally conceived, but not fully developed, by the authors in a recent monograph [35]. Here, the theory underlying such an idea is further elaborated and cast into the formalisms of functional analysis.

5.3.1. Continuum formulation

To originate our considerations from a general standpoint, it proves convenient first to rephrase the left-hand side of the reduced-order equilibrium equation Eq.(21) as the action of a certain linear operator $\boldsymbol{G} : L_2(\Omega)^s \to \mathbb{R}^{n_u}$ on the stress field over the RVE:

$$\int_\Omega \boldsymbol{B}_i^{*T} \boldsymbol{\sigma} \, d\Omega = \langle \boldsymbol{B}_i^*, \boldsymbol{\sigma} \rangle_{L_2(\Omega)} = (\boldsymbol{G}[\boldsymbol{\sigma}])_i \quad i = 1, 2 \dots n_u. \tag{32}$$

Invoking now the orthogonal decomposition of $L_2(\Omega)^s$ induced by this operator, one obtains:

$$L_2(\Omega)^s = \mathcal{N}(\boldsymbol{G}) \oplus \text{span}\{\boldsymbol{B}_i^*\}_{i=1}^{n_u}, \tag{33}$$

where $\mathcal{N}(\boldsymbol{G})$ stands for the nullspace of $\boldsymbol{G}$. Since the RVE equilibrium equation has a vanishing right-hand side term, it follows that $\mathcal{N}(\boldsymbol{G})$ is actually the space of *statically admissible stress fields*. Its orthogonal complement, $\text{span}\{\boldsymbol{B}_i^*\}_{i=1}^{n_u}$, can be therefore construed as the abovementioned *space of statically inadmissible stresses. The key fact here is that such a space is inherently n_u-dimensional and, thus, there is no need to perform any dimensionality reduction whatsoever over* unconverged *snapshots to arrive at the desired basis: the strain-displacement functions $\{\boldsymbol{B}_1^*, \boldsymbol{B}_2^* \dots \boldsymbol{B}_{n_u}^*\}$ themselves are linearly independent (albeit not orthogonal) and can thereby serve this very purpose.*

According to the preceding decomposition, any $\boldsymbol{\sigma} \in L_2(\Omega)^s$ can be resolved as (see Figure 2):

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^{ad} + \boldsymbol{\sigma}^{in}, \quad \text{with } \left\langle \boldsymbol{\sigma}^{ad}, \boldsymbol{\sigma}^{in} \right\rangle_{L_2(\Omega)} = 0, \tag{34}$$

where $\boldsymbol{\sigma}^{ad} \in \mathcal{N}(\boldsymbol{G})$ and $\boldsymbol{\sigma}^{in} \in \text{span}\{\boldsymbol{B}_i^*\}_{i=1}^{n_u}$ stand for the statically admissible and statically inadmissible components of $\boldsymbol{\sigma}$, respectively. Following the standard approach, the statically admissible component $\boldsymbol{\sigma}^{ad}$ —i.e., the stress solution we wish to calculate for a given input $\boldsymbol{\epsilon_M}$— is forced to lie in the span of the POD modes $\boldsymbol{\Psi}_i$ $(i = 1, 2 \dots n_\sigma)$ obtained from converged snapshots:

$$\boldsymbol{\sigma}^{ad} \approx \boldsymbol{\sigma}^* = \sum_{i=1}^{n_\sigma} \boldsymbol{\Psi}_i c_i^{ad}, \tag{35}$$

$c_i^{ad} \in \mathbb{R}$ $(i = 1, 2 \dots n_\sigma)$ being the corresponding modal coefficients. The non-equilibrated component $\boldsymbol{\sigma}^{in}$, on the other hand, resides naturally in the span

of the reduced strain-displacement functions, so we can directly write—i.e., without introducing further approximations—:

$$\boldsymbol{\sigma}^{in} = \sum_{i=1}^{n_u} \boldsymbol{B}_i^* c_i^{in}, \tag{36}$$

with $c_i^{in} \in \mathbb{R}$ $(i = 1, 2 \dots n_u)$. The low-dimensional approximation required in the proposed integration method, denoted in what follows by $\boldsymbol{\sigma}^{ex*}$ (the appended superscript "ex" means "stress approximated in the *expanded* space"), is finally obtained as the sum of Eq.(35) and Eq.(36) :

$$\boldsymbol{\sigma}^{ex*} = \sum_{i=1}^{n_\sigma} \boldsymbol{\Psi}_i c_i^{ad} + \sum_{j=1}^{n_u} \boldsymbol{B}_j^* c_j^{in}. \tag{37}$$

Substituting the above approximation into the equilibrium equation, one

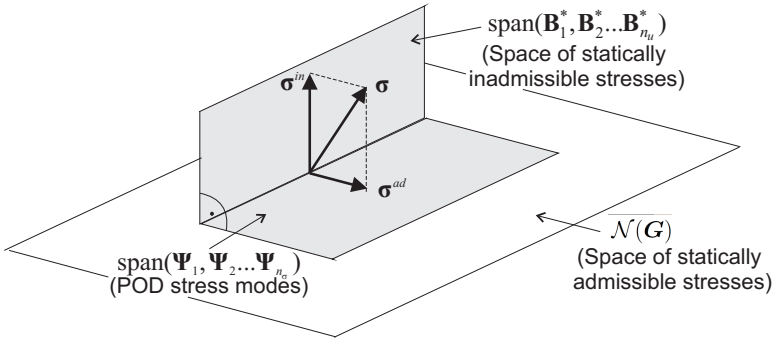


Figure 2: Expanded space approach. The stress approximation space is expanded so that it embraces, not only the span of the stress POD modes, but also the span of the reduced strain-displacement functions $\{\boldsymbol{B}_1^*, \boldsymbol{B}_2^* \dots \boldsymbol{B}_{n_u}^*\}$. The reduced-order RVE equilibrium problem boils down to find the reduced displacement fluctuations vector $\boldsymbol{U}^*$ that makes the *non-equilibrated* component $\boldsymbol{\sigma}^{in}$ to vanish ($\boldsymbol{\sigma}^{in}(\boldsymbol{U}^*, \boldsymbol{\epsilon_M}) = \boldsymbol{0}$).

gets:

$$\int_\Omega \boldsymbol{B}_i^{*T} \boldsymbol{\sigma}^{ex*} = \sum_{j=1}^{n_u} \left(\int_\Omega \boldsymbol{B}_i^{*T} \boldsymbol{B}_j^* d\Omega \right) c_j^{in} = \boldsymbol{0}, \quad i = 1, 2 \dots n_u. \tag{38}$$

Since $\{\boldsymbol{B}_1^*, \boldsymbol{B}_2^* \dots \boldsymbol{B}_{n_u}^*\}$ are linearly independent functions, it becomes immediately clear that the above equations holds only if:

$$c_j^{in}(\boldsymbol{\epsilon_M}, \boldsymbol{U}^*) = 0, \quad j = 1, 2 \dots n_u, \tag{39}$$

i.e., if the n_u coefficients multiplying $\boldsymbol{B}_i^* \in L_2(\Omega)^s$ $(i = 1, 2 \dots n_u)$ are identically zero. In adopting the proposed integration approach, thus, the reduced-order RVE equilibrium problem (21) is transformed into the problem of finding, for a given input macroscopic strain tensor $\boldsymbol{\epsilon_M}$, the reduced displacement fluctuations vector $\boldsymbol{U}^* \in \mathbb{R}^{n_u}$ that makes the non-equilibrated component $\boldsymbol{\sigma}^{in}$ (defined in Eq.(36)) to vanish.

In a nutshell, *the ill-posedness exhibited by the discrete problem when adopting the standard approach of using only POD modes is eliminated by expanding the stress approximation space so that it embraces also the span of the reduced strain-displacement functions* (or strain modes[11]) $\boldsymbol{B}_i^* \in L_2(\Omega)^s$ $(i = 1, 2 \dots n_u)$:

$$\mathcal{V}_\sigma^{apr} = \mathcal{V}_\sigma^* \oplus \text{span}\{\boldsymbol{B}_i^*\}_{i=1}^{n_u} = \text{span}\{\overbrace{\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \dots \boldsymbol{\Psi}_{n_\sigma}}^{n_\sigma \text{ stress modes}}, \overbrace{\boldsymbol{B}_1^*, \boldsymbol{B}_2^* \dots \boldsymbol{B}_{n_u}^*}^{n_u \text{ strain modes}}\}. \quad (40)$$

5.3.2. Discrete formulation

In typical finite element implementations, both stresses and gradients of shape functions are only calculated and stored at the Gauss points of the underlying spatial discretization. For practical reasons, thus, it proves imperative to reformulate the above explained *expanded space* strategy and treat both magnitudes as spatially discrete variables, defined only at such Gauss points.

The *discrete* counterparts of the continuously defined fields $\boldsymbol{\sigma} \in L_2(\Omega)^s$ and $\boldsymbol{B}_i^* \in L_2(\Omega)^s$ $(i = 1, 2 \dots n_u)$ will be denoted by $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{n_g \cdot s}$ and $\boldsymbol{\mathcal{B}^*} = \begin{bmatrix} \boldsymbol{\mathcal{B}_1^*} & \boldsymbol{\mathcal{B}_2^*} & \cdots & \boldsymbol{\mathcal{B}_{n_u}^*} \end{bmatrix} \in \mathbb{R}^{n_g \cdot s \times n_u}$, and termed the *global stress vector*, and the *global matrix of strain modes*, respectively. The *global stress vector* $\boldsymbol{\mathcal{S}}$ is constructed by stacking the stress vectors $\boldsymbol{\sigma}(\mathbf{x}_g; \cdot) \in \mathbb{R}^s$ $(g = 1, 2 \dots n_g)$ at the Gauss points of the finite element grid into a single column vector:

$$\boldsymbol{\mathcal{S}} := \begin{bmatrix} \boldsymbol{\sigma}^T(\mathbf{x}_1; \cdot) & \boldsymbol{\sigma}^T(\mathbf{x}_2; \cdot) & \cdots & \boldsymbol{\sigma}^T(\mathbf{x}_{n_g}; \cdot) \end{bmatrix}^T. \quad (41)$$

Similarly, the global matrix of strain modes $\boldsymbol{\mathcal{B}^*}$ is constructed as:

$$\boldsymbol{\mathcal{B}^*} := \begin{bmatrix} \boldsymbol{B}^{*T}(\mathbf{x}_1) & \boldsymbol{B}^{*T}(\mathbf{x}_2) & \cdots & \boldsymbol{B}^{*T}(\mathbf{x}_{n_g}) \end{bmatrix}^T. \quad (42)$$

[11] Indeed, functions $\boldsymbol{B}_i^* \in L_2(\Omega)^s$ $(i = 1, 2 \dots n_u)$ can be viewed as fluctuating *strain* modes, since they are the symmetric gradient of the displacement fluctuation modes, see Eq. 23.

Having definitions (41) and (42) at hand, the approximation of Eq.(22) by Gauss quadrature can be written as:

$$\begin{aligned} &\int_\Omega \boldsymbol{B}_i^{*T}(\mathbf{x})\boldsymbol{\sigma}(\mathbf{x};\cdot)\,d\Omega \approx \sum_{g=1}^{n_g} w_g \boldsymbol{B}_i^{*T}(\mathbf{x}_g)\boldsymbol{\sigma}(\mathbf{x}_g;\cdot) = 0 \\ &\Rightarrow \boldsymbol{\mathcal{B}}_{\boldsymbol{i}}^{*T}\boldsymbol{W}\boldsymbol{\mathcal{S}} = 0, \qquad i = 1,2\ldots n_u, \end{aligned} \tag{43}$$

where $\boldsymbol{W}$ is a diagonal matrix containing the weights at each Gauss point:

$$\boldsymbol{W} := \begin{bmatrix} w_1\boldsymbol{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & w_2\boldsymbol{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & w_{n_g}\boldsymbol{I} \end{bmatrix} \tag{44}$$

(here, $\boldsymbol{I}$ denotes the $s\,\mathrm{x}\,s$ identity matrix). Assuming that $w_g > 0$ ($g = 1,2\ldots n_g$) —Gauss quadrature rules with negative weights are excluded from our considerations—, and using the Cholesky decomposition of $\boldsymbol{W}$ ($\boldsymbol{W} = \boldsymbol{W}^{1/2}\boldsymbol{W}^{1/2}$), one can reexpress Eq.(43) as

$$\boldsymbol{\mathcal{B}}^{*T}\boldsymbol{W}\boldsymbol{\mathcal{S}} = (\boldsymbol{\mathcal{B}}^{*T}\boldsymbol{W}^{1/2})(\boldsymbol{W}^{1/2}\boldsymbol{\mathcal{S}}) = \mathbf{0}. \tag{45}$$

Defining now the *weighted* global stress vector and *weighted* matrix of strain modes as

$$\boldsymbol{\Sigma} := \boldsymbol{W}^{1/2}\boldsymbol{\mathcal{S}} = \begin{bmatrix} \sqrt{w_1}\boldsymbol{\sigma}^T(\mathbf{x}_1;\cdot) & \sqrt{w_2}\boldsymbol{\sigma}^T(\mathbf{x}_2;\cdot) & \cdots & \sqrt{w_{n_g}}\boldsymbol{\sigma}^T(\mathbf{x}_{n_g};\cdot) \end{bmatrix}^T, \tag{46}$$

and

$$\mathbb{B}^* := \boldsymbol{W}^{1/2}\boldsymbol{\mathcal{B}}^* = \begin{bmatrix} \sqrt{w_1}\boldsymbol{B}^{*T}(\mathbf{x}_1) & \sqrt{w_2}\boldsymbol{B}^{*T}(\mathbf{x}_2) & \cdots & \sqrt{w_{n_g}}\boldsymbol{B}^{*T}(\mathbf{x}_{n_g}) \end{bmatrix}^T \tag{47}$$

respectively, and inserting these definitions into Eq.(41), one finally arrives at:

$$\mathbb{B}^{*T}\boldsymbol{\Sigma} = \mathbf{0}, \tag{48}$$

or equivalently,

$$\mathbb{B}_i^{*T}\boldsymbol{\Sigma} = 0, \qquad i = 1,2\ldots n_u, \tag{49}$$

which shows that any statically admissible *weighted* stress vector is orthogonal, in the sense of the standard euclidean inner product, to the *weighted* strain modes $\mathbb{B}_i^{*T}$ ($i = 1,2\ldots n_u$).

Comparing Eq.(48) with Eq.(32), it becomes clear that $\mathbb{B}^{*T}$ plays the same role as operator $\boldsymbol{G}$ in Eq.(32). In analogy with Eq.(33), thus, we can write

$$\mathbb{R}^{n_g \cdot s} = \mathcal{N}(\mathbb{B}^{*T}) \oplus \text{Range}(\mathbb{B}^*) \tag{50}$$

where $\mathcal{N}(\mathbb{B}^{*T})$ and $\text{Range}(\mathbb{B}^*)$ denote the null space and the range (or column space) of $\mathbb{B}^{*T}$ and $\mathbb{B}^*$, respectively, and consequently decompose any $\boldsymbol{\Sigma} \in \mathbb{R}^{n_g \cdot s}$ as

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma^{ad}} + \boldsymbol{\Sigma^{in}} \tag{51}$$

with $\boldsymbol{\Sigma^{ad}} \in \mathcal{N}(\mathbb{B}^{*T})$ and $\boldsymbol{\Sigma^{in}} \in \text{Range}(\mathbb{B}^*)$. As in the continuous case (see Eq.(35)), the statically admissible component $\boldsymbol{\Sigma^{ad}}$ is now *approximated* by a linear combination of POD basis vectors obtained from converged stress snapshots[12]:

$$\boldsymbol{\Sigma^{ad}} \approx \boldsymbol{\Sigma}^* = \sum_{i=1}^{n_\sigma} c_i^{ad} \boldsymbol{\Psi}_i = \boldsymbol{\Psi} \boldsymbol{c}^{ad}, \tag{52}$$

where $\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Psi}_1 & \boldsymbol{\Psi}_2 & \cdots & \boldsymbol{\Psi}_{n_\sigma} \end{bmatrix}$ denotes the (weighted) stress basis matrix and $\boldsymbol{c}^{ad} \in \mathbb{R}^{n_\sigma}$ stands for the vector of modal coefficients associated to such a basis matrix. Likewise, since the non-equilibrated component $\boldsymbol{\Sigma^{in}}$ pertains to the column space of $\mathbb{B}^*$, we can directly write

$$\boldsymbol{\Sigma^{in}} = \mathbb{B}^* \boldsymbol{c}^{in}, \tag{53}$$

where $\boldsymbol{c}^{in} \in \mathbb{R}^{n_u}$. The low-dimensional (weighted) stress vector $\boldsymbol{\Sigma}^{ex*}$ required in the proposed integration method is finally obtained as the sum of Eq.(53) and Eq.(52).

$$\boldsymbol{\Sigma} \approx \boldsymbol{\Sigma}^{ex*} = \boldsymbol{\Psi} \boldsymbol{c}^{ad} + \mathbb{B}^* \boldsymbol{c}^{in}, \tag{54}$$

or in a more compact format:

$$\boldsymbol{\Sigma}^{ex*} = \boldsymbol{\Psi}^{ex} \boldsymbol{c}. \tag{55}$$

where

$$\boldsymbol{\Psi}^{ex} := \begin{bmatrix} \boldsymbol{\Psi} & \mathbb{B}^* \end{bmatrix}, \tag{56}$$

[12]The methodology for obtaining these modes using the SVD is similar to that explained in Section 3.2 for the displacement fluctuation modes.

and

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}^{ad} \\ \boldsymbol{c}^{in} \end{bmatrix}. \tag{57}$$

The matrix $\boldsymbol{\Psi}^{ex} \in \mathbb{R}^{n_g \cdot s \times (n_u + n_\sigma)}$ defined by Eq.(56) will be hereafter called the *expanded* basis matrix for the (weighted) stresses, whereas $\boldsymbol{c} \in \mathbb{R}^{n_\sigma + n_u}$ will be correspondingly termed the *expanded* vector of modal coefficients. Inserting approximation (54) into Eq.(48), and considering that $\mathbb{B}^{*T}\boldsymbol{\Psi} = \mathbf{0}$ and that $\mathbb{B}^{*T}$ is a full rank matrix, one finally arrives at the same equilibrium condition derived in the continuum case (see Eq. 39):

$$\boldsymbol{c}^{in}(\boldsymbol{U}^*, \boldsymbol{\epsilon_M}) = \mathbf{0}. \tag{58}$$

Once the above equation is solved for $\boldsymbol{U}^*$, the desired equilibrated stress vector $\boldsymbol{\Sigma}^*$ is obtained by evaluating Eq.(52):

$$\boldsymbol{\Sigma}^* = \boldsymbol{\Psi}\boldsymbol{c}^{ad}(\boldsymbol{U}^*, \boldsymbol{\epsilon_M}). \tag{59}$$

6. The High-Performance ROM

The next step in the development of the proposed integration scheme is to deduce closed-form expressions for the vectors of modal coefficients $\boldsymbol{c}^{ad} \in \mathbb{R}^{n_\sigma}$ and $\boldsymbol{c}^{in} \in \mathbb{R}^{n_u}$ in terms of the stress values computed at a set of $p = \mathcal{O}(n_u)$ *pre-specified sampling points* (to be chosen among the set of Gauss points of the underlying finite element mesh). To this end, we need first to introduce some notation and terminology.

6.1. Gappy vectors

Let $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2 \ldots \mathcal{I}_p\} \subset \{1, 2 \cdots n_g\}$ denote the set of indices of sampling points. Notationally, we write $\hat{\boldsymbol{\Sigma}}_{(\mathcal{I})} \in \mathbb{R}^{p \cdot s}$ to designate the subvector of $\boldsymbol{\Sigma}$ containing the rows associated to these sampling points; viz.:

$$\hat{\boldsymbol{\Sigma}}_{(\mathcal{I})} := \left[\sqrt{w_{\mathcal{I}_1}}\boldsymbol{\sigma}^T(\mathbf{x}_{\mathcal{I}_1}, \cdot) \quad \sqrt{w_{\mathcal{I}_2}}\boldsymbol{\sigma}^T(\mathbf{x}_{\mathcal{I}_2}, \cdot) \quad \cdots \quad \sqrt{w_{\mathcal{I}_p}}\boldsymbol{\sigma}^T(\mathbf{x}_{\mathcal{I}_p}, \cdot)\right]^T \tag{60}$$

(When confusion is not apt to arise, the parenthetical subscript indicating the set of sampling indices will be dropped, and we shall simply write $\hat{\boldsymbol{\Sigma}}$). It proves conceptually advantageous to regard this *restricted* or "gappy" — a terminology that goes back to the work of Everson et al. [28]— stress vector $\hat{\boldsymbol{\Sigma}}_{(\mathcal{I})}$ as the result of the application of a certain boolean operator

$\boldsymbol{\mathcal{P}}_{(\mathcal{I})} : \mathbb{R}^{n_g \cdot s} \to \mathbb{R}^{p \cdot s}$ over the full vector $\boldsymbol{\Sigma}$ (i.e., $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{\Sigma}$). We call $\boldsymbol{\mathcal{P}}_{(\mathcal{I})}$ the *selection operator* associated to sampling indices $\mathcal{I}$. This operator can be of course applied to any $\mathbf{Y} \in \mathbb{R}^{\mathbf{n_g \cdot s \times z}}$ ($z \in \mathbb{N}$). For instance, the restricted matrix of weighted strain modes is defined as $\hat{\mathbb{B}}^* := \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\mathbb{B}^*$. Furthermore, it is straighforward to show that

$$\boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{\mathcal{P}}_{(\mathcal{I})}^T = \boldsymbol{I}, \tag{61}$$

(here $\boldsymbol{I}$ is the $(n_g \cdot s)$x$(n_g \cdot s)$ identity matrix) and that

$$\boldsymbol{\mathcal{P}}_{(\mathcal{I})}(\mathbf{AY}) = (\boldsymbol{\mathcal{P}}_{(\mathcal{I})}\mathbf{A}\boldsymbol{\mathcal{P}}_{(\mathcal{I})}^{\mathbf{T}})(\boldsymbol{\mathcal{P}}_{(\mathcal{I})}\mathbf{Y}) \tag{62}$$

for any $\mathbf{A} \in \mathbb{R}^{\mathbf{n_g \cdot s \times n_g \cdot s}}$ and $\mathbf{Y} \in \mathbb{R}^{\mathbf{n_g \cdot s \times z}}$.

6.2. Least-squares fit

In the spirit of classical polynomial quadrature, such as Newton-Cotes formulae [36], the modal coefficients $\boldsymbol{c}^{ad} \in \mathbb{R}^{n_\sigma}$ and $\boldsymbol{c}^{in} \in \mathbb{R}^{n_u}$ are determined by *fitting* the low-dimensional approximation (54) to the weighted stresses calculated at the pre-specified sampling points. It should be noticed that, the variable subject to approximation —the stress— being a vector-valued function, the total number of discrete points to be fitted does not coincide with the number of spatial sampling points (p), but rather is equal to the product of such a number times the number of stress components (s). The well-posedness of the fitting problem, thus, demands that $p \cdot s \geq n_\sigma + n_u$, i.e., the number of discrete points must be equal or greater than the number of parameters to be adjusted. For the equality to hold, both $n_\sigma + n_u$ and p have to be multiple of s; thus, an *exact fit* is in general not possible for arbitrary values of n_σ and n_u, and recourse to an *approximate fit* is to be made. In this respect, we follow here the standard approach of using a *least-squares*, best-fit criterion, i.e., minimization of the squares of the deviations between "observed" ($\hat{\boldsymbol{\Sigma}}$) and fitted ($\hat{\boldsymbol{\Sigma}}^{ex*} = \hat{\boldsymbol{\Psi}}\boldsymbol{a} + \hat{\mathbb{B}}^*\boldsymbol{b}$) values (in our context, "observed" signifies "calculated through the pertinent constitutive equation"). This minimization problem can be stated as:

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}^{ad} \\ \boldsymbol{c}^{in} \end{bmatrix} = \arg \min_{\boldsymbol{a} \in \mathbb{R}^{n_\sigma}, \boldsymbol{b} \in \mathbb{R}^{n_u}} \| \hat{\boldsymbol{\Sigma}} - \left(\hat{\boldsymbol{\Psi}}\boldsymbol{a} + \hat{\mathbb{B}}^*\boldsymbol{b} \right) \| \tag{63}$$

where $\| \cdot \|$ stands for the standard euclidean norm. Let $\hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}} = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{\Psi}^{ex} = [\hat{\boldsymbol{\Psi}} \ \hat{\mathbb{B}}^*]$ be the *gappy* expanded basis matrix, and suppose that the sampling indices $\mathcal{I}$ have been chosen so that $\hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}}$ has full rank, i.e.:

$$\text{rank}(\hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}}) = \text{rank}([\hat{\boldsymbol{\Psi}} \ \hat{\mathbb{B}}^*]) = n_\sigma + n_u. \tag{64}$$

Then, it can be shown (see, for instance, Ref. [23]) that the solution of this standard, least-squares problem is provided by the following vector of coefficients:

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}^{ad} \\ \boldsymbol{c}^{in} \end{bmatrix} = \hat{\boldsymbol{\Psi}}^{ex\dagger} \hat{\boldsymbol{\Sigma}}, \tag{65}$$

where

$$\hat{\boldsymbol{\Psi}}^{ex\dagger} := \overbrace{(\hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}T} \hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}})^{-1}}^{\hat{\boldsymbol{M}}^{-1}} \hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}T} \tag{66}$$

is the so-called *pseudo-inverse* of matrix $\hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}}$.

Recall that our ultimate aim is to derive closed-form expressions for $\boldsymbol{c}^{in}$ and $\boldsymbol{c}^{ad}$ as functions of $\hat{\boldsymbol{\Sigma}}$. Thus, it remains to extricate these two sub-vectors from expression (65). This can be done by first partitioning both $\hat{\boldsymbol{M}} = \hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}T} \hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}}$ and $\hat{\boldsymbol{\Psi}}^{\boldsymbol{ex}T}$ in terms of the gappy stress basis matrix $\hat{\boldsymbol{\Psi}}$ and the gappy matrix of strain modes $\hat{\mathbb{B}}^*$:

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}^{ad} \\ \boldsymbol{c}^{in} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Psi}}^T \hat{\boldsymbol{\Psi}} & \hat{\boldsymbol{\Psi}}^T \hat{\mathbb{B}}^* \\ \hat{\mathbb{B}}^{*T} \hat{\boldsymbol{\Psi}} & \hat{\mathbb{B}}^{*T} \hat{\mathbb{B}}^* \end{bmatrix}^{-1} \begin{bmatrix} \hat{\boldsymbol{\Psi}}^T \\ \hat{\mathbb{B}}^{*T} \end{bmatrix} \hat{\boldsymbol{\Sigma}}. \tag{67}$$

Invoking the blockwise inverse formula for 2x2 block symmetric matrices [11], and upon tedious algebra —that has been relegated to Appendix C— one finally arrives at the following expressions for $\boldsymbol{c}^{ad}$ and $\boldsymbol{c}^{in}$

$$\boldsymbol{c}^{ad} = \hat{\boldsymbol{\Psi}}^{\dagger} (\hat{\boldsymbol{\Sigma}} - \hat{\mathbb{B}}^* \boldsymbol{c}^{in}), \tag{68}$$

$$\boldsymbol{c}^{in} = \mathbf{S}^{-\mathbf{1}} \hat{\mathbb{B}}^{*\mathbf{T}} \left(\boldsymbol{I} - \hat{\boldsymbol{\Psi}} \hat{\boldsymbol{\Psi}}^{\dagger} \right) \hat{\boldsymbol{\Sigma}}, \tag{69}$$

where $\hat{\boldsymbol{\Psi}}^{\dagger}$ denotes the pseudoinverse of the gappy stress basis matrix $\hat{\boldsymbol{\Psi}}$:

$$\hat{\boldsymbol{\Psi}}^{\dagger} = (\hat{\boldsymbol{\Psi}}^T \hat{\boldsymbol{\Psi}})^{-1} \hat{\boldsymbol{\Psi}}^T \tag{70}$$

and $\mathbf{S} := \hat{\mathbb{B}}^{*\mathbf{T}} (\boldsymbol{I} - \hat{\boldsymbol{\Psi}} \hat{\boldsymbol{\Psi}}^{\dagger})$ (note that $\mathbf{S}$ is invertible by virtue of the hypothesis represented by Eq.(64)).

6.2.1. Reconstruction matrix

Let us first examine expression (68) for the modal coefficients $\boldsymbol{c}^{ad}$ —those that multiply the *statically admissible* component of the global stress vector.

Since, at the solution, $\boldsymbol{c}^{in} = \mathbf{0}$, we have that:

$$\boldsymbol{c}^{ad} = \hat{\boldsymbol{\Psi}}^{\dagger}(\hat{\boldsymbol{\Sigma}} - \hat{\mathbb{B}}^{*}\overbrace{\boldsymbol{c}^{in}}^{=\mathbf{0}}) = \hat{\boldsymbol{\Psi}}^{\dagger}\hat{\boldsymbol{\Sigma}}. \tag{71}$$

(Notice that this result can also be obtained by directly solving minimization problem (63) with $\boldsymbol{b} = \mathbf{0}$). Substitution of this equation into Eq.(59) yields:

$$\boldsymbol{\Sigma}^{*} = \boldsymbol{\Psi}\boldsymbol{c}^{ad} = \overbrace{\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{\dagger}}^{\boldsymbol{R}}\hat{\boldsymbol{\Sigma}} = \boldsymbol{R}\,\hat{\boldsymbol{\Sigma}}, \tag{72}$$

where

$$\boldsymbol{R} := \boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{\dagger} = \boldsymbol{\Psi}(\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^{T}. \tag{73}$$

Inspection of Eq.(72) reveals that the matrix $\boldsymbol{R} \in \mathbb{R}^{n_g \cdot s \times p \cdot s}$ defined above is the operator that allows one to *reconstruct* the (weighted) *statically admissible* stress vector $\boldsymbol{\Sigma}^{*} \in \mathbb{R}^{n_g \cdot s}$ using only the (weighted) stress values ($\hat{\boldsymbol{\Sigma}} \in \mathbb{R}^{p \cdot s}$) calculated at the pre-selected sampling points $\mathcal{I}$. For this reason, we shall use the term *weighted reconstruction matrix* (or simply reconstruction matrix) to refer to this operator. It must be emphasized here that this matrix only depends on the POD stress basis matrix $\boldsymbol{\Psi}$ and on the selected sampling indices $\mathcal{I}$ —i.e., it is independent of the input parameter, the macro-strain $\boldsymbol{\epsilon_M}$—and, therefore, it can be pre-computed *offline*.

6.3. "Hyperreduced" RVE equilibrium equation

As for the expression for the set of "statically inadmissible" coefficients $\boldsymbol{c}^{in} \in \mathbb{R}^{n_u}$, we know that, at the solution, these coefficients must vanish; thus, from Eq.(69), we have

$$\boldsymbol{c}^{in}(\boldsymbol{U}^{*}, \boldsymbol{\epsilon_M}) = \mathbf{S}^{-\mathbf{1}}\hat{\mathbb{B}}^{*\mathbf{T}}\left(\boldsymbol{I} - \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger}\right)\hat{\boldsymbol{\Sigma}}(\boldsymbol{U}^{*}, \boldsymbol{\epsilon_M}) = \mathbf{0}. \tag{74}$$

Since $\mathbf{S}$ is a nonsingular matrix, the above condition is equivalent to

$$\hat{\mathbb{B}}^{*T}\left(\boldsymbol{I} - \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger}\right)\hat{\boldsymbol{\Sigma}}(\boldsymbol{U}^{*}, \boldsymbol{\epsilon_M}) = \mathbf{0}. \tag{75}$$

Furthermore, examination of Eq.(73) and Eq.(75) readily shows that the bracketed term $\hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger}$ in Eq.(75) is nothing but the submatrix of the reconstruction matrix $\boldsymbol{R}$ formed by the rows associated to sampling points $\mathcal{I}$, i.e.:

$$\hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger} = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}(\boldsymbol{\Psi}\hat{\boldsymbol{\Psi}}^{\dagger}) = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{R} = \hat{\boldsymbol{R}}. \tag{76}$$

Substitution of expression (76) into Eq.(75) finally leads to:

$$\hat{\mathbb{B}}^{*^T}(\boldsymbol{I}-\hat{\boldsymbol{R}})\hat{\boldsymbol{\Sigma}}(\boldsymbol{U}^*,\boldsymbol{\epsilon_M})=\mathbf{0}. \tag{77}$$

As previously noted (see Figure 2), the purpose of enforcing condition $\boldsymbol{c}^{in}(\boldsymbol{U}^*,\boldsymbol{\epsilon_M})=\mathbf{0}$ is to ensure that the stress solution lies entirely in the space of equilibrated stresses. Equation (77) can be viewed, thus, as the "hyperreduced" form of the original RVE equilibrium equation.

Observation 6.1. *The "hyperreduced" qualifier —coined by D. Ryckelynck [58, 59]— is used here to indicate that Eq.(77) is the result of two subsequent steps of complexity reduction: firstly, in the number of degrees of freedom (when passing from the finite element model to the ROM that employs standard Gauss quadrature), and, secondly, in the number of integration points (when passing from this standard ROM to what we have baptized*[13] *"High-Performance" ROM). This double complexity reduction can be better appreciated by rephrasing both Eq.(77) and the FE equation (9) in a format similar to that of Eq.(48), viz.:*

$$\overbrace{\hat{\mathbb{B}}^{*^T}(\boldsymbol{I}-\hat{\boldsymbol{R}})}^{\hat{\mathbb{B}}^{**^T}}\hat{\boldsymbol{\Sigma}}=\hat{\mathbb{B}}^{**^T}\hat{\boldsymbol{\Sigma}}=\mathbf{0}, \tag{78}$$

and

$$\sum_{g=1}^{n_g} w_g\boldsymbol{B}^T(\boldsymbol{x}_g)\boldsymbol{\sigma}(\boldsymbol{x}_g,;)=\mathbb{B}^T\boldsymbol{\Sigma}=\mathbf{0}, \tag{79}$$

respectively (here, $\mathbb{B}\in\mathbb{R}^{n_g\cdot s\times n\cdot d}$ is the finite element counterpart of $\mathbb{B}^*$, defined in Eq.(42)). With Eq.(79), Eq.(48) and Eq.(78) at our disposal, the abovementioned process of complexity reduction can be symbolically represented as

$$\overbrace{\mathbb{B}^T\boldsymbol{\Sigma}=\mathbf{0}}^{FEM} \quad \overset{1^{st}reduc.}{\underset{n\cdot d\to n_u}{\Longrightarrow}} \quad \overbrace{\mathbb{B}^{*^T}\boldsymbol{\Sigma}=\mathbf{0}}^{Stand.\ ROM} \quad \overset{2^{nd}reduc.}{\underset{n_g\to p}{\Longrightarrow}} \quad \overbrace{\hat{\mathbb{B}}^{**^T}\hat{\boldsymbol{\Sigma}}=\mathbf{0}}^{HP\text{-}ROM}, \tag{80}$$

[13] The term *High-Performance, Reduced-Order Model* (HP-ROM) is used to highlight the tremendous gains in performance that affords this model over the standard ROM, let alone over the full-order, finite model. In the numerical example shown in Section 9, we report speedup factors of above three order of magnitudes.

the relation between $\mathbb{B} \in \mathbb{R}^{n_g \cdot s \times n \cdot d}$, $\mathbb{B}^* \in \mathbb{R}^{n_g \cdot s \times n_u}$ *and* $\hat{\mathbb{B}}^{**} \in \mathbb{R}^{p \cdot s \times n_u}$ *being* $\mathbb{B}^* = \mathbb{B}\boldsymbol{\Phi}$ *and*

$$\hat{\mathbb{B}}^{**} := (\boldsymbol{I} - \hat{\boldsymbol{R}})\hat{\mathbb{B}}^* = (\boldsymbol{I} - \hat{\boldsymbol{R}})\boldsymbol{\mathcal{P}}_{(\mathcal{I})}\mathbb{B}^*, \tag{81}$$

with $p = \mathcal{O}(n_u) << n_g = \mathcal{O}(n)$. *It is interesting to see how the reduction in complexity of the RVE equilibrium equation is reflected in the gradual reduction of the dimensions of the "B" operators that act on the weighted vector of stresses.*

6.3.1. Physical interpretation

Aside from a "compressed" version of the original, full-order cell condition, the *hyperreduced* equation (77) can be alternatively interpreted as a balance between "observed" and "fitted" internal forces at the selected sampling points. Such an interpretation becomes readily identifiable by realizing that the product $\hat{\boldsymbol{R}}\hat{\boldsymbol{\Sigma}}$ appearing in Eq.(77) is but the (weighted) vector of *fitted* stresses at the selected sampling points. Indeed, by virtue of Eq.(72) and, considering the properties of the selection operator $\boldsymbol{\mathcal{P}}_{(\mathcal{I})}$, we have that

$$\hat{\boldsymbol{R}}\hat{\boldsymbol{\Sigma}} = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}(\boldsymbol{R}\hat{\boldsymbol{\Sigma}}) = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{\Sigma}^* = \hat{\boldsymbol{\Sigma}}^*. \tag{82}$$

Using the above equality, Eq.(77) is expressible as $\hat{\mathbb{B}}^{*^T}\hat{\boldsymbol{\Sigma}} = \hat{\mathbb{B}}^{*^T}\hat{\boldsymbol{\Sigma}}^*$, or, reverting to the original, summation notation as

$$\sum_{j \in \mathcal{I}} w_j \boldsymbol{B}^{*^T}(\mathbf{x}_j)\boldsymbol{\sigma}(\mathbf{x}_j;\cdot) = \sum_{j \in \mathcal{I}} w_j \boldsymbol{B}^{*^T}(\mathbf{x}_j)\boldsymbol{\sigma}^*(\mathbf{x}_j;\cdot). \tag{83}$$

Note that both sides of the above equation represent the same physical quantity, namely, the sum of internal forces, in reduced coordinates, at the sampling Gauss points $\{\mathbf{x}_{\mathcal{I}_1}, \mathbf{x}_{\mathcal{I}_2} \cdots \mathbf{x}_{\mathcal{I}_p}\}$. The difference lies in the stresses employed for computing these internal forces. In the left-hand side, they are calculated using "observed" stresses $\boldsymbol{\sigma}$ —stresses that arises directly from evaluating the corresponding constitutive equation—, whereas, in the right-hand side, "fitted" stresses $\boldsymbol{\sigma}^*$ are used —that is, stresses obtained from fitting the approximation constructed using the POD stress basis functions $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \ldots \boldsymbol{\Psi}_{n_\sigma}$ to the observed data. Thus, the HP-ROM equilibrium condition (83) is telling us that, at the solution, *the sum of internal forces —at the pre-selected sampling points— computed using either observed or fitted stresses*[14] *must coincide.*

[14]It should be mentioned in this respect that, in general, $\boldsymbol{\sigma}^*(\mathbf{x}_j;\cdot) \neq \boldsymbol{\sigma}(\mathbf{x}_j;\cdot)$ since the

6.4. Jacobian matrix

Needless to say, the dependence of the stresses on the reduced vector of reduced displacement fluctuations $\boldsymbol{U}^*$ is in general non-linear, and, thereby, an iterative method is required for solving Eq.(77). Here we employ the standard *Newton-Raphson* procedure. The iterative scheme corresponding to this procedure is given by the following expression (the parenthetical superscript indicates iteration number):

$$\boldsymbol{U}^{*(k+1)} = \boldsymbol{U}^{*(k)} - \boldsymbol{K}^{*(k)^{-1}} \boldsymbol{F}^{*(k)}, \tag{84}$$

where

$$\boldsymbol{F}^{*(k)} = \hat{\mathbb{B}}^{*^T} (\boldsymbol{I} - \hat{\boldsymbol{R}})\, \hat{\boldsymbol{\Sigma}}(\boldsymbol{\epsilon_M}, \boldsymbol{U}^{*(k)}) \tag{85}$$

and

$$\boldsymbol{K}^{*(k)} = \hat{\mathbb{B}}^{*^T} (\boldsymbol{I} - \hat{\boldsymbol{R}})\, \hat{\mathbb{C}}(\boldsymbol{\epsilon_M}, \boldsymbol{U}^{*(k)})\, \hat{\mathbb{B}}^*. \tag{86}$$

In the above equation, $\hat{\mathbb{C}} \in \mathbb{R}^{p\cdot s \times p\cdot s}$ denotes a block diagonal matrix containing the algorithmic, constitutive tangent matrices at each sampling point:

$$\hat{\mathbb{C}} := \begin{bmatrix} \boldsymbol{C}(\mathbf{x}_{\mathcal{I}_1}; \cdot) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C}(\mathbf{x}_{\mathcal{I}_2}; \cdot) & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{C}(\mathbf{x}_{\mathcal{I}_p}; \cdot) \end{bmatrix}. \tag{87}$$

6.4.1. Positive definiteness

Because of its relevance in the overall robustness of the proposed method, it is worthwhile at this point to digress and discuss thoroughly the spectral properties of the Jacobian matrix represented by Eq.(86). In particular, it would be interesting to ascertain whether *positive definiteness* of the algorithmic tangent matrices $\boldsymbol{C}(\mathbf{x}_{\mathcal{I}_1}; \cdot)$, $\boldsymbol{C}(\mathbf{x}_{\mathcal{I}_2}; \cdot), \cdots \boldsymbol{C}(\mathbf{x}_{\mathcal{I}_p})$ at the selected sampling points, and thus of matrix $\hat{\mathbb{C}}$, ensures positive definiteness of the Jacobian matrix $\boldsymbol{K}^*$ —as it occurs when using classical Gauss quadrature rules with positive weights—, and, if not, which remedies can be applied to obtain such desirable property.

number of data items to be fitted ($p \cdot s$) is always greater than the number of stress modes (n_σ). Observed and fitted stresses coincide only when the stress vector $\boldsymbol{\Sigma}$ one wishes to approximate pertains to the column space of the stress basis matrix $\boldsymbol{\Psi}$

Positive definiteness of the Jacobian matrix (86) requires that the function defined as

$$F(\mathbf{U}) = \mathbf{U}^{\mathbf{T}}\boldsymbol{K}^{*}\mathbf{U} = (\hat{\mathbb{B}}^{*}\mathbf{U})^{\mathbf{T}}(\boldsymbol{I} - \hat{\boldsymbol{R}})\,\hat{\mathbb{C}}(\hat{\mathbb{B}}^{*}\mathbf{U}) \tag{88}$$

be positive for all non-zero $\mathbf{U} \in \mathbb{R}^{\mathbf{n_u}}$. Since $\hat{\mathbb{B}}^{*}$ is a full rank matrix —by virtue of Eq.(64)—, condition $F(\mathbf{U}) > \mathbf{0}$ is equivalent to:

$$G(\mathbf{V}) = \mathbf{V}^{\mathbf{T}}(\boldsymbol{I} - \hat{\boldsymbol{R}})\,\hat{\mathbb{C}}\,\mathbf{V} > \mathbf{0} \tag{89}$$

for all non-zero $\mathbf{V} \in \text{Range}(\hat{\mathbb{B}}^{*})$.

To go further, we need to demonstrate that $\hat{\boldsymbol{R}} \in \mathbb{R}^{n_g \cdot s \times n_g \cdot s}$ —recall that $\hat{\boldsymbol{R}}$ is the matrix that maps the vector of "observed" stresses $\hat{\boldsymbol{\Sigma}}$ to the vector of *fitted* stresses $\hat{\boldsymbol{\Sigma}}^{*}$— actually represents an *orthogonal projection*[15] onto the column space of the gappy stress basis matrix $\hat{\boldsymbol{\Psi}}$. This can be shown by simply noting that $\hat{\boldsymbol{R}}$ is, on the one hand, symmetric:

$$\hat{\boldsymbol{R}}^{T} = (\hat{\boldsymbol{\Psi}}(\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^{T})^{T} = \hat{\boldsymbol{\Psi}}(\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-T}\hat{\boldsymbol{\Psi}}^{T} = \hat{\boldsymbol{R}} \tag{90}$$

and, on the other hand, idempotent:

$$\hat{\boldsymbol{R}}^{2} = (\hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger})^{2} = \hat{\boldsymbol{\Psi}}\overbrace{(\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}}}^{=\boldsymbol{I}}(\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^{T} = \hat{\boldsymbol{\Psi}}(\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^{T} = \hat{\boldsymbol{R}}. \tag{91}$$

With this property at hand, we can decompose any $\mathbf{V} \in \text{Range}(\hat{\mathbb{B}}^{*})$ as

$$\mathbf{V} = \mathbf{V}^{||} + \mathbf{V}^{\perp} \tag{92}$$

where $\mathbf{V}^{||} = \hat{\boldsymbol{R}}\mathbf{V} \in \text{Range}(\hat{\boldsymbol{\Psi}})$ —the component of $\mathbf{V}$ along the column space of $\hat{\boldsymbol{\Psi}}$— and $\mathbf{V}^{\perp} = (\boldsymbol{I} - \hat{\boldsymbol{R}})\mathbf{V}$ —the component of $\mathbf{V}$ along the orthogonal complement of $\text{Range}(\hat{\boldsymbol{\Psi}})$. Introducing the above decomposition into Eq.(89), we arrive at

$$G = \mathbf{V}^{\perp T}\hat{\mathbb{C}}\mathbf{V}^{\perp} + \mathbf{V}^{\perp\mathbf{T}}\,\hat{\mathbb{C}}\mathbf{V}^{||}. \tag{93}$$

While the first term $\mathbf{V}^{\perp T}\hat{\mathbb{C}}\mathbf{V}^{\perp}$ in the preceding equation is, in virtue of the positive definiteness of $\hat{\mathbb{C}}$, eminently positive for all nonzero $\mathbf{V}^{\perp} \in \mathbb{R}^{\mathbf{p}\cdot\mathbf{s}}$, nothing can be said in principle about the second term $\mathbf{V}^{\perp T}\,\hat{\mathbb{C}}\mathbf{V}^{||}$: numerical experience shows that the sign and relative magnitude of this term depends further on the chosen *set of sampling indices* $\mathcal{I}$.

[15] $\hat{\boldsymbol{R}}$ is the so-called "hat" matrix of linear regression models [52].

Remark 6.1. *From the above observation, it follows that the positive definiteness of the Jacobian matrix* $\boldsymbol{K}^*$ is determined*, not only by the spectral properties of* $\hat{\mathbb{C}}$*, but —not surprisingly— also by the number and the location within the RVE of the sampling points employed in the integration.*

The foregoing remark naturally leads to wonder whether it is possible to select the sampling indices $\mathcal{I}$ so as to ensure the positive definiteness of $\boldsymbol{K}^*$ (assuming, obviously, that $\hat{\mathbb{C}}$ enjoys this property). To shed light on this question, let us first divide Eq.(93) by $\mathbf{V}^{\perp T}\hat{\mathbb{C}}\mathbf{V}^{\perp}$ (notice that hypothesis (64) precludes the possibility of $\mathbf{V}^{\perp}$ being zero)

$$\bar{G} = \frac{G}{\mathbf{V}^{\perp T}\hat{\mathbb{C}}\mathbf{V}^{\perp}} = 1 + \frac{\mathbf{V}^{\perp T}\hat{\mathbb{C}}\mathbf{V}^{||}}{\mathbf{V}^{\perp T}\hat{\mathbb{C}}\mathbf{V}^{\perp}}. \tag{94}$$

Suppose now, for the sake of argument, that $\hat{\mathbb{C}}$ is also symmetric. Such being the case, the above equation can be legitimately rewritten as:

$$\bar{G} = 1 + \cos(\mathbf{V}^{\perp}, \mathbf{V}^{||})_{\hat{\mathbb{C}}} \frac{\|\mathbf{V}^{||}\|_{\hat{\mathbb{C}}}}{\|\mathbf{V}^{\perp}\|_{\hat{\mathbb{C}}}}, \tag{95}$$

where

$$\cos(\mathbf{V}^{\perp}, \mathbf{V}^{||})_{\hat{\mathbb{C}}} = \frac{\langle \mathbf{V}^{\perp}, \mathbf{V}^{||} \rangle_{\hat{\mathbb{C}}}}{\|\mathbf{V}^{\perp}\|_{\hat{\mathbb{C}}} \|\mathbf{V}^{||}\|_{\hat{\mathbb{C}}}}. \tag{96}$$

In the above equation, $\langle \cdot, \cdot \rangle_{\hat{\mathbb{C}}}$ symbolizes the inner product defined by $\hat{\mathbb{C}}$ (i.e., $\langle \mathbf{x}, \mathbf{y} \rangle_{\hat{\mathbb{C}}} = \mathbf{x}^{\mathbf{T}}\hat{\mathbb{C}}\mathbf{y}$), whereas $\| \cdot \|_{\hat{\mathbb{C}}}$ denotes the norm associated to such an inner product ($\|\mathbf{x}\|^{\mathbf{2}}_{\hat{\mathbb{C}}} = \langle \mathbf{x}, \mathbf{x} \rangle_{\hat{\mathbb{C}}}$). From Eq.(94), it can be deduced that a sufficient (yet not necessary) condition for $\bar{G} > 0$, and thus for $\boldsymbol{K}^*$ to be positive definite, is that

$$\frac{\|\mathbf{V}^{||}\|_{\hat{\mathbb{C}}}}{\|\mathbf{V}^{\perp}\|_{\hat{\mathbb{C}}}} < 1 \tag{97}$$

for all nonzero $\mathbf{V} \in \text{Range}(\hat{\mathbb{B}}^*)$, or equivalently (setting $\mathbf{V} = \hat{\mathbb{B}}^*\mathbf{U}$):

$$\frac{\|\hat{\boldsymbol{R}}\hat{\mathbb{B}}^*\mathbf{U}\|_{\hat{\mathbb{C}}}}{\|(\boldsymbol{I} - \hat{\boldsymbol{R}})\hat{\mathbb{B}}^*\mathbf{U}\|_{\hat{\mathbb{C}}}} < 1 \tag{98}$$

for all nonzero $\mathbf{U} \in \mathbb{R}^{\mathbf{n_u}}$.

Useful guidelines on how to choose $\mathcal{I}$ so as to make positive definite the Jacobian matrix $\boldsymbol{K}^*$ can be inferred from inequality (98). Firstly, given a fixed number of sampling points p, expression (98) indicates that such points should be selected so that the columns of the gappy *strain* basis matrix $\hat{\mathbb{B}}^* = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\mathbb{B}^* = [\hat{\mathbb{B}}_1^* \, \hat{\mathbb{B}}_2^* \ldots \hat{\mathbb{B}}_{n_u}^*]$ are, loosely speaking, "as orthogonal as possible" to $\text{Range}(\hat{\boldsymbol{R}}) = \text{Range}(\hat{\boldsymbol{\Psi}})$ —the column space of the gappy *stress* basis matrix $\hat{\boldsymbol{\Psi}} = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{\Psi}$. In so doing, the factor defined as

$$f_{\hat{\mathbb{C}}} := \frac{\sqrt{\sum_{i=1}^{n_u} \|\hat{\boldsymbol{R}}\hat{\mathbb{B}}_i^*\|_{\hat{\mathbb{C}}}^2}}{\sqrt{\sum_{i=1}^{n_u} \|(\boldsymbol{I} - \hat{\boldsymbol{R}})\hat{\mathbb{B}}_i^*\|_{\hat{\mathbb{C}}}^2}}, \tag{99}$$

would diminish, and so would, consequently, the left-hand side of inequality Eq.(98). In practice, however, factor $f_{\hat{\mathbb{C}}}$ cannot be used as a criterion for guiding the selection of sampling points, simply because it is defined in terms of the norm induced by $\hat{\mathbb{C}}$, and this matrix virtually changes at every time step and iteration. One has to be content to estimate this factor using other norm; for instance, employing the standard euclidean norm $\|\cdot\|$, one gets

$$f_{\hat{\mathbb{C}}} \sim f_F := \frac{\sqrt{\sum_{i=1}^{n_u} \|\hat{\boldsymbol{R}}\hat{\mathbb{B}}_i^*\|^2}}{\sqrt{\sum_{i=1}^{n_u} \|(\boldsymbol{I} - \hat{\boldsymbol{R}})\hat{\mathbb{B}}_i^*\|^2}} = \frac{\|\hat{\boldsymbol{R}}\hat{\mathbb{B}}^*\|_F}{\|(\boldsymbol{I} - \hat{\boldsymbol{R}})\hat{\mathbb{B}}^*\|_F}, \tag{100}$$

where $\|\cdot\|_F$ stands for the Frobenius norm.

Aside from seeking orthogonality between $\hat{\mathbb{B}}^*$ and $\hat{\boldsymbol{R}}$, expression (99) suggests that another way of lowering factor f_F may be to reduce the ratio defined as

$$\beta_F := \frac{\|\hat{\boldsymbol{R}}\|_F}{\|\boldsymbol{I} - \hat{\boldsymbol{R}}\|_F}. \tag{101}$$

Since $\hat{\boldsymbol{R}}$ and, consequently, $\boldsymbol{I} - \hat{\boldsymbol{R}}$ are matrices representing orthogonal projections, we have that $\|\hat{\boldsymbol{R}}\|_F = \sqrt{\text{rank}(\hat{\boldsymbol{R}})} = \sqrt{n_\sigma}$ and $\|\boldsymbol{I} - \hat{\boldsymbol{R}}\|_F = \sqrt{p \cdot s - n_\sigma}$. Therefore,

$$\beta_F = \sqrt{\frac{n_\sigma}{p \cdot s - n_\sigma}}. \tag{102}$$

Observation 6.2. *From the above expression, thus, one can conclude that increasing the number of sampling points p while keeping the number of stress modes n_σ constant also contributes to reduce factor f_F in Eq.(98), and, hence, to improve the spectral properties (positive defineteness) of the Jacobian matrix $\boldsymbol{K}^*$. Notice that this property is totally consistent with the fact that, in the limiting case of taking all Gauss points ($p = n_g$), the reduced matrices $\hat{\boldsymbol{R}}$ and $\hat{\mathbb{B}}^*$ degenerate into their full order counterparts $\boldsymbol{R}$ and $\mathbb{B}^*$, for which the condition $\boldsymbol{R}\,\mathbb{B}^* = \boldsymbol{0}$ holds —they span subspaces that are mutually orthogonal—, hence making $f_F = f_{\hat{\mathbb{C}}} = 0$.*

7. Selection of sampling points

The last theoretical issue to be discussed in the present work is the selection —among the full set of Gauss points of the underlying finite element mesh— of appropriate *sampling or interpolation points*. At the very least, the set of sampling indices $\mathcal{I} = \{i_1, i_2 \dots i_p\}$ must be chosen so that the *gappy* expanded basis matrix has full rank (see section 6.2):

$$\text{rank}(\hat{\boldsymbol{\Psi}}^{ex}_{(\mathcal{I})}) = \text{rank}([\hat{\boldsymbol{\Psi}}_{(\mathcal{I})} \;\; \hat{\mathbb{B}}^*_{(\mathcal{I})}]) = n_\sigma + n_u. \tag{103}$$

Any set of sampling indices fulfilling this necessary condition is said to be *admissible*.

7.1. Optimality criteria

7.1.1. Accuracy

As in any other model reduction problem, the overriding concern when choosing the sampling points is the *accuracy of the approximation*: we would like to position such points so that maximum similarity between the "high-fidelity", finite element solution and its reduced-order counterpart is obtained. More specifically, since the output of interest in our BVP is the *macroscopic* stress response, the aim is to find the set of sampling points $\mathcal{I}$ that minimizes the following error estimate:

$$E_{M,\sigma}(\boldsymbol{\Psi}, \mathcal{I}) := \sqrt{\sum_{i=1}^{n_{snp}} \|\boldsymbol{\sigma_M}^i - \boldsymbol{\sigma_M}^{*\,i}(\boldsymbol{\Psi}, \mathcal{I})\|^2}, \tag{104}$$

where $\boldsymbol{\sigma_M}^i = \boldsymbol{\sigma_M}({}^t\boldsymbol{\epsilon_M}_k{}^j)$ denotes the finite element, macroscopic stress response corresponding to the the $k-th$ ($k = 1, 2 \dots n_{stp}$) time step of the

"training"[16] strain trajectory ${}^t\boldsymbol{\epsilon}_{\boldsymbol{M}}{}^j$ ($j = 1, 2 \dots n_{hst}$); and $\boldsymbol{\sigma}_{\boldsymbol{M}}{}^{*i}(\boldsymbol{\Psi}, \mathcal{I})$ its low-dimensional approximation.

Using the Cauchy-Schwarz inequality, and approximating the pertinent integrals by Gauss quadrature, we can write

$$\begin{aligned} E_{M,\sigma}{}^2 &\leq \frac{1}{V} \sum_{i=1}^{n_{snp}} \int_\Omega \|\boldsymbol{\sigma}^i(\mathbf{x}) - \boldsymbol{\sigma}^{*i}(\mathbf{x}; \boldsymbol{\Psi}, \mathcal{I})\|^2 \, d\Omega \\ &\approx \frac{1}{V} \sum_{i=1}^{n_{snp}} \sum_{g=1}^{n_g} \|\sqrt{w_g}\boldsymbol{\sigma}^i(\mathbf{x}_g) - \sqrt{w_g}\boldsymbol{\sigma}^{*i}(\mathbf{x}_g; \boldsymbol{\Psi}, \mathcal{I})\|^2 \\ &= \frac{1}{V} \sum_{i=1}^{n_{snp}} \|\boldsymbol{\Sigma}^i - \boldsymbol{\Sigma}^{*i}(\boldsymbol{\Psi}, \mathcal{I})\|^2 2 = \frac{1}{V} \|\boldsymbol{X} - \boldsymbol{X}^*(\boldsymbol{\Psi}, \mathcal{I})\|_F^2 \end{aligned} \tag{105}$$

where $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{\Sigma}^1 & \boldsymbol{\Sigma}^2 & \cdots & \boldsymbol{\Sigma}^{n_{snp}} \end{bmatrix}$ and $\boldsymbol{X}^*(\boldsymbol{\Psi}, \mathcal{I}) = \boldsymbol{R}_{(\mathcal{I})}(\boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{X})$. The error estimate for the macroscopic stresses defined in Eq.(104) is, thus, bounded above by the Frobenius norm of the difference between the (weighted) stress snapshot matrix $\boldsymbol{X}$ and its oblique projection, $\boldsymbol{X}^*$, onto Range($\boldsymbol{\Psi}$). This bound, hereafter designated by e_σ, admits the following decomposition[17]

$$e_\sigma{}^2 = \frac{1}{V} \|\boldsymbol{X} - \boldsymbol{X}^{*\perp}(\boldsymbol{\Psi})\|_F^2 + \frac{1}{V} \|\boldsymbol{X}^{*\perp}(\boldsymbol{\Psi}) - \boldsymbol{X}^*(\boldsymbol{\Psi}, \mathcal{I})\|_F^2, \tag{106}$$

$\boldsymbol{X}^{*\perp}$ being the *orthogonal* projection of $\boldsymbol{X}$ onto the range of $\boldsymbol{\Psi}$, i.e., $\boldsymbol{X}^{*\perp} = \boldsymbol{\Psi}\boldsymbol{\Psi}^T\boldsymbol{X}$. Note that the first term of the right-hand side of Eq.(106) only depends on the stress basis matrix, but not on the employed sampling indices; it provides, thus, an estimate of the *stress truncation error*. The term that actually measures the quality, in terms of accuracy, of a given set of admissible sampling points is the second one—it provides an (a priori) estimate of the *stress reconstruction error*. We shall denote this term by e_σ^{rec}:

$$e_\sigma^{rec} := \frac{1}{V} \|\boldsymbol{X}^{*\perp}(\boldsymbol{\Psi}) - \boldsymbol{X}^*(\boldsymbol{\Psi}, \mathcal{I})\|_F. \tag{107}$$

For this reason —and also because the cost of evaluating expression Eq.(107)

[16] The term "training", which, incidentally, is borrowed from the neural network literature [38], is used throughout the text to refer to the offline generation of snapshots.

[17] This decomposition follows easily from the Pythagorean theorem (just notice that $\boldsymbol{\Psi}^T(\boldsymbol{X} - \boldsymbol{X}^{*\perp}) = \boldsymbol{0}$).

is independent of the number of Gauss points[18], and therefore significantly lower than in the case of the original error estimate $E_{M,\sigma}$—, we shall use in what follows e_σ^{rec} as error estimator for guiding the selection of sampling points.

7.1.2. Spectral properties

Yet the optimality of a given set of sampling points cannot be measured only in terms of accuracy of the approximation. As demonstrated in Section 6.4, the number and particular placement of such points influence also the *spectral properties* (positive definiteness) of the Jacobian matrix of the equilibrium equation, and therefore, the convergence characteristics of the accompanying Newton-Raphson algorithm. We saw that, to preserve the positive definiteness of the full-order Jacobian matrix, one should strive to choose the sampling indices $\mathcal{I}$ so as to make the factor —defined previously in Eq.(100)—:

$$f_F(\boldsymbol{\Psi}, \mathbb{B}^*, \mathcal{I}) = \frac{\|\hat{\boldsymbol{R}}_{(\mathcal{I})}\hat{\mathbb{B}}^*_{(\mathcal{I})}\|_F}{\|(\boldsymbol{I} - \hat{\boldsymbol{R}}_{(\mathcal{I})})\hat{\mathbb{B}}^*_{(\mathcal{I})}\|_F} \tag{108}$$

as small as possible.

7.2. Optimization approach: basic and stabilizing sampling points

Unfortunately, the minimization of the approximation error represented by expression Eq.(107) and the minimization of Eq.(108) are in general conflicting goals. For instance, numerical experiments show that when the selection is driven exclusively by accuracy considerations, the resulting Jacobian matrix becomes indefinite at certain states of deformation —especially when inelastic deformations are severe—, leading occasionally to convergence failures. These goals must be therefore balanced in order to arrive at an accurate and at the same time robust solution scheme.

To accomodate these conflicting requirements, we propose here a heuristic strategy that basically consists in treating the minimization of Eq.(107) and

[18]Indeed, since $\boldsymbol{\Psi}$ is column-wise orthogonal, minimizing Eq.(107) is equivalent to minimizing the difference between the coefficients of the respective projections. The number of coefficients is equal to the number of snapshots, which is normally much lower than the number of Gauss points.

Eq.(108) as two separated, sequential problems —in the spirit of the so-called "greedy" optimization algorithms[19] [44]. The set of sampling points is assumed to be divided into two disjoint subsets $\mathcal{I}_\sigma$ and $\mathcal{I}_B$:

$$\mathcal{I} = \mathcal{I}_\sigma \cup \mathcal{I}_B. \tag{109}$$

The first subset $\mathcal{I}_\sigma = \{i_1, i_2, \cdots i_{p_\sigma}\}$ is obtained as the minimizer of the error estimation given in Eq.(107), viz.:

$$\mathcal{I}_\sigma = \arg \min_{\mathcal{K} \subseteq \{1,2 \ldots n_g\}} e_\sigma^{rec}(\boldsymbol{\Psi}, \mathcal{K}). \tag{110}$$

Once the set $\mathcal{I}_\sigma$ is determined, the remaining sampling indices $\mathcal{I}_B = \{j_1, j_2 \cdots j_{p_B}\}$ ($p_\sigma + p_B = p$) are calculated as

$$\mathcal{I}_B = \arg \min_{\mathcal{K} \subseteq \{1,2 \ldots n_g\}} f_F(\boldsymbol{\Psi}, \mathbb{B}^*, \mathcal{I}_\sigma \cup \mathcal{K}). \tag{111}$$

Remark 7.1. *It must be noted here that the minimization problem represented by Eq.(110) is in essence the same problem addressed in (standard) interpolatory-based, model reduction approaches for determining, given a set of empirical basis functions, the optimal location of associated interpolations points. For this reason, we shall refer to the set of points arising from solving this minimization problem as the* standard or basic sampling points *—these are the Best Interpolation Points of Nguyen et al. [54], or the "magic points" of Maday et al. [46].. By contrast, the necessity of introducing points that attempt to solve problem (111) is a consequence of expanding the stress approximation space in the first place —the main innovative feature of our approach—, and it is therefore not present in other model reduction strategies. We shall call* $\{\boldsymbol{x}_{\mathcal{I}_B(1)}, \boldsymbol{x}_{\mathcal{I}_B(2)} \ldots \boldsymbol{x}_{\mathcal{I}_B(p_B)}\}$ *the set of stabilizing sampling points.*

The number of *basic sampling points* must satisfy the necessary condition $p_\sigma \geq n_\sigma/s$. In general, taking $p_\sigma = n_\sigma$ suffices to ensure highly satisfactory approximations. How many, on the other hand, *stabilizing sampling points* have to be added to safely render positive definite the Jacobian matrix —for at least a representative range of macroscopic state deformations— is a

[19] A greedy method is any algorithm that solves the problem by making the locally optimal choice at each step with the hope of finding the global optimum.

question that can only be answered empirically. In the examples presented in the next section, it has been found that a conservative answer is to use as many stabilizing sampling points as displacement basis modes ($p_B = n_u$).

To deal with the discrete minimization problem (110), we have used here the *Hierarchical Interpolation Points* (HPI) method proposed by Nguyen et al. [54]. The essence of this method is to construct, in a *greedy* fashion, the set of indices by solving a *sequence* of one-dimensional minimization problems. The minimization problem (111) is also addressed using a heuristic based on the greedy paradigm. In particular, the $k-th$ ($k = 1, 2 \dots p_B$) index is selected by solving the following, one-dimensional minimization problem:

$$\mathcal{I}_B(k) = \arg \min_{j \in \{1,2 \dots n_g\}} f_F(\boldsymbol{\Psi}, \mathbb{B}^*, \mathcal{K}(j)) \tag{112a}$$

$$\mathcal{K} = \mathcal{I}_\sigma \cup \{\mathcal{I}_B(1), \mathcal{I}_B(2) \cdots \mathcal{I}_B(k-1), j\}. \tag{112b}$$

8. Summary

Lastly, for the reader's convenience and easy reference, the online reduced-order problem, along with the offline steps that leads to the the hyperreduced operators appearing in the online problem, are summarized in Boxes 8.1 and 8.2.

9. Numerical results

This section is intended to illustrate the performance and assess the efficiency of the proposed model reduction strategy in solving the fine scale BVP corresponding to a *porous metal material* under plane strain conditions.

9.1. Microstructure description

The voids are elliptical in shape (with eccentricity equal to 0.3), randomly distributed (with porosity equal to 0.3), and have aligned major axes ranging in length —according to the cumulative probability distribution displayed in figure 3.b— from 0.2 to 1.5 mm. The mechanical behavior of the metal matrix is modeled by a rate-independent, Von Mises elastoplastic model endowed with the following non-linear, isotropic hardening saturation law (consult Ref. [61] for details on the implementation of this elastoplastic model):

$$\sigma_u(\alpha) = \sigma_0 + \bar{H}\alpha + (\sigma_\infty - \sigma_0)(1 - exp(-\alpha\delta)). \tag{118}$$

1. Compute FE displacement fluctuations and stress snaphots for representative, input macro-strain histories. Apply —see Appendix B— the elastic/inelastic POD to the resulting snapshot matrices to obtain the displacement fluctuation and stress basis matrices ($\boldsymbol{\Phi} \in \mathbb{R}^{n\cdot d\times n_u}$ and $\boldsymbol{\Psi} \in \mathbb{R}^{n_g\cdot s\times n_\sigma}$, respectively).
2. Calculate the weighted matrix of fluctuating strain modes $\mathbb{B}^* \in \mathbb{R}^{n_g\cdot s\times n_\sigma}$ using Eqs. (26) and (47).
3. Select a set $\mathcal{I}$ of sampling indices optimal for the basis matrices $\boldsymbol{\Psi}$ and $\mathbb{B}^*$ following the procedure sketched in Section 7.
4. Finally, using $\boldsymbol{\Psi}$, $\mathbb{B}^*$ and $\mathcal{I}$, construct the *hyperreduced*-order matrices $\hat{\mathbb{B}}^{**} \in \mathbb{R}^{p\cdot s\times n_u}$ and $\mathbb{T} \in \mathbb{R}^{s\times p\cdot s}$; the expressions for these matrices read:

$$\hat{\mathbb{B}}^{**} = (\boldsymbol{I} - \hat{\boldsymbol{R}})\hat{\mathbb{B}}^* = (\boldsymbol{I} - \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{R})(\boldsymbol{\mathcal{P}}_{(\mathcal{I})}\mathbb{B}^*) \tag{113}$$

and

$$\mathbb{T} := \frac{1}{V}\left[\sqrt{w_1}\boldsymbol{I} \quad \sqrt{w_2}\boldsymbol{I} \quad \cdots \quad \sqrt{w_{n_g}}\boldsymbol{I}\right]\boldsymbol{R} \tag{114}$$

where $\boldsymbol{R} = \boldsymbol{\Psi}(\hat{\boldsymbol{\Psi}}^T\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^T$ and $\hat{\boldsymbol{\Psi}} = \boldsymbol{\mathcal{P}}_{(\mathcal{I})}\boldsymbol{\Psi}$.

Box 8.1: Offline stage. Pre-computation of reduced basis and hyperreduced operators.

Here, σ_u stands for the yield stress, $\alpha \geq 0$ denotes the equivalent plastic strain; and $\sigma_0 = 75.0\ MPa$, $\sigma_\infty = 100.0\ MPa$, $\delta = 2500.0$ and $\bar{H} = 5000\ MPa$ are material constants. The Young's modulus and Poisson's coefficient, on the other hand, are equal to $E_m = 75\,GPa$ and $\nu_m = 0.3$, respectively (these material constants corresponds approximately to Aluminum).

9.2. RVE and finite element discretization

The size of the RVE was determined by conducting finite element analyses on square domains of increasing size subject to vanishing displacement fluctuations boundary conditions. It was found that the macroscopic stress responses calculated under representative macroscopic strain paths (stretching along the longitudinal and transversal directions, and shearing) of all samples above 20x20 mm^2 were practically indistinguishable. This fact indicates that any subvolume of 20x20 mm^2 (or greater) can be considered as a Representative Volume Element (RVE) of the porous material under study.

1. *Initial data*: $\boldsymbol{U}_n^* \in \mathbb{R}^{n_u}$ (reduced vector of displacement fluctuations at t_n), $\boldsymbol{\epsilon}_{\boldsymbol{M}n}$ (macroscopic strain vector at t_n), and $\{\boldsymbol{\xi}_n(\mathbf{x}_{\mathcal{I}_1}), \boldsymbol{\xi}_n(\mathbf{x}_{\mathcal{I}_2}), \ldots \boldsymbol{\xi}_n(\mathbf{x}_{\mathcal{I}_p})\}$ (internal variables at t_n *at the selected sampling points*).
2. *Input data*: $\boldsymbol{\epsilon}_{\boldsymbol{M}n+1}$ (macroscopic strain vector at t_{n+1})
3. Given the above initial and input data, find $\boldsymbol{U}_{n+1}^* \in \mathbb{R}^{n_u}$ such that

$$\hat{\mathbb{B}}^{**^T} \hat{\boldsymbol{\Sigma}}(\boldsymbol{\epsilon}_{\boldsymbol{M}n+1}, \boldsymbol{U}_{n+1}^*) = \mathbf{0}, \tag{115}$$

   where

$$\hat{\boldsymbol{\Sigma}} = \left[\sqrt{w_{\mathcal{I}_1}}\boldsymbol{\sigma}^T(\mathbf{x}_{\mathcal{I}_1}, \cdot) \quad \sqrt{w_{\mathcal{I}_2}}\boldsymbol{\sigma}^T(\mathbf{x}_{\mathcal{I}_2}, \cdot) \quad \cdots \quad w_{\mathcal{I}_p}\boldsymbol{\sigma}^T(\mathbf{x}_{\mathcal{I}_p}, \cdot)\right]^T \tag{116}$$

   (here, $\boldsymbol{\sigma}(\mathbf{x}_{\mathcal{I}_i}, \cdot)$ denotes the stress vector evaluated at the $i-th$ sampling point through the corresponding constitutive equation).
4. *Output data:* Once Eq.(116) has been solved for $\boldsymbol{U}_{n+1}^*$, update the macroscopic stress vector as

$$\boldsymbol{\sigma}_{\boldsymbol{M}}|_{n+1} = \mathbb{T}\, \hat{\boldsymbol{\Sigma}}(\boldsymbol{\epsilon}_{\boldsymbol{M}n+1}, \boldsymbol{U}_{n+1}^*). \tag{117}$$

Box 8.2: Online stage (solution of the hyperreduced-order RVE equilibrium problem for given macroscopic strains).

The finite element discretization corresponding to the particular 20x20 mm^2 RVE employed in the ensuing simulations is shown in figure 3.a. The number of (four-node bilinear) elements is $n_e = 9746$, and the number of nodes $n = 11825$. The employed quadrature formula, on the other hand, is the standard 2x2 Gauss rule, the total number of Gauss points amounting thus to $n_g = 4\, n_e = 38984$. To overcome incompressibility issues while maintaining the displacement-based formulation presented in the preceding sections, the commonly known as "B-bar" approach is adopted[20]. The constitutive differential equations are integrated in time using the classical (fully

[20]This means that, in this case, the reduced "B-matrix" $\boldsymbol{B}^*(\mathbf{x})$ appearing in the formulation of the HP-ROM is not constructed using the gradients of the shape functions, as indicated by Eq.(23), but rather using the modified "B-matrix" emanating from the three-field Hu-Washizu variational principle [61]

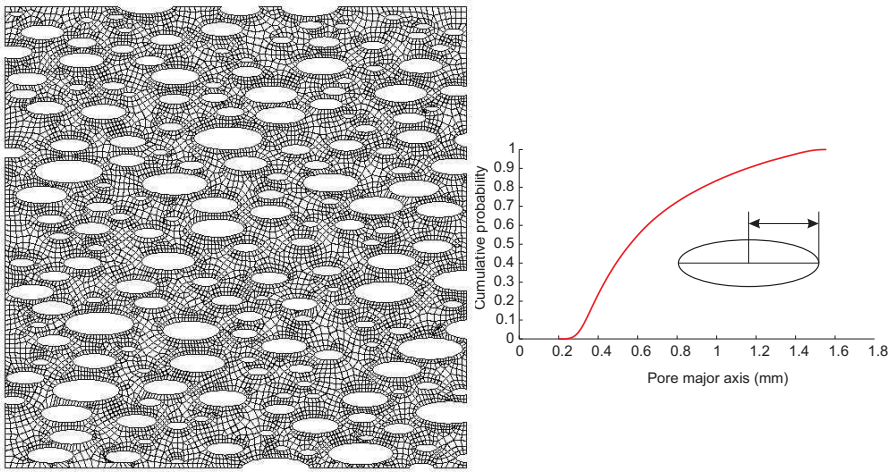


Figure 3: a) Finite element mesh of the RVE corresponding to the porous metal material. b) Cumulative probability distribution followed by the length of the pore major axes.

implicit) backward-Euler scheme.

9.3. Sampling of parameter space

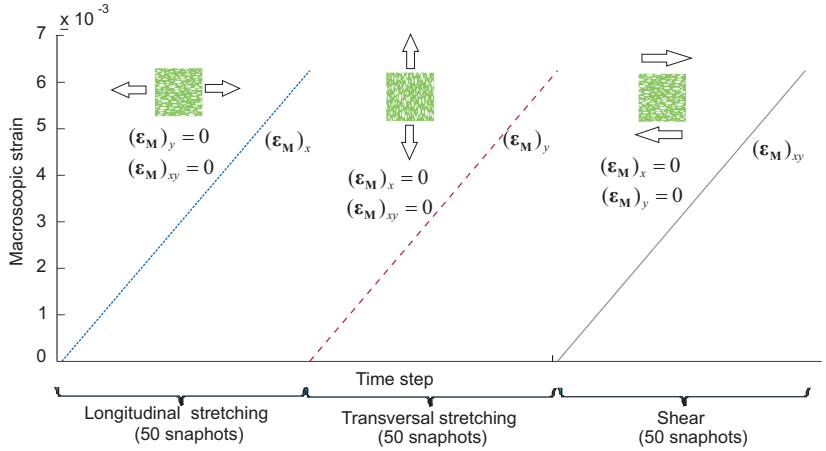


Figure 4: Macro-strain trajectories used for generating the displacement and stress snapshots.

The first step in the process of constructing the reduced basis is the *sampling of the input parameter space*; we saw in Section 3.1 that, in the fine-scale BVP, this process amounts to select *representative macroscopic strain histories*. The three macroscopic strain histories ($n_{hst} = 3$) used in the case under study are depicted in figure 4. In each of these strain trajectories, one of the (independent) strain components follows a linear ascending path while the magnitude of the other two components is set to zero. The time domain for each strain history is discretized into $n_{stp} = 50$ equally spaced steps, resulting in a total number of $n_{snp} = n_{hst} \cdot n_{stp} = 150$ snapshots.

Remark 9.1. *The task of sampling the input parameter space is somehow akin to the experimental process whereby material parameters of standard phenomenological models are calibrated in a laboratory. In this analogy, the RVE plays the role of the corresponding experimental specimen, whereas the macro-strain training trajectories represent the loading paths of the pertinent calibration tests. Hindsight and elementary physical considerations can therefore aid in restricting the number of strain histories (and hence of snapshots) necessary to characterize the response. For instance, if the behavior of the materials that compose the RVE is governed by rate-independent constitutive models (as in the case at hand), we know beforehand that it is not necessary to study the response under varying rates of deformation.*

Strategies for efficiently sampling the input parameter space *in general model reduction contexts can be found in Refs. [12, 13, 15, 42].*

9.4. Dimensionality reduction: a priori error analysis

The finite element displacement fluctuation and stress fields computed at each time step of the input strain trajectories shown above are multiplied by their corresponding weighting matrices ($\bar{\mathbf{M}}$ and $\boldsymbol{W}^{1/2}$) and stored, in the snapshot matrices $\bar{\boldsymbol{X}}_u \in \mathbb{R}^{n \cdot d \times n_{snp}}$ ($n \cdot d = 11825 \cdot 2 = 23650$) and $\boldsymbol{X} \in \mathbb{R}^{n_g \cdot s \times n_{snp}}$ ($n_g \cdot s = 38984 \cdot 4 = 155936$), respectively. Then, these matrices are subjected to the SVD-based, elastic/inelastic dimensionality reduction process sketched in Section 3.2.1 —and described more in detail in Appendix B— in order to generate an optimal set of basis vectors for both the displacements fluctuation and stress solution spaces.

To elucidate which of these basis vectors constitute the "essential" modes of the response, we plot in Figure 5 the dimensionless POD truncation error estimates defined, for the displacement fluctuations, as:

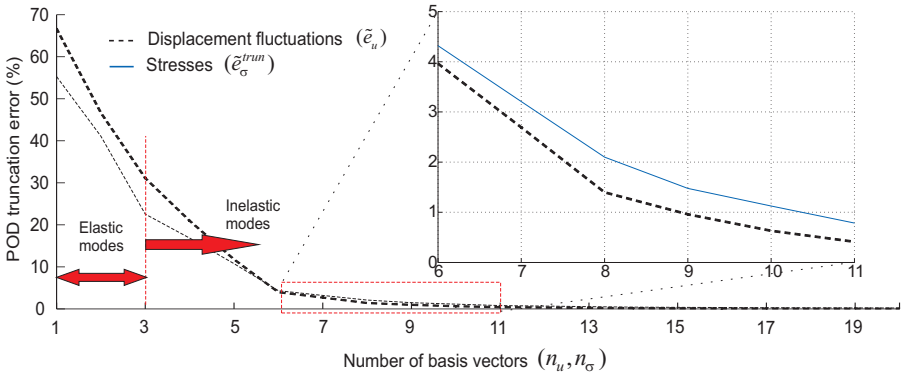


Figure 5: POD truncation error estimates $\tilde{e}_u$ (for the displacement fluctuations, see Eq.(119)) and $\tilde{e}_\sigma^{trun}$ (for the stresses, see Eq.(120)) versus number of basis vectors employed in the approximation (n_u and n_σ, respectively). The portion between 6 and 11 modes is shown in magnified form.

$$\tilde{e}_u(n_u) := \frac{\|\bar{\boldsymbol{X}}_u - \bar{\boldsymbol{X}}_u^*(n_u)\|_F}{\|\bar{\boldsymbol{X}}_u\|_F}, \tag{119}$$

and for the stresses:

$$\tilde{e}_\sigma^{trun}(n_\sigma) := \frac{\|\boldsymbol{X} - \boldsymbol{X}^{*\perp}(n_\sigma)\|_F}{\|\boldsymbol{X}\|_F}, \tag{120}$$

$\bar{\boldsymbol{X}}_u^*(n_u)$ and $\boldsymbol{X}^{*\perp}(n_\sigma)$ being the orthogonal projection of $\bar{\boldsymbol{X}}_u$ and $\boldsymbol{X}$ onto the span of the first n_u and n_σ basis vectors, respectively. It can be observed in Figure 5 that both error measures decrease monotonically with increasing order of truncation —this is a mere consequence of the optimality properties of the SVD—, and at approximately the same rate; the decay is more pronounced from 1 to 6 modes, and becomes more gradual thereafter, tending asymptotically to zero as the number of modes increases. The truncation error for both stresses and displacement fluctuations at $n_\sigma = n_u = 6$ is around 5%. In terms of dimensionality reduction, this means that the data contained in the snapshot matrices can be "compressed" to a factor of $(n_u/n_{snp}) \cdot 100 = (6/150) \cdot 100 = 4\%$ and still retain 95% of the information —the *essential* information. The first 6 basis functions (3 elastic and 3 inelastic) for both stresses and displacement fluctuations, therefore, are to be regarded as *essential* modes in the characterization of the mechanical

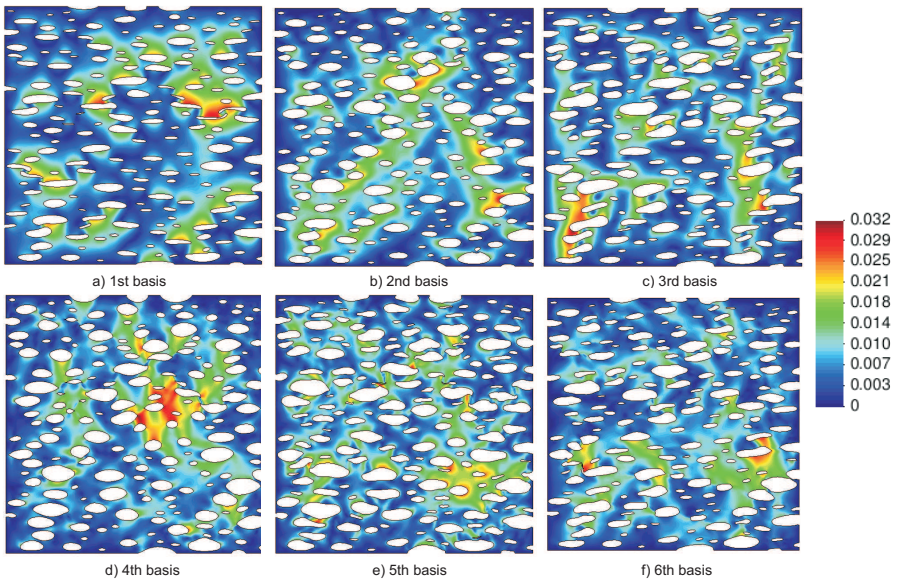


Figure 6: Contour plots of the euclidean norm of the first 6 displacement fluctuations modes ($\|\boldsymbol{\Phi}_i\|$, $i = 1, 2 \dots 5$). Deformed shapes are scaled up by a factor of 15.

response of the concerned RVE. By way of illustration, we plot in Figure 6 the contour plots of the euclidean norm of such 6 *essential* displacement fluctuations modes ($\|\boldsymbol{\Phi}_i\|$, $i = 1, 2 \dots 6$).

9.5. Sampling points

9.5.1. Basic sampling points

Once the stress and displacement fluctuation basis vectors have been determined, the next *offline* step consists in the selection —among the full set of finite element Gauss points—of an optimal set of *sampling points*. Following the strategy described in Section 7.2, we carry out such a selection by first computing the location of what we have called *basic sampling points* $\{\mathbf{x}_{\mathcal{I}_\sigma(1)}, \mathbf{x}_{\mathcal{I}_\sigma(2)} \dots \mathbf{x}_{\mathcal{I}_\sigma(p_\sigma)}\}$. To assess the efficiency of the employed Hierarchical Interpolation Points Method, abbreviated HIPM, we plot in Figure 7 the estimates for both the POD truncation (shown previously in Figure 5) and *total stress error* versus the number of stress modes n_σ (in using this algorithm, it is assumed that $p_\sigma = n_\sigma$). The total stress error estimate is

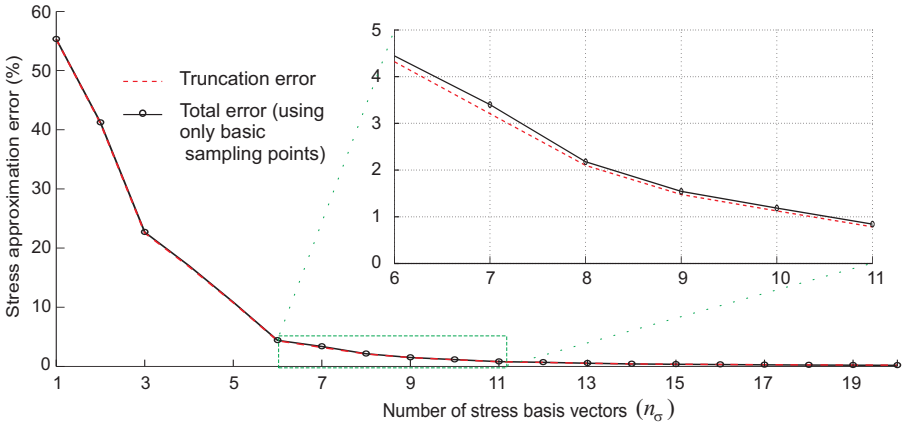


Figure 7: Estimates for the POD truncation ($\tilde{e}_\sigma^{trun}$, see Eq.(120)) and total ($\tilde{e}_\sigma$, see Eq.(121)) stress error versus number of basis vectors employed in the approximation (n_σ). The total error estimate is computed using only the set of *basic sampling points* ($\tilde{e}_\sigma = \tilde{e}_\sigma(n_\sigma, \mathcal{I}_\sigma)$, with $p_\sigma = n_\sigma$). The portion between 6 and 11 modes is shown in magnified form.

defined as

$$\tilde{e}_\sigma(n_\sigma, \mathcal{I}_\sigma) := \frac{\|\boldsymbol{X} - \boldsymbol{X}^*(n_\sigma, \mathcal{I}_\sigma)\|_F}{\|\boldsymbol{X}\|_F} \tag{121}$$

where $\boldsymbol{X}^*(n_\sigma, \mathcal{I}_\sigma)$ denotes the *oblique* projection (calculated using sampling points $\mathcal{I}_\sigma$) of $\boldsymbol{X}$ onto the span of the first n_σ basis vectors $(\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \ldots \boldsymbol{\Psi}_{n_\sigma})$. It can be appreciated in Figure 7 that both the total error and the truncation error curves are practically coincident, a fact that indicates that the contribution of the *reconstruction error*:

$$\tilde{e}_\sigma^{rec} = \sqrt{\tilde{e}_\sigma^2 - \tilde{e}_\sigma^{trun\,2}} = \frac{\|\boldsymbol{X}^{*\perp}(n_\sigma) - \boldsymbol{X}^*(n_\sigma, \mathcal{I}_\sigma)\|_F}{\|\boldsymbol{X}\|_F} \tag{122}$$

(the error introduced as a result of using only p_σ sampling points instead of the entire set of finite element Gauss points, see Section 7.1.1) is *negligible in comparison to the discrepancies due to truncation of the POD basis*. For $n_\sigma = p_\sigma = 6$, for instance, the reconstruction error is less than 3% of the total stress error. In view of these results, it becomes clear that *further refinements in the algorithm for selecting the basic sampling points are in principle not necessary: the employed HIPM optimization algorithm, however*

heuristic, satisfactorily fulfills this purpose. If one wishes to lower the stress approximation error, it is far more effective to simply increase the level of truncation.

9.5.2. Stabilizing sampling points

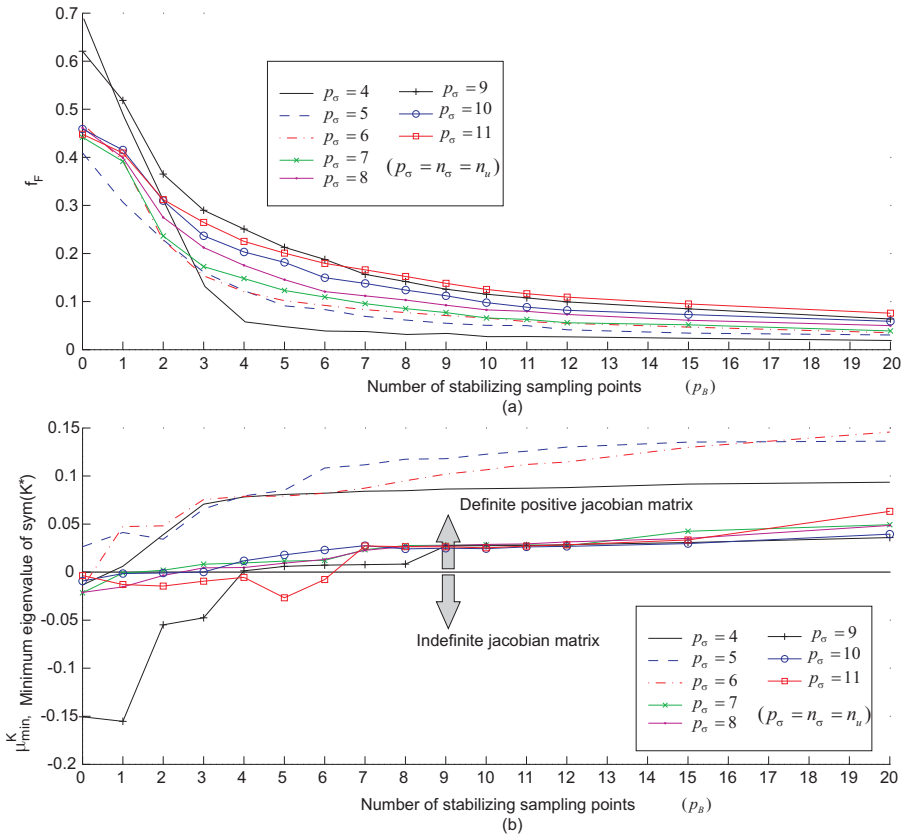


Figure 8: a) Factor f_F (defined in Eq.(99)) versus number of stabilizing sampling points p_B for varying numbers of basic sampling points p_σ (with $p_\sigma = n_\sigma = n_u$). b) Minimum eigenvalue μ^K_{min} (over all time steps and iterations for each p_σ) of the symmetric part of the reduced-order Jacobian matrix $\boldsymbol{K}^*$ versus number of stabilizing sampling points p_B.

Concerning what we have termed "stabilizing sampling points", Figure 8.a contains the graphs, for varying levels of truncation, of factor f_F defined

in Eq.(100) as a function of the number of stabilizing sampling points p_B. To study the influence of including such points on the spectral properties — positive defineteness— of the stiffness matrix, these graphs are accompanied, see figure 8.b, by the plots of the minimum eigenvalue μ^K_{min} (over all time steps and iterations for each case) of the symmetric part of the reduced-order Jacobian matrix $\boldsymbol{K}^*$ versus p_B. It can be seen that f_F decreases monotonically as the number of stabilizing sampling points increases, and such a decrease is reflected, as theoretically anticipated in Section 6.4.1, in the improvement of the spectral properties of the reduced-order Jacobian matrix (higher μ^K_{min} as p_B raises). For clarity, the minimum number of stabilizing sampling points required, for each level of truncation, to render positive definite $\boldsymbol{K}^*$ is plotted in Figure 9. From this plot, it can be gleaned that, roughly, the higher

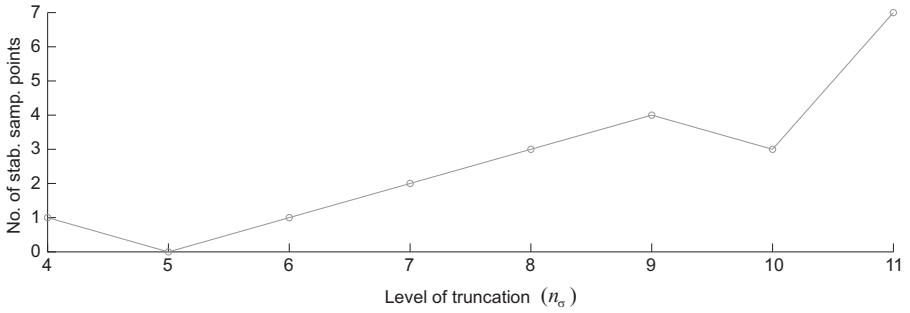


Figure 9: Minimum number of stabilizing sampling points required to make the Jacobian matrix $\boldsymbol{K}^*$ definite positive for each level of truncation $n_\sigma = n_u = p_\sigma$ (deduced from Figure 8).

the level of truncation (and thus the number of basic sampling points), the more stabilizing sampling points appear to be needed to ensure the positive definiteness of $\boldsymbol{K}^*$. For $n_\sigma = p_\sigma = 6$, adding just one stabilizing sampling points suffices, while for $n_\sigma = p_\sigma = 11$, 7 points are required.

Observation 9.1. *The values shown in Figure 9 correspond to the minimum p_B that leads to positive definite $\boldsymbol{K}^*$ when the prescribed strain path coincides with any of the "training" strain trajectory (displayed in Figure 4). Unfortunately, there is no guarantee that the Jacobian matrix will also exhibit this desirable property for prescribed strain histories different from the training ones. Thus, in view of such uncertainty, and in the interest of robustness, it is preferable to stay on the side of "caution" in this regard and use more*

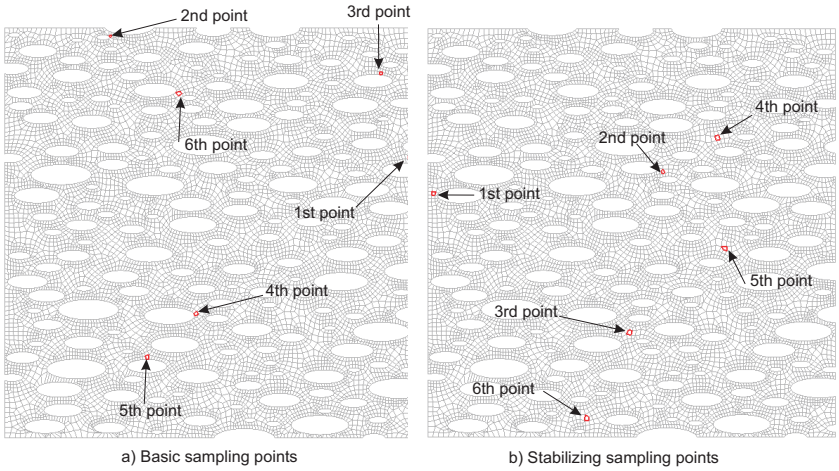


Figure 10: Location within the RVE of the finite elements (marked in red) that contains the first $p_\sigma = p_B = 6$ *basic* and *stabilizing* sampling points.

stabilizing sampling points that the minimum number indicated by the analysis based on the training strain trajectories. It is the authors' experience that a "safe" estimate for p_B is to simply take $p_B = p_\sigma$ —that is, equal number of basic and stabilizing sampling points. In adopting such a rule, the authors have not observed any convergence failures whatsoever, neither in the example under consideration nor in other cases not shown here.

The location of the first $p_\sigma = 6$ basic sampling points and the corresponding $p_B = 6$ stabilizing sampling points is depicted in Figure 10.

9.6. A posteriori errors: consistency analysis

The error measures displayed previously in Figures 5 and 7 only depend on the outcome of the SVD of the snapshot matrices; they can be calculated, thus, *before* actually constructing the reduced-order model. Error analyses based on such measures serve the useful purpose of providing a first hint of how many stress and displacement fluctuations modes are needed to satisfactorily replicate the full-order, finite element solution, and thereby, of prospectively evaluating the *viability* of the reduced basis approach itself.

However, these *a priori* error estimates do not tell the whole story. Expression (121) for the stress approximation error presumes that the stress

solution at the chosen sampling points is the one provided by the finite element model, thus ignoring the fact that, actually, in the reduced-order model, and for the general case of nonlinear, dissipative materials, the stress information at such points at a given time step is already polluted by truncation (in displacement fluctuations and stresses) and reconstruction (in stresses) errors originated in previous time steps. To quantify the extent to which this amalgam of accumulated errors affects the predictions furnished by the HP-ROM, it is necessary to perform a *consistency analysis*.

Generally speaking, a reduced basis approximation is said to be *consistent* if, in the limit of no truncation, it introduces no additional error in the solution of the same problem for which the data used in constructing the basis functions were acquired [14]. In the BVP under consideration, thus, consistency implies that, when using as input macro-strain paths the same trajectories employed in the "training" process, results obtained with the HP-ROM should converge, as n_σ and n_u increase, to the solution furnished by the full-order, finite element model. This condition can be checked by

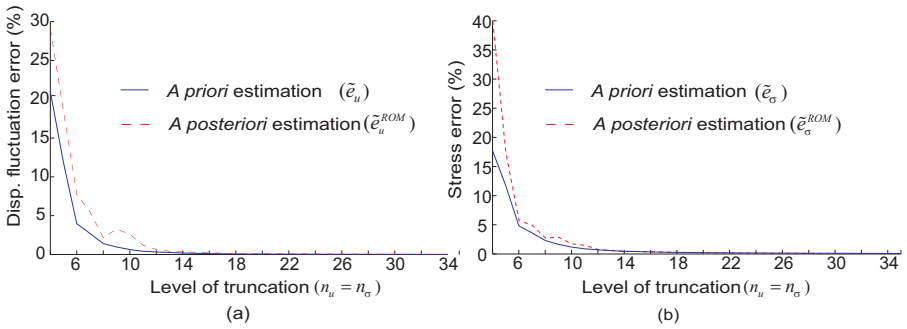


Figure 11: Comparison of the evolution of a priori and a posteriori error measures versus the level of truncation (using $n_u = n_\sigma = p_\sigma = p_B$). a) Displacement fluctuations (see Eqs. 119 and 123). b) Stresses (see Eqs. 121 and 124)

studying the evolution of the error measures defined as

$$\tilde{e}_u^{ROM}(n_u, n_\sigma, \mathcal{I}) := \frac{\|\bar{\boldsymbol{X}}_u - \bar{\boldsymbol{X}}_u^{*ROM}(n_u, n_\sigma, \mathcal{I})\|_F}{\|\bar{\boldsymbol{X}}_u\|_F}, \tag{123}$$

for the displacement fluctuations, and

$$\tilde{e}_\sigma^{ROM}(n_u, n_\sigma, \mathcal{I}_\sigma) := \frac{\|\boldsymbol{X} - \boldsymbol{X}^{*ROM}(n_u, n_\sigma, \mathcal{I}_\sigma)\|_F}{\|\boldsymbol{X}\|_F}, \tag{124}$$

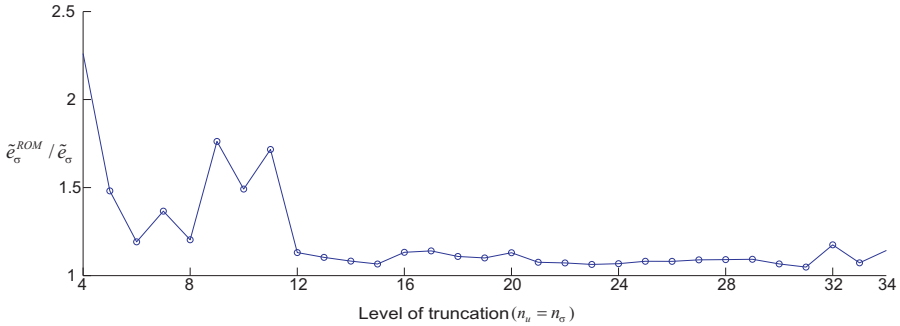


Figure 12: Ratio $\tilde{e}_\sigma^{ROM}/\tilde{e}_\sigma$ between the a posteriori and a priori measures for the stress approximation error against the level of truncation (using $n_u = n_\sigma = p_\sigma = p_B$).

for the stresses. (The superscript "ROM" is appended to highlight that, unlike $\bar{\boldsymbol{X}}_u^*$ and $\boldsymbol{X}^*$ in Eqs. (119) and (121), $\bar{\boldsymbol{X}}_u^{*ROM}$ and $\boldsymbol{X}^{*ROM}$ are matrices of displacement fluctuation and stress snapshots computed using the HP-ROM). Figures 11.a and 11.b contain the graphs of these *a posteriori* error measures, along with their respective *a priori* counterparts $\tilde{e}_u$ (Eq. 119) and $\tilde{e}_\sigma$ (Eq. 121), versus the level of truncation. It becomes clear from these graphs that consistency, in the sense given above, is observed in terms of both displacement fluctuations and stresses: the a posteriori error measures $\tilde{e}_u^{ROM}$ and $\tilde{e}_\sigma^{ROM}$ mimic essentially the decreasing tendency of their a priori counterparts $\tilde{e}_u$ and $\tilde{e}_\sigma$, respectively. It can be seen also that the *a priori error estimations* $\tilde{e}_u$ and $\tilde{e}_\sigma$ *constitute (rather tight) lower bounds* for *their a posteriori counterparts* $\tilde{e}_u^{ROM}$ and $\tilde{e}_\sigma^{ROM}$, respectively. This can be better appreciated, for the stresses, in Figure 12, where the ratio $\tilde{e}_\sigma^{ROM}/\tilde{e}_\sigma$ versus the level of truncation is plotted.

The degree of approximation that can be achieved using the proposed HP-ROM is quantified in a more "engineering" fashion in Figure 13, where we plot, for the case of the first training strain history (stretching in the longitudinal direction), the longitudinal, *macroscopic* stress-strain curves computed using the FE model, on the one hand, and the HP-ROM with $n_u = n_\sigma = 6, 7, 8$ modes, on the other hand. Observe that the maximum deviation from the FE response when using 6 modes (3 elastic and 3 inelastic) takes place at the onset of plastic yielding and is below 8%; remarkably, as deformation continues, this deviation gradually diminishes, being practically

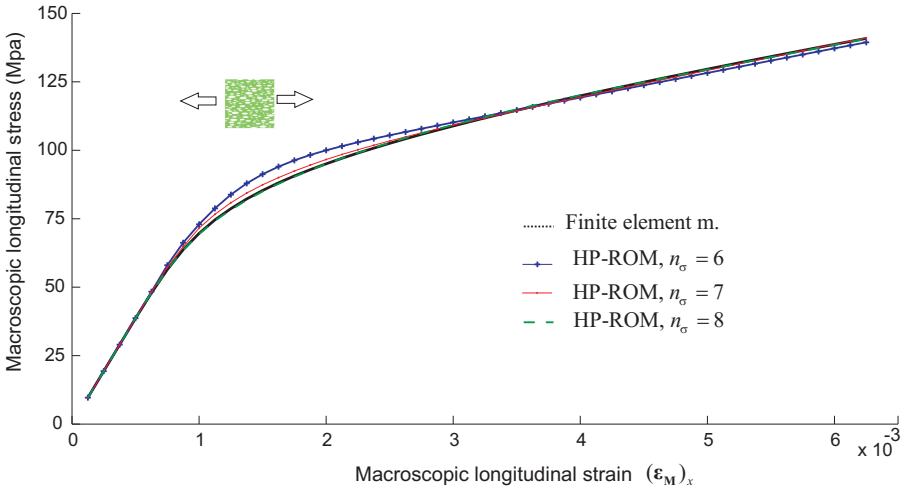


Figure 13: Longitudinal macroscopic stress versus longitudinal macroscopic strain computed using FEM and HP-ROM with $n_\sigma = n_u = 6, 7, 8$.

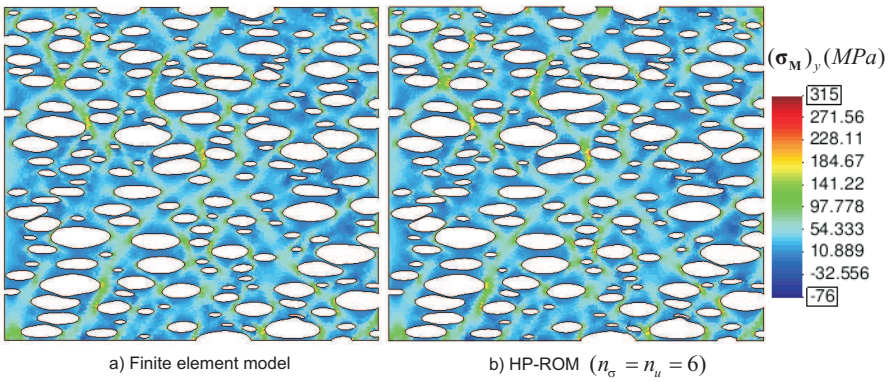


Figure 14: Contour plot of transversal stresses computed at the end of the first "training" strain history using a) FEM (b) HP-ROM with $n_\sigma = n_u = 6$. Deformed shapes are exaggerated (by a factor of 20).

negligible at the end of the process. Furthermore, by just increasing the order of truncation to $n_\sigma = n_u = 8$, differences between the HP-ROM and the FEM responses become virtually imperceptible at all levels of deformation.

Resemblance between HP-ROM and FEM results can also be appreciated in terms of stress distribution in the contour plots shown in Figure 14. Visually, there are no discernible differences between the two contour plots.

9.7. "Training" errors

The studies presented in the preceding subsections were aimed at examining the errors incurred in approximating the *snapshot solution space* $\mathcal{V}_u^{snp}$ by the reduced-order subspace $\mathcal{V}_u^* \in \mathcal{V}_u^{snp}$ spanned by the POD basis vectors —in the terminology of Section 3.1—, and to check that when $\mathcal{V}_u^* \to \mathcal{V}_u^{snp}$, the solution provided by the HP-ROM converges to that obtained with the FEM. But recall that the snapshot space $\mathcal{V}_u^{snp}$ is but a (presumably representative) subspace of $\mathcal{V}_u^\epsilon$, the manifold of $\mathcal{V}_u^h$ induced by the parametric dependence of the fine-scale BVP on the prescribed macroscopic strain history. Consequently, in general —for an arbitrary input strain trajectory— the HP-ROM solution will not converge to the solution provided by the FEM. To complete the error assessment analysis, thus, it is necessary to estimate also the errors inherent to the sampling of the parameter space —we call them *training errors*— and judge whether the selected training strain trajectories generate a snapshot subspace that is indeed *representative* of such a solution space[21] $\mathcal{V}_u^\epsilon$.

Ideally, one should carry out this error assessment by picking up, guided by some sound, statistically-based procedure, a sufficiently large set of strain paths and by comparing the solutions computed by the FEM and HP-ROM under such input strain paths for varying levels of truncation. Such a degree of rigor, however, is beyond the scope of the present work. Here, we limit ourselves here to analyze the quality of the HP-ROM approximation obtained for two different input strain histories, namely, a *uniaxial compression test*, and a *biaxial loading/unloading test*.

[21] To put it in less mathematical terms —by appealing to the the analogy, introduced in Remark 9.1, between the training of the RVE reduced-order model and the calibration of standard phenomenological models— we have "calibrated" our HP-ROM using the training tests displayed previously in Figure 4, and we have shown that the model is able to *exactly* replicate the behavior of the RVE in these tests when $n_u = n_\sigma$ is sufficiently large. Similarly to the situation encountered when dealing with standard phenomenological models, it remains now to assess the capability of the proposed HP-ROM to predict the behavior of the RVE under conditions different from those used in the "calibration" (training) process.

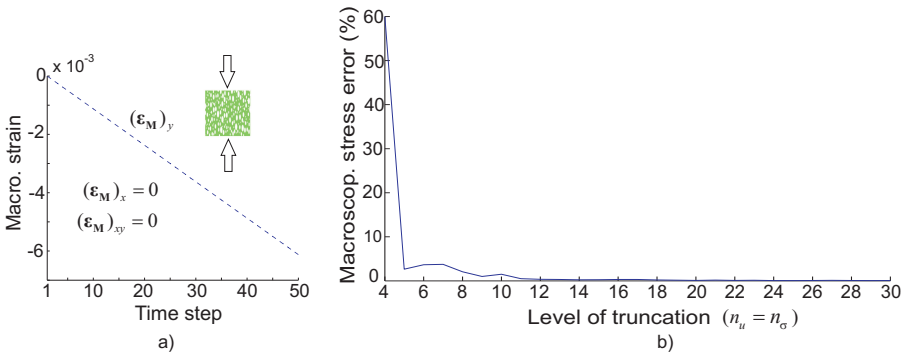


Figure 15: a) First strain trajectory employed for assessing training errors. b) Plot of the macroscopic error estimator $\tilde{E}^{ROM}_{\sigma,M}$ (see Eq.(125)) corresponding to this testing trajectory versus level of truncation ($n_\sigma = n_u$)

9.7.1. Uniaxial compression

The first strain path employed for the assessment is displayed in Figure 15.a; it represents a monotonic compression in the transversal direction (the model, see Figure 4, was trained using only stretching and shear, but not compression, tests). For purposes of evaluating the quality of the HP-ROM approximation, it is convenient to introduce the following macroscopic[22] stress error estimate:

$$\tilde{E}^{ROM}_{\sigma,M} := \sqrt{\frac{\sum_{i=1}^{n^t_{stp}} \|\boldsymbol{\sigma_M}^i - \boldsymbol{\sigma}_M^{*\,i,ROM}(n_\sigma, n_u, \mathcal{I})\|^2}{\sum_{i=1}^{n^t_{stp}} \|\boldsymbol{\sigma_M}^i\|^2}}, \tag{125}$$

where $\boldsymbol{\sigma_M}^i$ and $\boldsymbol{\sigma}_M^{*\,i,ROM}$ denote the macroscopic stress at the $i-th$ time step computed by the FEM and the HP-ROM, respectively. This error estimate is plotted in Figure 15.b versus the level of truncation $n_u = n_\sigma$. Observe that the error goes to zero as the number of employed modes increase. In this particular case, thus, *there is no additional error due to sampling of the*

[22]Recall that the output of interest in solving the fine-scale BVP is the macroscopic stress tensor; thus, the error estimate defined in Eq.(125) ($\tilde{E}^{ROM}_{\sigma,M}$) provides a more meaningful indication of the quality of the approximation than the stress error measure defined previously in Eq.(124) ($\tilde{e}^{ROM}_{\sigma}$). The latter is more suited for examining convergence properties of the HP-ROM approximation, since the minimization problem that underlies the SVD is posed in terms of the Frobenis norm.

parameter space.

Remark 9.2. *This simple example fittingly illustrates one of the acclaimed advantages of POD/Galerkin reduced-order approaches over "black box" methods such as artificial neural networks —that are also based on the partitioned offline-online computational paradigm—: POD/Galerkin reduced-order approaches preserve the "physics" of the problem one wishes to model and, as a consequence, are able to make physically-based extrapolations. For instance, in this case, the reduced-order model is able to exactly replicate (for sufficiently large $n_u = n_\sigma$) the macroscopic compressive behavior of the RVE, even though no information regarding this deformational state has been supplied to the model in the calibration (training) phase; the HP-ROM is "aware", figuratively speaking, that the matrix material in the RVE exhibits similar behavior in tension and compression (J2 plasticity).*

9.7.2. Biaxil loading/unloading test

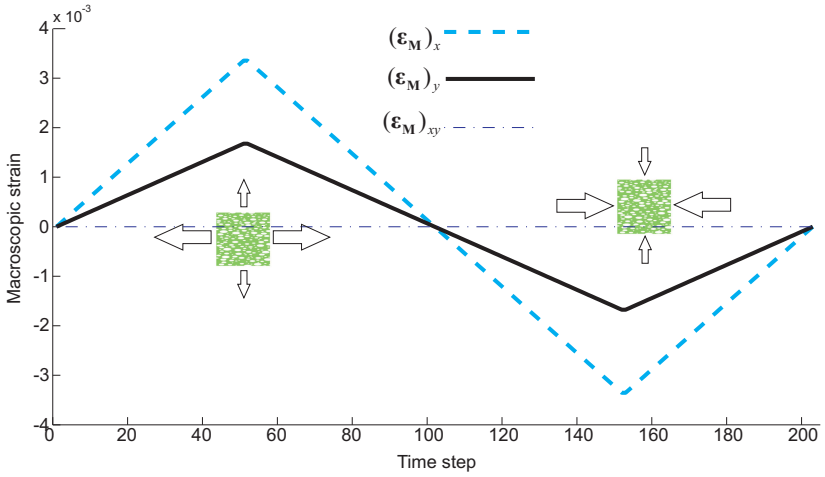


Figure 16: Second strain trajectory employed for assessing training errors.

A more severe test for assessing errors associated to the training process is provided by the strain trajectory shown in Figure 16. Indeed, while the training strain histories of Figure 4 only included monotonic, uniaxial stretching, the strain history displayed in Figure 16 consists of a cycle of *biaxial, loading/unloading stretching* (time steps 1 to 100) and biaxial loading/unloading

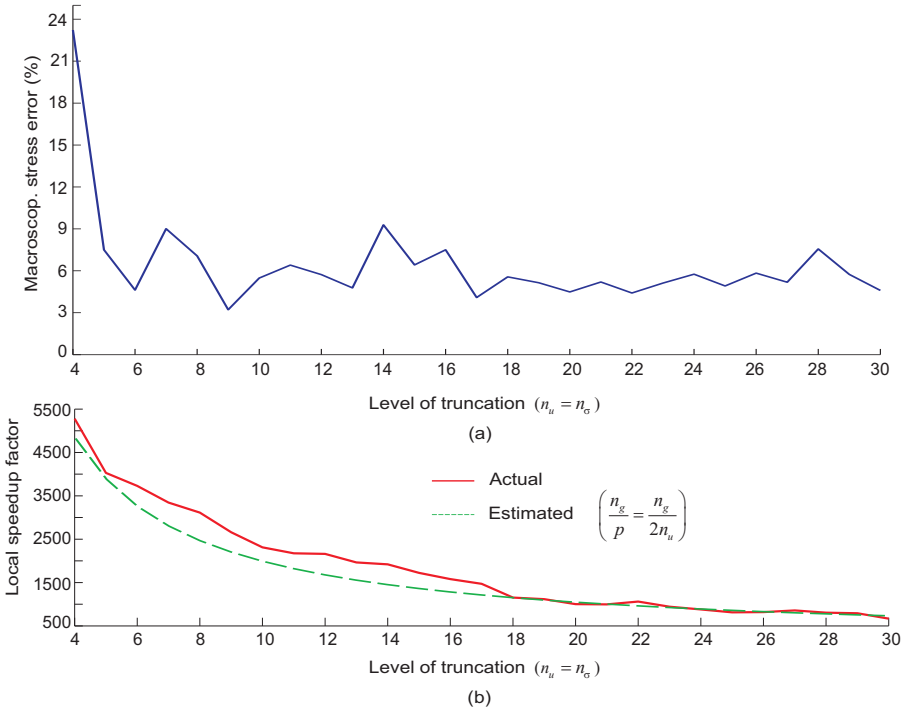


Figure 17: a) Macroscopic error estimator $\tilde{E}^{ROM}_{\sigma,M}$ (see Eq.(125)) versus level of truncation ($n_\sigma = n_u$) for the case of testing trajectory shown in Figure 16,. b) Local speedup factor S_{loc} (defined in Eq.(126)) reported for this case versus level of truncation. This plot is accompanied by the graph of the ratio n_g/p, where $n_g = 38984$ is the total number of Gauss points of the finite element mesh, and $p = 2n_\sigma$ the number of sampling points employed for numerically integrating the HP-ROM.

compression (time steps 101 to 200). The graph of the macroscopic error estimator (125) corresponding to this input strain path as a function of the level of truncation is represented in Figure 17.a. It can be readily perceived that, in this case, and in contrast to the situation encountered in the previously discussed input strain trajectory, the macroscopic stress does not go to zero as the number of POD modes included in the basis increases. Rather, the graph drops sharply from 24% to approximately 5% at $n_\sigma = n_u = 5$ (second inelastic mode), and then fluctuates erratically, with no apparent trend, between 3% and 10% —a level of accuracy that, nevertheless, may be deemed

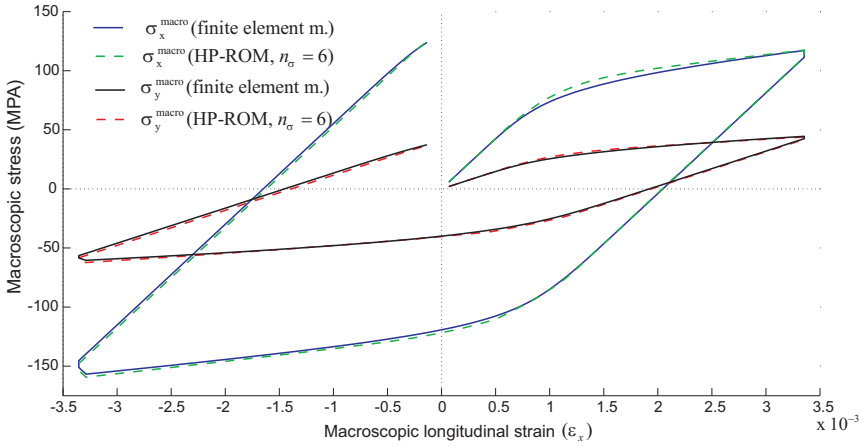


Figure 18: Longitudinal and transversal macroscopic stress versus longitudinal macroscopic strain computed using the FEM and the HP-ROM with $n_\sigma = n_u = 6$ (for the case of the testing trajectory shown in Figure 16)

.

more than acceptable in most practical applications. A more clear picture of the accuracy of the approximation for the particular case of $n_\sigma = n_u = 6$ can be obtained from the stress-strain diagrams shown in figure 18.

9.8. Speedup analysis

Lastly, we turn our attention to one of the main concerns of the present work: the issue of computational efficiency. For a given error level, how many times can the proposed HP-ROM speed up the calculation of the RVE response with respect to the reference finite element model? Let us define the local *speedup factor* as the ratio

$$S_{loc} := \frac{t_{FE}(n, n_g)}{t_{ROM}(n_\sigma, n_u, p)}, \tag{126}$$

where t_{FE} and t_{ROM} denote the CPU times required to compute the FE and HP-ROM macro-stress responses, respectively, induced by a given input strain history[23] In Figure 17.b, we show the graph of the speedup factor

[23]The computational cost associated to the offline stage —generation of snapshots plus the comparatively negligible expenses of applying the POD and selecting the sampling

reported in the the case of the input strain path of Figure 16 as a function of the number of POD modes included in the analysis (recall in this respect that $n_u = n_\sigma = p/2$). We plot also in Figure 17.b the ratio n_g/p, i.e., the relation between the total number of integration points in the finite element model ($n_g = 38984$) and in the reduced order model (p). It can be gleaned from Figure 17.b that the *reported speedup factors are of the same order of magnitude as the ratio* n_g/p; i.e.:

$$S_{loc} \sim \frac{n_g}{p} = \frac{n_g}{2n_\sigma}, \tag{127}$$

(this indicates that the evaluation of the stresses at the integration points dominates the total computational cost). Although these results are no doubt influenced and biased by the particular programming language and coding style employed —we use an in-house, non-vectorized Matlab program operating in a Linux platform—, and, consequently, this trend may not be exactly observed when using other programming languages and/or platforms, they serve to provide an idea of the tremendous *gains in performance* that can be achieved using the proposed ROM; for $n_\sigma = p = 6$ modes, for instance, the *computational cost is reduced by a factor above* 3600, while still capturing 95% of the full-order, high-fidelity information —the *essential information*.

10. Concluding remarks

One of the the most striking features of the proposed reduced-order model is perhaps the conceptual simplicity of the RVE equilibrium equation in its hyperreduced-order form: the sum of (reduced) internal forces at the preselected sampling points must give identical result either calculated using *observed* stresses or *fitted* stresses. Although this condition appears, in hindsight, rather reasonable, even obvious —it ensures maximum resemblance between reduced-order and full-order responses at the sampling points— it seems difficult to arrive at it without the benefit of the integration procedure —based on the notion of *expanded approximation space*— advocated in the present paper.

points—has been deliberately ruled out from this speedup analysis because, in two-scale homogenization contexts, the RVE equilibrium problem is to be solved a sheer number of times and, consequently, this overhead cost is quickly amortized.

The hyperreduced form of the RVE equilibrium equation excels not only in its conceptual simplicity; the corresponding solution scheme is also very simple to implement. Taking as departure point an existing FE code, one only has to replace the typical loop over elements in the FE code by a loop over the pre-selected sampling points $\{\mathbf{x}_{\mathcal{I}_1}, \mathbf{x}_{\mathcal{I}_2}, \ldots \mathbf{x}_{\mathcal{I}_p}\}$. The stress vectors and corresponding constitutive tangent matrices obtained at each stage of the loop are stored in the gappy weighted vector $\hat{\boldsymbol{\Sigma}}$ and the matrix $\hat{\mathbb{C}}$, respectively, and, then the residual vector and the Jacobian matrix are computed as $\hat{\mathbb{B}}^{**^T}\hat{\boldsymbol{\Sigma}}$ and $\hat{\mathbb{B}}^{**^T}\hat{\mathbb{C}}\,\hat{\mathbb{B}}^{*}$, respectively. Notice that *no assembly process is needed*, nor has one to worry about imposing *boundary conditions*. Once convergence is achieved, the macroscopic stress value is simply calculated as $\boldsymbol{\sigma_M} = \mathbb{T}\hat{\boldsymbol{\Sigma}}$. It should be emphasized again that the *operation count in both solving this hyperreduced RVE equation and updating the macroscopic stress vector depends exclusively on the reduced dimensions* n_u and p (number of fluctuation modes and number of sampling points, respectively). Likewise, *storage of history data (internal variables) is only required at the pre-selected sampling points*. Computational savings accrue, thus, not only in terms of number of operations, but also in terms of memory requirements.

The success of the proposed homogenization strategy is predicated on the assumption that displacement and stress fields induced by the parametric dependence on the input macroscopic strain can be approximated by (relatively) low-dimensional functions. Numerical results shown in the preceding section seem to suggest that, in general, this assumption may be expected to hold in the case of *materials governed by strain hardening laws in the small strain regime* —regardless of the geometrical complexity, number and distribution of heterogeneities within the RVE. However, it is by no means apparent that this conclusion can be easily extended to more (kinematically and phenomenologically) complex scenarios, involving large deformations, strain localization, decohesion, etc. For instance, can the deformational behavior of an RVE affected by multiple propagating cracks be represented also in a parsimonious manner, as in the case of strain hardening? Or will the number of modes necessary to accurately replicate its response combinatorially increase with the number of potential crack paths (i.e., with the geometrical complexity of the RVE)? Undoubtedly, in these complex, nonlinear scenarios, the task of sampling the parameter space —i.e., of choosing the macro-strain histories at which to obtain the snapshots— will become quite complicated, due to the richness of possible deformational patterns (void closure, propagating discontinuities ...), and thus difficult to carry out

on intuitive basis (as it has been done in the present paper). Therefore, it would be desirable to systematize this crucial task, as well as to provide some statistical means to *certify*, so to speak, the representativeness of the chosen snapshots. Likewise, topological variations caused by these phenomena may render POD-based compression algorithms inappropriate for collapsing the dimensions of the snapshots space; in such cases, nonlinear dimensionality reduction methods such as the Isomap algorithm [43] may be more suitable. Research in these fronts is currently in progress and will be reported in forthcoming publications.

Acknowledgements

The research leading to these results has received funding from, on the one hand, the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 320815, Advanced Grant Project COMP-DES-MAT, and, on the other hand, the Spanish Ministry of Science and Innovation under grant BIA2011-24258.

A. Proper Orthogonal Decomposition

The formal statement of the POD problem goes as follows: given the ensemble of snapshots $\{\boldsymbol{u}^1, \boldsymbol{u}^2, \dots \boldsymbol{u}^{n_{snp}}\}$, find a set of $n_u < n_{snp}$ orthogonal basis functions $\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \dots \boldsymbol{\Phi}_{n_u}\}$ ($\boldsymbol{\Phi}_i \in \mathcal{V}_u^{snp}$) such that the error defined as

$$e_u(n_u) := \sqrt{\sum_{k=1}^{n_{snp}} \|\boldsymbol{u}^k - \mathbf{P}^* \boldsymbol{u}^k\|^2_{L_2(\Omega)}} \tag{128}$$

is minimized. Here, $\mathbf{P}^* \boldsymbol{u}^k$ represents the projection of $\boldsymbol{u}^k$ onto the subspace spanned by the basis functions $\{\boldsymbol{\Phi}_i\}_{i=1}^{n_u}$, and $\|\cdot\|_{L_2(\Omega)}$ symbolizes the L_2 norm. We shall denote by $\boldsymbol{\Phi}_i$ ($i = 1, 2 \dots n_u$) the column vector containing the values of basis function $\boldsymbol{\Phi}_i$ at the nodes of the underlying finite element mesh. Likewise, the matrix formed by such vectors, $\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \cdots \end{bmatrix} \in \mathbb{R}^{n \cdot d \times n_u}$, will be hereafter called the *reduced basis matrix*.

The reduced basis matrix $\boldsymbol{\Phi}$ can be computed from the snapshot matrix $\boldsymbol{X}_u$ by means of the Singular Value Decomposition (SVD) as follows (see

Ref. [16] for more details). We first define the matrix[24]

$$\mathbf{M}_{IJ} := \int_{\Omega} N_I(\mathbf{x}) N_J(\mathbf{x})\, d\Omega \qquad I, J = 1, 2 \ldots n. \tag{129}$$

Let $\mathbf{M} = \bar{\mathbf{M}}^T \bar{\mathbf{M}}$ be the Cholesky decomposition of $\mathbf{M}$, and let $\bar{\boldsymbol{X}}_u$ denote the matrix defined as:

$$\bar{\boldsymbol{X}}_u := \bar{\mathbf{M}} \boldsymbol{X}_u. \tag{130}$$

Then, we compute the reduced SVD [37] of $\bar{\boldsymbol{X}}_u$, that is, the factorization

$$\bar{\boldsymbol{X}}_u = \bar{\boldsymbol{U}} \bar{\boldsymbol{S}} \bar{\boldsymbol{V}}^T, \tag{131}$$

where $\bar{\boldsymbol{V}} \in \mathbb{R}^{n_{snp} \times r}$ (r is the rank of $\bar{\boldsymbol{X}}_u$) and $\bar{\boldsymbol{U}} \in \mathbb{R}^{n \cdot d \times r}$ stand for the matrices of *right and left singular vectors*, respectively; and $\bar{\boldsymbol{S}} \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing the *singular values* of $\boldsymbol{X}_u$. The $i-th$ column of the reduced basis matrix $\boldsymbol{\Phi}$ is finally related to the $i-th$ *left singular vector* of $\bar{\boldsymbol{X}}_u$ through expression

$$\boldsymbol{\Phi}_i = \bar{\mathbf{M}}^{-1} \bar{\boldsymbol{U}}_i, \quad i = 1, 2 \ldots n_u. \tag{132}$$

B. Elastic/Inelastic reduced basis matrix

This appendix is devoted to provide further details concerning the actual numerical implementation of the elastic/inelastic partitioned strategy, presented in Section 3.2.1, for the computation of the *reduced basis matrices* $\boldsymbol{\Phi}$ (displacement fluctuations). The steps to arrive at the desired matrix basis $\boldsymbol{\Phi}$ are summarized in the following.

1. Compute finite element stress solutions for representative, input macro-strain histories.
2. Store the displacement fluctuation solutions computed at each time step of these macro-strain trajectories in the displacement fluctuations *snapshot matrix* $\boldsymbol{X}_u \in \mathbb{R}^{n \cdot d \times n_{snp}}$:

$$\boldsymbol{X}_u = \begin{bmatrix} \boldsymbol{U}^1 & \boldsymbol{U}^2 & \cdots & \boldsymbol{U}^{n_{snp}} \end{bmatrix} \tag{133}$$

[24] Note that, except for the density factor, this matrix $\mathbf{M}$ is similar to the "mass matrix" appearing in finite element implementations of dynamical problems. For implementational purposes, one can simply use a lumped version of such a matrix.

3. Pick up from $\boldsymbol{X}_u$ a minimum of m_e ($m_e = 6$ for 3D problems, and $m_e = 3$ for plane strain) linearly independent columns corresponding to purely elastic solutions. Store these columns in a matrix $\boldsymbol{Z}_u^{el}$.
4. Perform the *reduced singular value decomposition* (SVD) of the matrix defined as
$$\bar{\boldsymbol{Z}}_u^{el} := \bar{\mathbf{M}}\boldsymbol{Z}_u^{el} \tag{134}$$
where $\bar{\mathbf{M}}$ is the matrix of the Cholesky factorization of $\mathbf{M}$ ($\mathbf{M} = \bar{\mathbf{M}}^T\bar{\mathbf{M}}$). A basis matrix for Range($\boldsymbol{Z}_u^{el}$) is finally obtained as
$$\boldsymbol{D}^{el} := \bar{\mathbf{M}}^{-1}\bar{\boldsymbol{D}}^{el} \tag{135}$$
$\bar{\boldsymbol{D}}^{el} \in \mathbb{R}^{n\cdot d\times m_e}$ being the matrix of left singular vectors arising from the SVD of $\bar{\boldsymbol{Z}}_u^{el}$. In principle, $\boldsymbol{D}^{el}$ may be used as the desired elastic basis matrix $\boldsymbol{\Phi}^{el}$. However, $\boldsymbol{D}^{el}$ does not enjoy any optimality property with respect to $\boldsymbol{X}_u$—it is only optimal with respect to the matrix $\boldsymbol{Z}_u^{el}$ of chosen elastic snapshots.
5. For consistency in the approximation, thus, it is preferable to derive $\boldsymbol{\Phi}^{el}$ from the the "elastic component" of $\boldsymbol{X}_u$ —the orthogonal projection of $\boldsymbol{X}_u$ onto Range($\boldsymbol{D}^{el}$)—; the expression for this projection reads:
$$\boldsymbol{X}_u^{el} = \boldsymbol{D}^{el}(\boldsymbol{D}^{el^T}\mathbf{M}\boldsymbol{X}_u) \tag{136}$$
The elastic basis matrix can be finally calculated from $\boldsymbol{X}_u^{el}$ as:
$$\boldsymbol{\Phi}^{el} := \bar{\mathbf{M}}^{-1}\bar{\boldsymbol{\Phi}}^{el}. \tag{137}$$
where $\bar{\boldsymbol{\Phi}}^{el}$ is the matrix of left singular vectors emerging from the reduced SVD of $\bar{\boldsymbol{X}}_u^{el} := \bar{\mathbf{M}}\boldsymbol{X}_u^{el}$; i.e $\bar{\boldsymbol{X}}_u^{el} = \bar{\boldsymbol{\Phi}}^{el}\bar{\boldsymbol{S}}_u^{el}\bar{\boldsymbol{V}}_u^{el^T}$.
6. Calculate the "inelastic component" $\boldsymbol{X}_u^{in}$ of the snapshot matrix $\boldsymbol{X}_u$ as:
$$\boldsymbol{X}_u^{in} = \boldsymbol{X}_u - \boldsymbol{X}_u^{el}; \tag{138}$$
that is, $\boldsymbol{X}_u^{in}$ is the orthogonal projection of $\boldsymbol{X}_u$ onto the orthogonal complement, in Range($\boldsymbol{X}_u$), of Range($\boldsymbol{\Phi}^{el}$).
7. It is now on this *inelastic* snapshot matrix $\boldsymbol{X}_u^{in}$ that we apply the POD in order to identify and unveil the *essential* or most "energetic" inelastic fluctuation modes. This is done by first carrying out the reduced SVD of $\bar{\boldsymbol{X}}_u^{in} = \bar{\mathbf{M}}\boldsymbol{X}_u^{in}$:
$$\bar{\boldsymbol{X}}_u^{in} = \bar{\boldsymbol{D}}^{in}\bar{\boldsymbol{S}}_u^{in}\bar{\boldsymbol{V}}_u^{in^T}. \tag{139}$$

The $i-th$ POD basis vector of $\boldsymbol{X}_u^{in}$ is then given by:

$$\boldsymbol{\Phi}_i^{in} = \bar{\mathbf{M}}^{-1}\bar{\boldsymbol{D}}_i^{in}, \quad i = 1, 2 \ldots n_u - m_e. \tag{140}$$

8. The desired basis matrix $\boldsymbol{\Phi} \in \mathbb{R}^{n\cdot d \times n_u}$ adopts finally the form:

$$\boldsymbol{\Phi} = [\boldsymbol{\Phi}^{el}\ \boldsymbol{\Phi}^{in}] = [\overbrace{\boldsymbol{\Phi}_1^{el}\ \boldsymbol{\Phi}_2^{el}\ \cdots\ \boldsymbol{\Phi}_{m_e}^{el}}^{\text{Elastic modes}}\ \overbrace{\boldsymbol{\Phi}_1^{in}\ \boldsymbol{\Phi}_2^{in}\ \cdots\ \boldsymbol{\Phi}_{n_u-m_e}^{in}}^{\text{Essential inelastic modes}}] \tag{141}$$

C. Block matrix pseudoinverse of the expanded basis matrix

The inverse of a 2x2 *symmetric* block matrix is given by the following expression (see, for instance, Ref. [11]):

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B^T} & \mathbf{C} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A^{-1}} + \mathbf{A^{-1}BS^{-1}B^TA^{-1}} & \mathbf{-A^{-1}BS^{-1}} \\ \mathbf{-S^{-1}B^TA^{-1}} & \mathbf{S^{-1}} \end{bmatrix} \tag{142}$$

where

$$\mathbf{S} = \mathbf{C} - \mathbf{B^TA^{-1}B} \tag{143}$$

is the so-called Schur complement of $\mathbf{A}$ in $\mathbf{M}$. This formula can be used to derive closed-form expressions for the modal coefficients $\boldsymbol{c}^{ad}$ and $\boldsymbol{c}^{in}$ (see Section 6.2). The departure point is equation Eq.(67):

$$\begin{bmatrix} \boldsymbol{c}^{ad} \\ \boldsymbol{c}^{in} \end{bmatrix} = ([\hat{\boldsymbol{\Psi}}\ \hat{\mathbb{B}}^*])^{\dagger}\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} \hat{\boldsymbol{\Psi}}^T\hat{\boldsymbol{\Psi}} & \hat{\boldsymbol{\Psi}}^T\hat{\mathbb{B}}^* \\ \hat{\mathbb{B}}^{*T}\hat{\boldsymbol{\Psi}} & \hat{\mathbb{B}}^{*T}\hat{\mathbb{B}}^* \end{bmatrix}^{-1} \begin{bmatrix} \hat{\boldsymbol{\Psi}}^T \\ \hat{\mathbb{B}}^{*T} \end{bmatrix} \hat{\boldsymbol{\Sigma}}, \tag{144}$$

where $([\hat{\boldsymbol{\Psi}}\ \hat{\mathbb{B}}^*])^{\dagger}$ designates the pseudo-inverse of the *gappy* expanded basis matrix. By setting:

$$\mathbf{A} = \hat{\boldsymbol{\Psi}}^{\mathbf{T}}\hat{\boldsymbol{\Psi}}, \quad \mathbf{B} = \hat{\boldsymbol{\Psi}}^{\mathbf{T}}\hat{\mathbb{B}}^*, \qquad \mathbf{C} = \hat{\mathbb{B}}^{*\mathbf{T}}\hat{\mathbb{B}}^*, \tag{145}$$

and by inserting Eq.(142) into Eq.(144), one obtains upon expansion:

$$\boldsymbol{c}^{in} = \mathbf{S^{-1}}\left(\mathbf{-B^TA^{-1}}\hat{\boldsymbol{\Psi}}^{\mathbf{T}} + \hat{\mathbb{B}}^{*\mathbf{T}}\right)\hat{\boldsymbol{\Sigma}} \tag{146}$$

and

$$\begin{aligned} \boldsymbol{c}^{ad} &= \left(\left(\mathbf{A^{-1}} + \mathbf{A^{-1}BS^{-1}B^TA^{-1}}\right)\hat{\boldsymbol{\Psi}}^T - \mathbf{A^{-1}BS^{-1}}\hat{\mathbb{B}}^{*\mathbf{T}}\right)\hat{\boldsymbol{\Sigma}} \\ &= \mathbf{A^{-1}}\hat{\boldsymbol{\Psi}}^{\mathbf{T}}\hat{\boldsymbol{\Sigma}} + \mathbf{A^{-1}B}\overbrace{\mathbf{S^{-1}}\left(\mathbf{-B^TA^{-1}}\hat{\boldsymbol{\Psi}}^{\mathbf{T}} + \hat{\mathbb{B}}^{*\mathbf{T}}\right)\hat{\boldsymbol{\Sigma}}}^{\boldsymbol{c}^{\mathbf{in}}} \\ &= \mathbf{A^{-1}}\hat{\boldsymbol{\Psi}}^{\mathbf{T}}\hat{\boldsymbol{\Sigma}} + \mathbf{A^{-1}B}\boldsymbol{c}^{\mathbf{in}}. \end{aligned} \tag{147}$$

By substituting back Eq.(145) into the above equation, and taking into account that:

$$\hat{\boldsymbol{\Psi}}^{\dagger} = (\hat{\boldsymbol{\Psi}}^{T}\hat{\boldsymbol{\Psi}})^{-1}\hat{\boldsymbol{\Psi}}^{T} \tag{148}$$

one finally gets:

$$\boldsymbol{c}^{ad} = \hat{\boldsymbol{\Psi}}^{\dagger}(\hat{\boldsymbol{\Sigma}} - \hat{\mathbb{B}}^{*}\boldsymbol{c}^{in}), \tag{149}$$

$$\boldsymbol{c}^{in} = \mathbf{S}^{-\mathbf{1}}\hat{\mathbb{B}}^{*\mathbf{T}}\left(\boldsymbol{I} - \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger}\right)\hat{\boldsymbol{\Sigma}}, \tag{150}$$

where

$$\mathbf{S} = \hat{\mathbb{B}}^{*\mathbf{T}}(\boldsymbol{I} - \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^{\dagger})\hat{\mathbb{B}}^{*}. \tag{151}$$

References

[1] Abdulle, A., Bai, Y., 2012. Adaptive reduced basis finite element heterogeneous multiscale method.

[2] Abdulle, A., Bai, Y., 2012. Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems. Journal of Computational Physics.

[3] An, S., Kim, T., James, D., 2009. Optimizing cubature for efficient integration of subspace deformations. ACM transactions on graphics 27 (5), 165.

[4] Ashby, M., 1992. Physical modelling of materials problems. Materials Science and Technology 8 (2), 102–111.

[5] Astrid, P., 2004. Reduction of process simulation models: a proper orthogonal decomposition approach. Technische Universiteit Eindhoven.

[6] Astrid, P., Weiland, S., Willcox, K., Backx, T., 2008. Missing point estimation in models described by proper orthogonal decomposition. Automatic Control, IEEE Transactions on 53 (10), 2237–2251.

[7] Barrault, M., Maday, Y., Nguyen, N., Patera, A., 2004. An empirical interpolation'method: application to efficient reduced-basis discretization of partial differential equations. Comptes Rendus Mathematique 339 (9), 667–672.

[8] Bishop, C., en ligne), S. S., 2006. Pattern recognition and machine learning. Vol. 4. springer New York.

[9] Bohm, H., 1998. A short introduction to basic aspects of continuum micromechanics. CDL-FMD Report 3.

[10] Boyaval, S., 2007. Reduced-basis approach for homogenization beyond the periodic setting. Arxiv preprint math/0702674.

[11] Boyd, S., Vandenberghe, L., 2004. Convex optimization. Cambridge Univ Pr.

[12] Bui-Thanh, T., 2007. Model-constrained optimization methods for reduction of parameterized large-scale systems. Ph.D. thesis, Citeseer.

[13] Bui-Thanh, T., Willcox, K., Ghattas, O., 2008. Model reduction for large-scale systems with high-dimensional parametric input space. SIAM Journal on Scientific Computing 30 (6), 3270–3288.

[14] Carlberg, K., Bou-Mosleh, C., Farhat, C., 2011. Efficient non-linear model reduction via a least-squares petrov–galerkin projection and compressive tensor approximations. International Journal for Numerical Methods in Engineering 86 (2), 155–181.

[15] Carlberg, K., Farhat, C., 2008. A Compact Proper Orthogonal Decomposition Basis for Optimization-Oriented Reduced-Order Models. AIAA Paper 5964, 10–12.

[16] Carlberg, K., Farhat, C., 2011. A low-cost, goal-oriented compact proper orthogonal decompositionbasis for model reduction of static systems. International Journal for Numerical Methods in Engineering 86 (3), 381–402.

[17] Chaturantabut, S., Sorensen, D., 2009. Application of POD and DEIM on Dimension Reduction of Nonlinear Miscible Viscous Fingering in Porous Media.

[18] Chaturantabut, S., Sorensen, D., 2010. Discrete empirical interpolation for nonlinear model reduction. In: Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on. IEEE, pp. 4316–4321.

[19] Cook, R., 1995. Finite element modeling for stress analysis. John Wiley and Sons., New York.

[20] Couégnat, G., 2008. Approche multiéchelle du comportement mécanique de matériaux composites à renfort tissé. Ph.D. thesis, Université Sciences et Technologies-Bordeaux I.

[21] Cremonesi, M., Néron, D., Guidault, P.-A., Ladevèze, P., 2013. A pgd-based homogenization technique for the resolution of nonlinear multi-scale problems. Computer Methods in Applied Mechanics and Engineering 267, 275–292.

[22] de Souza Neto, E., Feijóo, R., 2006. Variational foundations of multi-scale constitutive models of solid: small and large strain kinematical formulation. LNCC Research & Development Report 16.

[23] DeVore, R., Iserles, A., Suli, E., 2001. Foundations of computational mathematics. Cambridge Univ Pr.

[24] Drago, A., Pindera, M., 2007. Micro-macromechanical analysis of heterogeneous materials: Macroscopically homogeneous vs periodic microstructures. Composites science and technology 67 (6), 1243–1263.

[25] Dvorak, G., Wafa, A., Bahei-El-Din, Y., 1994. Implementation of the transformation field analysis for inelastic composite materials. Computational Mechanics 14 (3), 201–228.

[26] Efendiev, Y., Galvis, J., Gildin, E., 2012. Local-global multiscale model reduction for flows in high-contrast heterogeneous media. Journal of Computational Physics.

[27] Efendiev, Y., Galvis, J., Thomines, F., 2012. A systematic coarse-scale model reduction technique for parameter-dependent flows in highly heterogeneous media and its applications.

[28] Everson, R., Sirovich, L., 1995. Karhunen–Loeve procedure for gappy data. Journal of the Optical Society of America A 12 (8), 1657–1664.

[29] Feyel, F., Chaboche, J., 2000. Fe-2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre sic/ti composite materials. Computer methods in applied mechanics and engineering 183 (3), 309–330.

[30] Fish, J., Shek, K., Pandheeradi, M., Shephard, M., 1997. Computational plasticity for composite structures based on mathematical homogenization: Theory and practice. Computer Methods in Applied Mechanics and Engineering 148 (1-2), 53–73.

[31] Galbally, D., Fidkowski, K., Willcox, K., Ghattas, O., 2010. Non-linear model reduction for uncertainty quantification in large-scale inverse problems. International Journal for Numerical Methods in Engineering 81 (12), 1581–1608.

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

[33] Grepl, M., Maday, Y., Nguyen, N., Patera, A., 2007. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. Mathematical Modelling and Numerical Analysis 41 (3), 575–605.

[34] Gross, D., Seelig, T., 2011. Fracture mechanics: with an introduction to micromechanics. Springer.

[35] Hernández, J. A., Oliver, J., Huespe, A., Caicedo, M., 2012. High-performance model reduction procedures in multiscale simulations. Monograph CIMNE (ISBN:978-84-9939640-6-1).
URL http://centrovnet.cimne.upc.edu/cvdata/cntr7/dtos/img/mdia/Download

[36] Hoffman, J. D., 2001. Numerical Methods for Engineers and Scientists. Marcel Dekker.

[37] Hogben, L., 2006. Handbook of linear algebra. Chapman & Hall/CRC.

[38] Hu, Y., Hwang, J., Perry, S., 2002. Handbook of neural network signal processing. The Journal of the Acoustical Society of America 111, 2525.

[39] Kim, T., James, D., 2009. Skipping steps in deformable simulation with online model reduction. In: ACM SIGGRAPH Asia 2009 papers. ACM, pp. 1–9.

[40] Kouznetsova, V., 2002. Computational homogenization for the multi-scale analysis of multi-phase materials.

[41] Krysl, P., Lall, S., Marsden, J., 2001. Dimensional model reduction in non-linear finite element dynamics of solids and structures. International Journal for Numerical Methods in Engineering 51 (4), 479–504.

[42] Kunisch, K., Volkwein, S., 2010. Optimal snapshot location for computing pod basis functions. ESAIM: Mathematical Modelling and Numerical Analysis 44 (3), 509.

[43] Li, Z., Wen, B., Zabaras, N., 2010. Computing mechanical response variability of polycrystalline microstructures through dimensionality reduction techniques. Computational Materials Science 49 (3), 568–581.

[44] Lovasz, L., Pelikan, J., Vesztergombi, K., 2003. Discrete Mathematics: Elementary and Beyond. Springer.

[45] Lubliner, J., 1990. Plasticity Theory. McMillan, New York.

[46] Maday, Y., Nguyen, N., Patera, A., Pau, G., 2007. A general, multipurpose interpolation procedure: the magic points.

[47] Maday, Y., Patera, A., Turinici, G., 2002. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods.

[48] Michel, J., Moulinec, H., Suquet, P., 1999. Effective properties of composite materials with periodic microstructure: a computational approach. Computer methods in applied mechanics and engineering 172 (1-4), 109–143.

[49] Michel, J., Suquet, P., 2003. Nonuniform transformation field analysis. International journal of solids and structures 40 (25), 6937–6955.

[50] Michel, J., Suquet, P., 2004. Computational analysis of nonlinear composite structures using the nonuniform transformation field analysis. Computer methods in applied mechanics and engineering 193 (48-51), 5477–5502.

[51] Monteiro, E., Yvonnet, J., He, Q., 2008. Computational homogenization for nonlinear conduction in heterogeneous materials using model reduction. Computational Materials Science 42 (4), 704–712.

[52] Montgomery, D., Runger, G., 2010. Applied statistics and probability for engineers. Wiley.

[53] Nguyen, N., 2008. A multiscale reduced-basis method for parametrized elliptic partial differential equations with multiple scales. Journal of Computational Physics 227 (23), 9807–9822.

[54] Nguyen, N., Patera, A., Peraire, J., 2008. A best points interpolation method for efficient approximation of parametrized functions. Int. J. Numer. Meth. Engng 73, 521–543.

[55] Quarteroni, A., Sacco, R., Saleri, F., 2000. Numerical Mathematics. Springer, New York.

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

[57] Rozza, G., 2009. Reduced basis methods for stokes equations in domains with non-affine parameter dependence. Computing and Visualization in Science 12 (1), 23–35.

[58] Ryckelynck, D., 2005. A priori hyperreduction method: an adaptive approach. Journal of computational physics 202 (1), 346–366.

[59] Ryckelynck, D., 2009. Hyper-reduction of mechanical models involving internal variables. International Journal for Numerical Methods in Engineering 77 (1), 75–89.

[60] Salomon, D., 2004. Data compression: the complete reference. Springer-Verlag New York Incorporated.

[61] Simo, J. C., Hughes, T. J. R., 1998. Computational inelasticity. Springer, New York.

[62] Yvonnet, J., He, Q., 2007. The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains. Journal of Computational Physics 223 (1), 341–368.

B.2 Paper #2

Title: Continuum approach to computational multiscale modeling of propagting fracture.
Authors:

- **J. Oliver**: Professor of Continuum Mechanics and Structural analysis at the Escola Tecnica Superior d'Enginyers de Camins, Canals i Ports (Civil Engineering School) of the Universitat Politècnica de Catalunya (Technical University of Catalonia BarcelonaTech). Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **E. Roubin**: Maître de conférence at the 3SR (Sols, Solides, Structures et Risques) and the IUT DGGC in Grenoble.
- **A. E. Huespe**: Professor of Mechanics at the Faculty of Chemical Engineering, Dept. of Materials, National University of Litoral, Santa Fe, Argentina. Independent researcher of Conicet at CIMEC (Centro de Investigaciones en Mecánica Computacional), National University of Litoral (UNL).
- **J. A. Hernández**: Assistant Professor of Structural Engineering and Strength of Materials at the School of Industrial and Aeronautic Engineering of Terrassa, of the Technical University of Catalonia. Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).

Journal of Computer Methods in Applied Mechanics and Engineering
Editors: Thomas J.R. Hughes, J. Tinsley Oden, Manolis Papadrakakis
ISSN: 0045-7825
Elsevier Editors
http://dx.doi.org/10.1016/j.cma.2015.05.012
Link to Publisher

Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 294 (2015) 384–427
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Continuum approach to computational multiscale modeling of propagating fracture

J. Oliver[a,b,*], M. Caicedo[b], E. Roubin[b], A.E. Huespe[b,c], J.A. Hernández[a,b]

[a] E.T.S. d'Enginyers de Camins, Canals i Ports, Technical University of Catalonia (BarcelonaTech), Spain
[b] Centre Internacional de Metodes Numerics en Enyinyeria (CIMNE), Campus Nord UPC, Edifici C-1, c/Jordi Girona 1-3, 08034 Barcelona, Spain
[c] CIMEC-UNL-CONICET, Güemes 3450, Santa Fe, Argentina

Available online 9 July 2015

Abstract

A new approach to two-scale modeling of propagating fracture, based on computational homogenization (FE^2), is presented. The specific features of the approach are: (a) a continuum setting for representation of the fracture at both scales based on the Continuum Strong Discontinuity Approach (CSDA), and (b) the use, for the considered non-smooth (discontinuous) problem, of the same computational homogenization framework than for classical smooth cases. As a key issue, the approach retrieves a characteristic length computed at the lower scale, which is exported to the upper one and used therein as a regularization parameter for a propagating strong discontinuity kinematics. This guarantees the correct transfer of fracture energy between scales and the proper dissipation at the upper scale. Representative simulations show that the resulting formulation provides consistent results, which are objective with respect to size and bias of the upper-scale mesh, and with respect to the size of the lower-scale RVE/failure cell, as well as the capability to model propagating cracks at the upper scale, in combination with crack-path-field and strain injection techniques. The continuum character of the approach confers to the formulation a minimal intrusive character, with respect to standard procedures for multi-scale computational homogenization.

Keywords: Multi-scale; Fracture; Propagating failure; Continuum Strong Discontinuity Approach (CSDA)

1. Introduction

Multi-scale computational modeling of solids, aiming at improving the predictive capabilities of mechanical models accounting for the description of the material at several scales, is a subject of increasing interest. A number of analytical and computational strategies have been developed in the past considering the description of the constitutive material at different scales, [1–14]. In most of them, multiscale description of the material itself

* Corresponding author at: E.T.S. d'Enginyers de Camins, Canals i Ports, Technical University of Catalonia (BarcelonaTech), Spain.
E-mail addresses: xavier.oliver@upc.edu (J. Oliver), mcaicedo@cimne.upc.edu (M. Caicedo), eroubin@cimne.upc.edu (E. Roubin), ahuespe@intec.unl.edu.ar (A.E. Huespe), jhortega@cimne.upc.edu (J.A. Hernández).

http://dx.doi.org/10.1016/j.cma.2015.05.012
0045-7825/

(e.g. computational material homogenization) and consequences and implications, on the overall modeling of the solid, of this specific description (computational multiscale modeling), are not generally considered in an integrated setting. In the context of a two scale (macro scale–micro/meso-scale) problem, computational homogenization of materials (FE^2) is generally regarded as a way of replacing, at the structural-scale, standard stress–strain phenomenological constitutive models equipped with internal variables, accounting for the micro/mesoscopic material morphology, by point wise overall stress–strain evaluations. The overall stresses are then obtained after solving an auxiliary problem, *the homogenization problem*, at the micro/meso-scale, in a manifold, the Representative Volume Element (RVE), endowed with a geometrical description of the material morphology. In turn, this RVE problem relies on some well-established paradigms, typically the classical Hill–Mandel principle [1,14,15] and the strain and stress homogenization concepts. More specifically: in this work we consider as starting point the modern *variational approach to multiscale homogenization* [16,17]. After this, the structural modeling proceeds at the macro/structural scale in a standard manner, with no further modifications.

This weak coupling makes sense for problems involving smooth – linear or nonlinear – material behavior, but the issue seems not to be so clear for non-smooth responses, like *material failure*, – typically fracture, de-cohesion, shear banding etc. – where the involved entities (strains, stresses, displacements) can be non-smooth or even unbounded [18]. For these non-smooth problems, two options emerge:

(a) Use the same homogenization paradigms than for smooth problems, with no specific modification. This approach has been strongly objected: even the existence of the RVE can be questioned, arguing that for fracture cases the material loses the statistical homogeneity [19], or, from another point of view, that the homogenized constitutive model lacks an internal length [20]. A crucial consequence of this issue is the lack of objectivity of the results with respect to the size of the RVE.
(b) Modify the homogenization paradigm towards a specific one for non-smooth problems. Selective RVE domain homogenization methods [21–25] or specific new homogenization paradigms [26,27] are possible ways to retrieve RVE-size objectivity of the results. However, sometimes this is done at the cost of a much higher complexity and intrusion in existing codes and loss of generality of the approach.

In this context, this work presents a new approach for computational multiscale analysis in non-smooth problems with the following features:

1. Extends the homogenization paradigms for smooth problems – typically the Hill–Mandel principle and the stress–strain homogenization procedures – to non-smooth problems, with no fundamental changes.
2. In both scales, a continuum (stress–strain) constitutive relationship is considered, instead of the most common discrete traction/separation-law, this contributing to provide a unified setting for smooth and non-smooth problems. This is achieved by resorting to the well-established Continuum Strong Discontinuity Approach (CSDA) to material failure [28,29,18].
3. As for the multiscale modeling issue, it involves a crucial additional entity: *a characteristic length*, which is point wise obtained from the geometrical features of the failure mechanism developed at the low scale. Introduction of a characteristic length in material homogenization schemes has been claimed as an ineluctable requirement for physical consistency [20], and some approaches to this subject can be found in recent works [30]. As a specific feature of the presented approach, for the non-smooth case this characteristic length is exported, in addition to the homogenized stresses and the tangent constitutive operator, to the macro-scale, and *considered the bandwidth of a propagating strain localization band, at that scale*.
4. Consistently with this characteristic length, a specific computational procedure, based on the *crack-path-field and strain injection techniques*, recently developed by the authors [31], is then used for modeling the onset and propagation of this localization band, at the macro-scale. This ensures the macro-scale mesh-size and micro-scale RVE-size objectivity of the results, and the proper energy dissipation at both scales.

In the remaining of this work a detailed description of the mechanical and computational elements of the proposed approach is presented. In Section 2, the multi-scale framework and the corresponding homogenization procedure, are described, whereas in Section 3 material failure propagation issues are addressed. Section 4 is devoted to present some representative numerical simulations to assess the performance of the proposed approach, and finally, in Section 5, some concluding remarks are stated.

Not to distract the reader's attention on issues that, though being crucial for the completeness of the work and the reproducibility of the results, are not in the core of the proposed approach, some appendices are added at the end of the

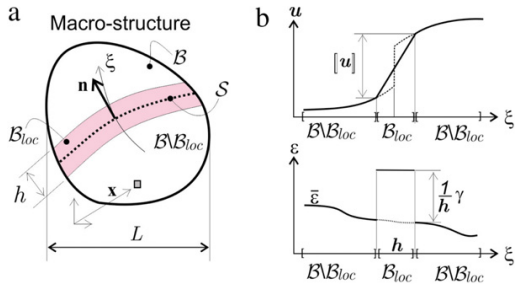


Fig. 1. Macroscopic (structural scale) body $\mathcal{B}$: (a) subdivision in a non-smooth domain, $\mathcal{B}_{loc}(t)$, and a smooth domain $\mathcal{B}\setminus\mathcal{B}_{loc}(t)$, (b) h-regularized, displacement and strain, discontinuity kinematics.

work. Appendix A, details a very simple technique for removing possible spurious unstable modes of the failure cell. Appendix B, refers to the extension to multi-scale problems of the *crack-path-field and strain injection techniques*, developed by the authors in previous works [31] for the purposes of modeling the intra-elemental propagation of strong discontinuities in mono-scale problems.

2. Multi-scale modeling setting

2.1. Macroscopic scale

Let us considered the body, $\mathcal{B}$, at the macroscopic (structural) scale (see Fig. 1). It is assumed that material points, $\mathbf{x}$, of the body belong, at the current time t, to either of the two different subdomains (see Fig. 1):

- Domain $\mathcal{B}\setminus\mathcal{B}_{loc}(t)$: the set of points at the macroscale, exhibiting smooth behavior at the current time. The infinitesimal strain field $\boldsymbol{\varepsilon}(\mathbf{x},t)$ is described, in rate form, as

$$\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) = (\boldsymbol{\nabla}_{\mathbf{x}}\otimes\dot{\mathbf{u}}(\mathbf{x},t))^{S} \equiv \boldsymbol{\nabla}^{S}\dot{\mathbf{u}}(\mathbf{x},t) \quad \forall\mathbf{x}\in\mathcal{B}\setminus\mathcal{B}_{loc}(t) \tag{1}$$

  where $\mathbf{u}(\mathbf{x},t)$ is the macroscale displacement field, t stands for the time or pseudo-time parameter and $(\cdot)^{S}$ stands for the symmetric counterpart of $(\cdot)$.
- Domain $\mathcal{B}_{loc}(t)$: the set of points exhibiting material failure and, therefore, some type of non-smooth behavior at the current time. The strain field at these points is assumed to be captured by a h-regularized strong/weak discontinuity kinematics, h being the width of the corresponding strain localization band (see Fig. 1(b)):

$$\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) = \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) + \frac{\kappa_{\mathcal{B}_{loc}}(\mathbf{x})}{h}\dot{\boldsymbol{\gamma}}(\mathbf{x},t) = \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) + \delta^{h}_{\mathcal{S}}(\mathbf{x})\dot{\boldsymbol{\gamma}}(\mathbf{x},t) \quad \forall\mathbf{x}\in\mathcal{B}_{loc}(t) \tag{2}$$

  where $\bar{\boldsymbol{\varepsilon}}(\mathbf{x},t)$ stands for regular (smooth) counterpart of the strain, $\kappa_{\mathcal{B}_{loc}}$ is a colocation (characteristic) function on $\mathcal{B}_{loc}$ ($\kappa_{\mathcal{B}_{loc}}(\mathbf{x}) = 1\ \forall\mathbf{x}\in\mathcal{B}_{loc}$; $\kappa_{\mathcal{B}_{loc}}(\mathbf{x}) = 0\ \forall\mathbf{x}\notin\mathcal{B}_{loc}$), so that the term $\delta^{h}_{\mathcal{S}}(\mathbf{x})$ becomes a h-regularized Dirac's delta function shifted to the center-line, $\mathcal{S}(t)$ (the macroscopic discontinuity-path at the current time), of band $\mathcal{B}_{loc}(t)$ (see Fig. 1(a)). Thus, in Eq. (2), the term $\delta^{h}_{\mathcal{S}}(\mathbf{x})\dot{\boldsymbol{\gamma}}(\mathbf{x}_{\mathcal{S}},t)$ is the non-smooth (discontinuous and h-regularized) localized counterpart of the strains; a space-discontinuous symmetric, second order tensor for the *weak-discontinuity case*. For the *strong-discontinuity case* – $\boldsymbol{\varepsilon}(\mathbf{x},t)$ stemming from a discontinuous displacement field – it takes the following specific format of a *rank-one tensor*, in terms of the macroscopic displacement jump at $\mathcal{S}(t)$, $\boldsymbol{\beta}(\mathbf{x}_{\mathcal{S}},t)$, and the normal to the discontinuity path, $\mathbf{n}(\mathbf{x}_{\mathcal{S}})$:

$$\begin{cases} \dot{\gamma}(\mathbf{x}_{\mathcal{S}},t) = (\underbrace{\dot{\boldsymbol{\beta}}(\mathbf{x}_{\mathcal{S}},t)}_{\dot{\xi}\mathbf{m}}\otimes\mathbf{n}(\mathbf{x}_{\mathcal{S}}))^{S} = \dot{\xi}(\mathbf{x},t)(\mathbf{m}\otimes\mathbf{n})^{S} \\ \|\mathbf{m}\| = \|\mathbf{n}\| = 1; \quad \left\|\dot{\boldsymbol{\beta}}\right\| = \dot{\xi} \end{cases} \tag{3}$$
$$\Rightarrow \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) = \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) + \delta^{h}_{\mathcal{S}}(\mathbf{x})\dot{\xi}(\mathbf{x},t)(\mathbf{m}\otimes\mathbf{n})^{S}$$

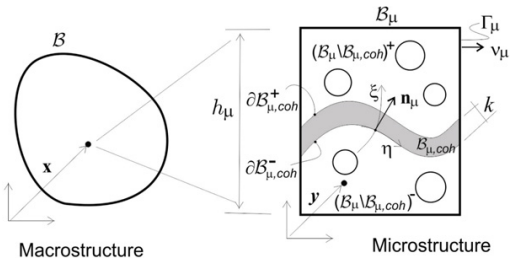


Fig. 2. Multiscale model: failure-cell with cohesive bands.

where notation $(\mathbf{m} \otimes \mathbf{n})^S \equiv \mathbf{m} \otimes^S \mathbf{n} = \frac{1}{2}(\mathbf{m} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{m})$ is used to denote the symmetric part of the tensor product $\mathbf{m} \otimes \mathbf{n}$. Transition, for a given material point, $\mathbf{x}$, from the weak discontinuity kinematics, in Eq. (2), to the strong discontinuity kinematics, in Eq. (3), is explained in detail in Section 2.5.

2.2. Microscopic scale

Let us now assume that the non-smooth behavior at the structural scale is produced, in turn, by a non-smooth behavior at the microscale level arising from some type of material failure at this lower scale. Therefore, the next issue is to endow the low-scale model with mechanisms to capture the onset and propagation of this material failure: typically discontinuous micro-displacement fields. For the sake of simplicity, and without prejudice to introduce further, more complex, options, it is considered that the microstructure is able to capture some dominant failure mechanisms of the material. For this purpose, a micro-failure-cell,[1] $\mathcal{B}_\mu(\mathbf{x})$, of typical size h_μ, is considered to exist at every material point $\mathbf{x} \in \mathcal{B}$ (see Fig. 2). It accounts for the material morphology at the lower scale (voids, inclusions etc.) as in regular homogenization procedures, but in addition, it is endowed with a *set of cohesive bands*, $\mathcal{B}_{\mu,coh} \subset \mathcal{B}_\mu$, of very small bandwidth ($k \ll h_\mu$), whose position and other geometric properties (typically the normal $\mathbf{n}_\mu$, see Fig. 2) are predefined. Activation (de-cohesion) at the current time, t, of a number of those bands, defines the current *subset of active bands*, $\mathcal{B}_{\mu,act}(t) \subset \mathcal{B}_{\mu,coh} \subset \mathcal{B}_\mu$ which constitutes the *activated microscopic failure mechanism*, at the current time and for the considered point $\mathbf{x} \in \mathcal{B}$ at the macro-scale.

Remark 2.1. In principle, there is no intrinsic limitation on the number of the cohesive bands at the failure cell. On one hand, their number and spatial position have to be sufficient to capture the dominant material failure mechanisms at the macro-scale, and, on the other hand, the associated computational cost clearly sets a limitation on the number of such bands. Therefore, an appropriated balance of both aspects has to be achieved. In the examples in Section 4, results obtained with relatively crude representations of the material morphology and failure mechanisms at the micro-scale provide a first approach to this issue.

In consequence, the following domains at the microscale are considered (see Fig. 2):

- Domain $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh}$: the set of points, $\mathbf{y}$, out of the cohesive bands. They are imposed to exhibit a smooth behavior described by a *continuum elastic model*, typically:

$$\boldsymbol{\sigma}_\mu(\mathbf{x}, \mathbf{y}, t) = \boldsymbol{\Sigma}^{elastic}(\boldsymbol{\varepsilon}_\mu) \equiv \boldsymbol{C}_\mu(\mathbf{y}) : \boldsymbol{\varepsilon}_\mu(\mathbf{x}, \mathbf{y}, t) \quad \forall \mathbf{x} \in \mathcal{B};\ \forall \mathbf{y} \in \mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh} \tag{4}$$

  where $\boldsymbol{\sigma}_\mu$ and $\boldsymbol{\varepsilon}_\mu$ stand, respectively, for the micro-stresses and micro-strains at the micro-scale point, $\mathbf{y}$, of the failure cell corresponding to the macro-scale point, $\mathbf{x}$, and $\boldsymbol{C}_\mu(\mathbf{y})$ is the micro elastic constitutive tensor.
- Domain $\mathcal{B}_{\mu,coh}$: the set of microscale cohesive bands. As for the material behavior, in this case one has to make a distinction of two cases:
  (a) The failure cell, $\mathcal{B}_\mu(\mathbf{x})$, is associated to a *non-smooth* material point at the macro-scale ($\mathbf{x} \in \mathcal{B}_{loc}(t)$). Without prejudice of using any other inelastic constitutive model, in the remaining of this work an *isotropic continuum*

[1] From now on, sub-index $(\cdot)_\mu$ will denote entities associated to the micro-scale.

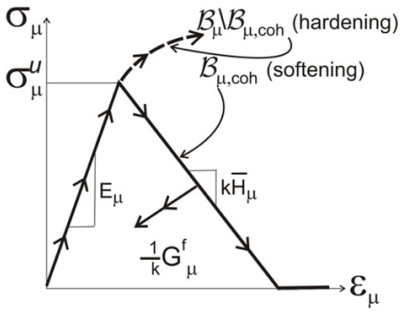


Fig. 3. Isotropic continuum damage model at the microscale (uniaxial version).

Box 2.1.
Isotropic continuum tensile-damage model with strain-softening

$$\boldsymbol{\sigma}_\mu = (1-d_\mu)\boldsymbol{C}_\mu : \boldsymbol{\varepsilon}_\mu = \frac{q_\mu}{r_\mu}\underbrace{\boldsymbol{C}_\mu : \boldsymbol{\varepsilon}_\mu}_{\overline{\boldsymbol{\sigma}}_\mu};$$

$$d_\mu(r_\mu) = 1-(q_\mu(r_\mu)/r_\mu) \geq 0;$$

$$\dot{r}_\mu = \lambda, \qquad r_\mu|_{t=0} = r_0 = \sigma_\mu^u/\sqrt{E_\mu};$$

$$\begin{cases} \dot{q}_\mu = k\overline{H}(r_\mu)\dot{r}_\mu; \quad q_\mu \geq 0; \quad q|_{t=0} = r_0 \\ k = \textit{regularization parameter (bandwidth of localized strains)} \end{cases}$$

$$\underbrace{\overline{H}(r_\mu)}_{\substack{\textit{Intrinsic}\\ \textit{softening}\\ \textit{modulus}}} = \begin{cases} -\dfrac{r_0^2}{2G_\mu^f}; & \textit{bilinear softening} \\ -{r_0}^2/G_\mu^f \exp\left(-\dfrac{kr_0}{G_\mu^f}(r_\mu - r_0)\right); & \textit{exponential softening} \end{cases}$$

$$\begin{cases} g(\boldsymbol{\varepsilon}_\mu, r_\mu) \equiv \sqrt{\overline{\boldsymbol{\sigma}}_\mu^+ : \boldsymbol{\varepsilon}_\mu} - r_\mu; \qquad \overline{\boldsymbol{\sigma}}_\mu^+ = \displaystyle\sum_{i=1}^{3} \left\langle \overline{\sigma}_{\mu,i} \right\rangle \mathbf{e}_i \otimes \mathbf{e}_i \\ \overline{\sigma}_{\mu,i} \rightarrow \textit{principal stress } \text{“}i\text{”} \\ \mathbf{e}_i \rightarrow \textit{principal direction } \text{“}i\text{”} \end{cases}$$

$$\dot{\lambda} \geq 0; \quad g \leq 0; \quad \dot{\lambda} g = 0; \quad \begin{matrix} \textit{Karush–Kuhn–Tucker} \\ \textit{(loading/unloading) conditions} \end{matrix}$$

Material parameters: ultimate stress: σ_μ^u; Young's modulus E_μ,
Poisson ratio, ν_μ, fracture energy: G_μ^f.
$\overline{\boldsymbol{\sigma}}_\mu^+$ is the positive part of the microscopic effective stress, d_μ is the damage variable and r_μ and q_μ are, respectively, the strain-like and stress-like internal variables.

damage model, exhibiting k-*regularized strain softening inelasticity* only for tensile stresses – *tensile-damage continuum damage model* [32] – will be considered (see also references [33] and [34]). It reads:

$$\boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y},t) = \boldsymbol{\Sigma}^{inelastic}(\boldsymbol{\varepsilon}_\mu, r_\mu) \equiv (1-d_\mu(r_\mu(\mathbf{y},t)))\,\boldsymbol{C}_\mu : \boldsymbol{\varepsilon}_\mu(\mathbf{x},\mathbf{y},t)$$
$$\mathbf{x} \in \mathcal{B}_{loc}(t), \forall \mathbf{y} \in \mathcal{B}_{\mu,coh} \tag{5}$$

where $d_\mu \in [0,1]$ is the, scalar, damage variable whose evolution is described in terms of the internal variable r_μ. More details about the model are given in Box 2.1 (see Fig. 3).

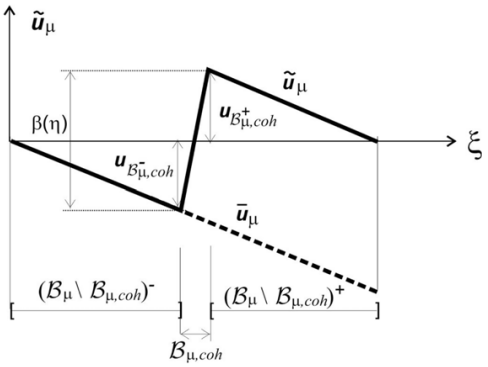


Fig. 4. Micro-displacement fluctuation fields.

(b) The failure cell, $\mathcal{B}_\mu(\mathbf{x})$, is associated to a *smooth* material point at the macro-scale, $(\mathbf{x} \in \mathcal{B} \setminus \mathcal{B}_{loc}(t))$. In this case, the inelastic model, in Eq. (5) is enforced to behave instantaneously elastically at the cohesive bands domain, $\mathcal{B}_{\mu,coh}$, i.e.:

$$\begin{aligned} &\boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y},t) = \boldsymbol{\Sigma}^{elastic}_{inst}(\boldsymbol{\varepsilon}_\mu) = \boldsymbol{C}^{elastic}_{inst} : \boldsymbol{\varepsilon}_\mu(\mathbf{x},\mathbf{y},t) \\ &\boldsymbol{C}^{elastic}_{inst} = (1-d_\mu)\,\boldsymbol{C}_\mu \quad \text{with } \dot{d}_\mu = 0 \\ &\forall \mathbf{x} \in \mathcal{B} \setminus \mathcal{B}_{loc}(t); \quad \forall \mathbf{y} \in \mathcal{B}_{\mu,coh}. \end{aligned} \tag{6}$$

Remark 2.2. Notice that, according to the previous setting, *the same RVE morphology is considered at all* $\mathcal{B}_\mu(\mathbf{x})$, both for $\mathbf{x} \in \mathcal{B}_{loc}(t)$ (failure cell) or $\mathbf{x} \in \mathcal{B} \setminus \mathcal{B}_{loc}(t)$, the only difference being the considered constitutive behavior at the cohesive bands, $\mathcal{B}_{\mu,coh}$, defined in Eqs. (5) and (6).

Let us now consider the micro displacement field, $\mathbf{u}_\mu$, at the cell described as

$$\mathbf{u}_\mu(\mathbf{x},\mathbf{y},t) = \mathbf{u}(\mathbf{x},t) + \boldsymbol{\varepsilon}(\mathbf{x},t)\cdot\mathbf{y} + \tilde{\mathbf{u}}_\mu(\mathbf{y},t) \tag{7}$$

where $\mathbf{u}(\mathbf{x},t)$ and $\boldsymbol{\varepsilon}(\mathbf{x},t)$ are, respectively, the macroscale displacements and strains at point $\mathbf{x}$ in $\mathcal{B}$, and $\tilde{\mathbf{u}}_\mu$ are the microscale displacement fluctuations. Considering a local coordinate system (ξ,η) aligned with the domain $\mathcal{B}_{\mu,coh}$ (see Fig. 2), the *smooth* part of the micro-fluctuation field, $\bar{\mathbf{u}}_\mu$, is defined as:

$$\begin{aligned} &\bar{\mathbf{u}}_\mu(\xi,\eta,t) = \tilde{\mathbf{u}}_\mu(\xi,\eta,t) - \mathcal{H}_{\mathcal{B}_{\mu,coh}}(\xi)\,\boldsymbol{\beta}_\mu(\eta,t) \qquad \text{(a)} \\ &\mathcal{H}_{\mathcal{B}_{\mu,coh}}(\xi) = \begin{cases} 0 & \forall \mathbf{y} \in (\mathcal{B} \setminus \mathcal{B}_{\mu,coh})^- \\ \dfrac{\xi}{k} & \forall \mathbf{y} \in \mathcal{B}_{\mu,coh} \\ 1 & \forall \mathbf{y} \in (\mathcal{B} \setminus \mathcal{B}_{\mu,coh})^+ \end{cases} \qquad \text{(b)} \\ &\boldsymbol{\beta}_\mu(\eta(\mathbf{y}),t)\big|_{\mathbf{y}\in\mathcal{B}_{\mu,coh}} \equiv [\![\,\tilde{\mathbf{u}}_\mu(\xi,\eta,t)\,]\!]^+_- \qquad \text{(c)} \end{aligned} \tag{8}$$

where $\mathcal{H}_{\mathcal{B}_{\mu,coh}}(\xi)$ is the k-regularized Heaviside function shifted to $\mathcal{B}_{\mu,coh}$, and $\boldsymbol{\beta}_\mu(\xi,\eta)$ is a (smooth) function arbitrarily defined excepting for the restriction in Eq. (8)(c).

Notation. $[\![(\cdot)(\xi,\eta)\,]\!]^+_- \equiv (\cdot)(\xi,\eta)|_{\xi=k} - (\cdot)(\xi,\eta)|_{\xi=0}$ stands for the *apparent jump* of $(\cdot)(\xi,\eta)$ between both sides of the cohesive band. From Eqs. (8) it turns out,

$$[\![\,\bar{\mathbf{u}}_\mu\,]\!]^+_- = [\![\,\tilde{\mathbf{u}}_\mu\,]\!]^+_- - \underbrace{[\![\,\mathcal{H}_{\mathcal{B}_{\mu,coh}}\,]\!]^+_-}_{=1}[\![\tilde{\mathbf{u}}_\mu]\!]^+_- = \mathbf{0} \tag{9}$$

and, therefore, function $\bar{\mathbf{u}}_\mu$ in Eq. (8) is smooth. Finally, from Eq. (8)(a),

$$\tilde{\mathbf{u}}_\mu(\mathbf{y},t) = \underbrace{\bar{\mathbf{u}}_\mu(\mathbf{y},t)}_{smooth} + \mathcal{H}_{\mathcal{B}_{\mu,k}}(\mathbf{y})\,\boldsymbol{\beta}_\mu(\mathbf{y},t) \tag{10}$$

(See a sketch in Fig. 4.) Eq. (10) constitutes the displacement counterpart of a k-regularized strong discontinuity kinematics [28], and proves that the cohesive bands-approach, herein proposed for the description of material failure at the microstructure, is consistent with consideration of a k-regularized strong discontinuity at the cohesive domain $\mathcal{B}_{\mu,coh}$.

In this context, one can resort to the well-established connection of the continuum modeling of material failure – based on stress–strain constitutive equations equipped with strain softening – and the discrete cohesive fracture mechanics, established in the Continuum Strong Discontinuity Approach (CSDA) to material failure [35]. *This states the equivalence, in the limit $k \to 0$, of the proposed approach and the one based on cohesive lines endowed with traction–separation laws* [36].
From Eq. (10) one obtains,

$$\begin{aligned}
\boldsymbol{\nabla}^S \tilde{\mathbf{u}}_\mu(\mathbf{y},t) &= \boldsymbol{\nabla}^S \bar{\mathbf{u}}_\mu(\mathbf{y},t) + \mathcal{H}_{\mathcal{B}_{\mu,k}}\, \boldsymbol{\nabla}^S \boldsymbol{\beta}_\mu(\mathbf{y},t) + \delta^k_{\mathcal{S}_\mu}(\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y}_{\mathcal{S}_\mu},t) \\
&= \underbrace{\bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{y},t)}_{(bounded)} + \underbrace{\delta^k_{\mathcal{S}_\mu}(\mathbf{y},t)\,(\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y}_{\mathcal{S}_\mu},t)}_{(unbounded)} \qquad \text{(a)} \\
\bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{y},t) &= \boldsymbol{\nabla}^S \bar{\mathbf{u}}_\mu(\mathbf{y},t) + \mathcal{H}_{\mathcal{B}_{\mu,k}}\, \boldsymbol{\nabla}^S \boldsymbol{\beta}_\mu(\mathbf{y},t) \qquad \text{(b)} \\
\delta^k_{\mathcal{S}_\mu}(\mathbf{y},t) &= \frac{1}{k}\,\kappa_{\mathcal{B}_{\mu,coh}}(\mathbf{y},t) \qquad \text{(c)}
\end{aligned} \tag{11}$$

where $\kappa_{\mathcal{B}_{\mu,coh}}(\mathbf{y})$ is the characteristic (colocation) function in Eq. (2) so that $\delta^k_{\mathcal{S}_\mu}(\mathbf{y},t)$ becomes a k-regularized Dirac's delta function placed at the center line, $\mathcal{S}_\mu$, of $\mathcal{B}_{\mu,coh}$. From Eqs. (7) and (11), the micro-strains $\boldsymbol{\varepsilon}_\mu$ can be written in terms of the macro-strains, $\boldsymbol{\varepsilon}(\mathbf{x},t)$, and the micro-fluctuations, $\tilde{\mathbf{u}}_\mu(\mathbf{y},t)$, as follows[2]:

$$\boldsymbol{\varepsilon}_\mu(\mathbf{x},\mathbf{y}) = \boldsymbol{\varepsilon}(\mathbf{x}) + \nabla^s \tilde{\mathbf{u}}_\mu(\mathbf{y}) = \underbrace{\boldsymbol{\varepsilon}(\mathbf{x}) + \bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{y})}_{(bounded)} + \underbrace{\delta^k_{\mathcal{S}_\mu}\,(\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y})}_{(unbounded)}. \tag{12}$$

Eq. (12) displays the classical format of a strain field in a k-regularized strong discontinuity kinematics, where the second term at the right-hand side becomes unbounded in the limit $k \to 0$.

2.3. Homogenization procedure

At this point, the RVE/failure-cell homogenization setting is tackled. It has to be emphasized that, in the proposed approach, *the homogenization setting is the same for smooth, $\mathbf{x} \in \mathcal{B} \setminus \mathcal{B}_{loc}(t)$, and non-smooth, $\mathbf{x} \in \mathcal{B}_{loc}(t)$*, material points of the macroscale and it *coincides with the one considered in the variational homogenization approach in smooth settings* [16]. Therefore, it relies on the following paradigms:

1. *Micro-strain homogenization*.
   According to this, the average value, over $\mathcal{B}_\mu$, of the micro-strains, $\boldsymbol{\varepsilon}_\mu(\mathbf{x},\mathbf{y})$, in Eq. (12) has to be equal to the macro strains, $\boldsymbol{\varepsilon}(\mathbf{x})$ i.e.

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu(\mathbf{x})} \boldsymbol{\varepsilon}_\mu(\mathbf{x},\mathbf{y})\, d\mathcal{B} \tag{13}$$

   were $\Omega_\mu = meas(\mathcal{B}_\mu)$ stands for the area/volume of the failure cell. Replacing Eq. (12) into (13) results in:

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \boldsymbol{\varepsilon}(\mathbf{x}) + \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu(\mathbf{x})} \nabla^s \tilde{\mathbf{u}}_\mu(\mathbf{y})\, d\mathcal{B} \tag{14}$$

   which is satisfied if:

$$\int_{\mathcal{B}_\mu(\mathbf{x})} \nabla^s \tilde{\mathbf{u}}_\mu(\mathbf{y})\, d\mathcal{B} = \int_{\Gamma_\mu(\mathbf{x})} (\tilde{\mathbf{u}}_\mu(\mathbf{y}) \otimes \mathbf{v}_\mu(\mathbf{y}))^S\, d\Gamma = \mathbf{0} \tag{15}$$

[2] From now on, dependence on time, t, will be omitted excepting when strictly necessary.

where $\mathbf{v}_\mu(\mathbf{y})$ stands for the outward normal to the failure cell, and $\Gamma_\mu \equiv \partial\mathcal{B}_\mu$ is its boundary (see Fig. 2). Thus, we define *the space, $\boldsymbol{\mathcal{V}}_\mu$, of kinematically admissible micro-displacement fluctuations*, as

$$\boldsymbol{\mathcal{V}}_\mu = \left\{ \tilde{\mathbf{u}}_\mu \middle| \int_{\Gamma_\mu} (\tilde{\mathbf{u}}_\mu \otimes \mathbf{v}_\mu)^S d\Gamma = \mathbf{0} \right\}. \tag{16}$$

Notice that the boundary conditions in Eq. (16) *coincide* exactly, for the proposed non-smooth setting, *with the ones considered in the variational homogenization approach* [16] in smooth settings, and termed *minimal kinematic constraint conditions*.

2. *Hill–Mandel variational principle*.

The balance of the macroscopic and microscopic stress power reads

$$\begin{aligned} \boldsymbol{\sigma}(\mathbf{x}) : \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) &= \frac{1}{\Omega_\mu} \int_{\mathcal{B}_\mu(\mathbf{x})} \boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y}) : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{x},\mathbf{y}) d\mathcal{B} = \\ &= \frac{1}{\Omega_\mu} \int_{\mathcal{B}_\mu(\mathbf{x})} \boldsymbol{\sigma}_\mu : (\dot{\boldsymbol{\varepsilon}}(\mathbf{x}) + \boldsymbol{\nabla}^S \dot{\tilde{\mathbf{u}}}_\mu(\mathbf{y})) d\mathcal{B}; \quad \forall \dot{\boldsymbol{\varepsilon}},\ \dot{\tilde{\mathbf{u}}}_\mu \in \boldsymbol{\mathcal{V}}_\mu \end{aligned} \tag{17}$$

which, for arbitrary variations $\dot{\boldsymbol{\varepsilon}} \neq \mathbf{0};\ \dot{\tilde{\mathbf{u}}}_\mu \equiv \mathbf{0}$ yields

$$\boldsymbol{\sigma}(\mathbf{x}) = \frac{1}{\Omega_\mu} \int_{\mathcal{B}_\mu(\mathbf{x})} \boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y}) d\mathcal{B} \tag{18}$$

and, for arbitrary variations of $\dot{\boldsymbol{\varepsilon}} \equiv \mathbf{0};\ \dot{\tilde{\mathbf{u}}}_\mu \neq \mathbf{0}$, leads to the self-equilibrated stress equation in variational form:

$$\int_{\mathcal{B}_\mu} \nabla^S \delta\tilde{\mathbf{u}}_\mu : \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\nabla}^S \tilde{\mathbf{u}}_\mu) d\mathcal{B} = \mathbf{0}; \quad \forall\ \delta\tilde{\mathbf{u}}_\mu \in \boldsymbol{\mathcal{V}}_\mu. \tag{19}$$

In the context of a finite element approximation, Eq. (16) is solved as

$$\int_{\mathcal{B}_\mu} \mathbf{B}^T \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \mathbf{B}\tilde{\mathbf{d}}_\mu) d\mathcal{B} = \mathbf{0} \tag{20}$$

where $\mathbf{B}(\mathbf{y})$ is the classical deformation matrix [37] and $\tilde{\mathbf{d}}_\mu$ are the nodal values of the fluctuations, compatible with the boundary condition in Eq. (16).

Again, it should be noticed that Eqs. (18)–(20) *have the same format than for purely smooth cases*.

2.4. Theoretical "a priori" analysis of the resulting homogenized constitutive model at the macroscale

Let us now perform a theoretical analysis by examining the consequences of applying the *homogenization procedure*s, based on the Hill–Mandel principle, to the totality of the failure cell, $\mathcal{B}_\mu$ (see Eqs. (13)–(19), when regularized strong discontinuities are present at the micro-scale being captured by the cohesive bands $\mathcal{B}_{\mu,coh}$, as displayed in Eqs. (12)). It is precisely in this feature (application of the homogenization in the complete domain of the failure cell) that the proposed approach differs from other approaches, as the one in [38].

Let us consider the failure cell displayed in Fig. 5. For the sake of generality, we will consider the cell composed of a matrix and inclusions (or voids), in addition to a number of cohesive bands defining the set $\mathcal{B}_{\mu,coh}$. Among them, there are n_{active} cohesive bands $\mathcal{B}^{(i)}_{\mu,act};\ i = 1 \ldots n_{active}$, which are in an inelastic softening state, defining a specific failure mode, $\mathcal{B}_{\mu,act} \subset \mathcal{B}_{\mu,coh}$, at the micro-cell.[3] Let us also denote by $\mathcal{S}^{(i)}_\mu$ the midline in the cohesive band i (see Fig. 5).

Therefore, it is fulfilled

$$\mathcal{B}_{\mu,act} = \bigcup_{i=1}^{i=n_{active}} \mathcal{B}^{(i)}_{\mu,act}; \qquad \mathcal{S}_\mu = \bigcup_{i=1}^{i=n_{active}} \mathcal{S}^{(i)}_\mu \tag{21}$$

[3] The remaining cohesive bands are discarded for the purposes of this analysis.

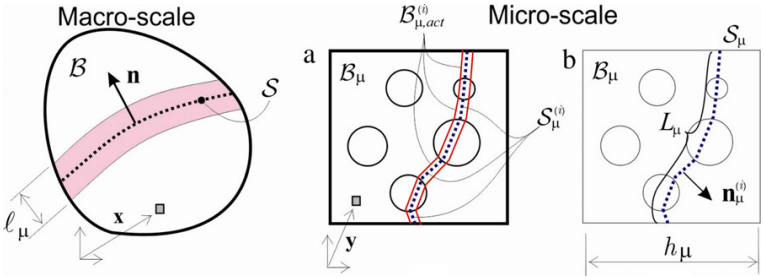


Fig. 5. Multiscale model: (a) failure cell with inclusions and cohesive bands; (b) geometrical characterization of the failure mode at the microscale.

where $\mathcal{S}_\mu$ can be regarded as the active failure/crack path at the micro-scale. From Eqs. (10) and (11), one can write

$$\begin{aligned} \tilde{\mathbf{u}}_\mu(\mathbf{y}) &= \underbrace{\bar{\mathbf{u}}_\mu(\mathbf{y})}_{smooth} + \mathcal{H}_{\mathcal{B}_\mu,k}(\mathbf{y})\,\boldsymbol{\beta}_\mu(\mathbf{y}) \qquad (a) \\ \boldsymbol{\nabla}^s\tilde{\mathbf{u}}_\mu(\mathbf{y}) &= \underbrace{\bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{y})}_{(bounded)} + \underbrace{\delta^k_{\mathcal{S}_\mu}(\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y})}_{(unbounded)}. \qquad (b) \end{aligned} \tag{22}$$

In this context, the following lemma can be stated:

Lemma.

$$\int_{\mathcal{B}_\mu} \bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{x},\mathbf{y})d\mathcal{B}_\mu = -\int_{\mathcal{S}_\mu} (\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y})d\mathcal{S}. \tag{23}$$

Proof. Replacement of Eq. (22) into Eq. (15) yields,

$$\begin{aligned} &\int_{\mathcal{B}_\mu} \nabla^s\tilde{\mathbf{u}}_\mu(\mathbf{y})d\mathcal{B} = \mathbf{0} \\ &\Rightarrow \int_{\mathcal{B}_\mu} \bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{y})d\mathcal{B} + \int_{\mathcal{B}_\mu} \delta^k_{\mathcal{S}_\mu}(\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y})d\mathcal{B} = \\ &\quad = \int_{\mathcal{B}_\mu} \bar{\boldsymbol{\varepsilon}}_\mu(\mathbf{y})d\mathcal{B} + \int_{\mathcal{S}_\mu} (\boldsymbol{\beta}_\mu \otimes \mathbf{n}_\mu)^S(\mathbf{y})\,d\mathcal{S} = \mathbf{0} \end{aligned} \tag{24}$$

where condition, in the limit ($k \to 0$),

$$\int_{\mathcal{B}_{\mu,coh}} \delta^k_{\mathcal{S}_\mu}(\cdot)(\mathbf{y})\,d\mathcal{B} \approx \int_{\mathcal{S}_\mu} (\cdot)(\mathbf{y})\,d\mathcal{S} \tag{25}$$

for sufficient regular functions, $(\cdot)(\mathbf{y})$ [39], has been considered. Eq. (24) proofs trivially Lemma (23).

Let us now compute the rate of homogenized stresses in Eq. (18), in the limit $k \to 0$ as

$$\begin{aligned} \dot{\boldsymbol{\sigma}}(\mathbf{x}) &= \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \dot{\boldsymbol{\sigma}}_\mu d\mathcal{B} = \frac{1}{\Omega_\mu}\left[\int_{\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh}} \dot{\boldsymbol{\sigma}}_\mu d\mathcal{B} + \underbrace{\int_{\mathcal{B}_{\mu,coh}} \dot{\boldsymbol{\sigma}}_\mu d\mathcal{B}}_{=0 \text{ as } k\to 0}\right] \\ &= \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh}} \boldsymbol{C}_\mu : \dot{\boldsymbol{\varepsilon}}_\mu d\mathcal{B}. \end{aligned} \tag{26}$$

In Eq. (26), the integral on the domain $\mathcal{B}_{\mu,coh}$ is neglected since its kernel, $\dot{\boldsymbol{\sigma}}_\mu$, is bounded whereas the measure of the integration domain tends to zero as $k \to 0$ (see Fig. 2). In addition, the elastic constitutive model in $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh}$ (see Eq. (4)) has been considered.

Now inserting Eq. (12), for $\mathbf{y} \in \mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh}$, into Eq. (26) yields

$$\dot{\boldsymbol{\sigma}}(\mathbf{x},t) \simeq \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,coh}} \underbrace{\boldsymbol{C}_\mu : (\dot{\boldsymbol{\varepsilon}} + \dot{\bar{\boldsymbol{\varepsilon}}}_\mu)}_{bounded}\, d\mathcal{B} \overset{k\to 0}{\simeq} \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{C}_\mu(\mathbf{y}) : (\dot{\boldsymbol{\varepsilon}}(\mathbf{x}) + \dot{\bar{\boldsymbol{\varepsilon}}}_\mu(\mathbf{x},\mathbf{y}))d\mathcal{B} \tag{27}$$

where, again, the smooth character of the smooth strains $\dot{\boldsymbol{\varepsilon}}$ and $\dot{\bar{\boldsymbol{\varepsilon}}}$ has been considered. Eq. (27) can be rephrased in terms of the *average value of the micro-elastic constitutive tensor*, $\overline{\boldsymbol{\mathcal{C}}}$, as:

$$\begin{aligned}
\dot{\boldsymbol{\sigma}}(\mathbf{x},t) &= \underbrace{\frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{C}_\mu(\mathbf{y})d\mathcal{B}}_{\overline{\boldsymbol{\mathcal{C}}}} : \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) + \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{C}_\mu(\mathbf{y}) : \dot{\bar{\boldsymbol{\varepsilon}}}_\mu(\mathbf{x},\mathbf{y})\, d\mathcal{B} \\
&= \overline{\boldsymbol{\mathcal{C}}} : \dot{\boldsymbol{\varepsilon}} + \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \overline{\boldsymbol{\mathcal{C}}} : \dot{\bar{\boldsymbol{\varepsilon}}}_\mu d\mathcal{B} + \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} (\boldsymbol{C}_\mu - \overline{\boldsymbol{\mathcal{C}}}) : \dot{\bar{\boldsymbol{\varepsilon}}}_\mu\, d\mathcal{B} \qquad \text{(a)} \\
\overline{\boldsymbol{\mathcal{C}}} &\equiv \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{C}_\mu(\mathbf{y})d\mathcal{B}. \qquad \text{(b)}
\end{aligned} \tag{28}$$

Now, replacing Eq. (23) into Eq. (28) yields,

$$\dot{\boldsymbol{\sigma}}(\mathbf{x},t) = \overline{\boldsymbol{\mathcal{C}}} : \dot{\boldsymbol{\varepsilon}} - \overline{\boldsymbol{\mathcal{C}}} : \frac{1}{\Omega_\mu}\underbrace{\int_{\mathcal{S}_\mu} (\dot{\boldsymbol{\beta}}_\mu \otimes \mathbf{n}_\mu)^S\, d\mathcal{S}}_{=L_\mu(\overline{\dot{\boldsymbol{\beta}}_\mu \otimes^s \mathbf{n}_\mu})_{\mathcal{S}_\mu}} + \underbrace{\frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} (\boldsymbol{C}_\mu - \overline{\boldsymbol{\mathcal{C}}}) : \dot{\bar{\boldsymbol{\varepsilon}}}_\mu d\mathcal{B}}_{\equiv \overline{\boldsymbol{\mathcal{C}}}:\dot{\boldsymbol{\chi}}(\mathbf{x},t)} \tag{29}$$

where $L_\mu = |\mathcal{S}_\mu|$ is the measure (length in 2D, area in 3D) of $\mathcal{S}_\mu$ and $\overline{(\cdot)}_{\mathcal{S}_\mu}$ stands for the average value of $(\cdot)$ in the activated macroscopic failure mechanism along the crack path $\mathcal{S}_\mu$.
Then, Eq. (29) can be rephrased giving rise to the homogenized constitutive model in Box 2.2.

Box 2.2.
Homogenized constitutive model at the macro-scale

$$\begin{cases} \text{homogenized} \\ \text{constitutive} \\ \text{equation (rate form)} \end{cases} \rightarrow \begin{cases} \dot{\boldsymbol{\sigma}}(\mathbf{x},t) = \overline{\boldsymbol{\mathcal{C}}} : [\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) + \dot{\boldsymbol{\chi}}(\mathbf{x},t) - \dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t)] \\ \qquad\qquad = \boldsymbol{\mathcal{C}}^{tang}_{hom}(\boldsymbol{\varepsilon}) : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) \end{cases} \qquad \text{(a)}$$

$$\begin{cases} \text{averaged elastic} \\ \text{constitutive} \\ \text{tensor} \end{cases} \rightarrow \quad \overline{\boldsymbol{\mathcal{C}}} \equiv \frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{C}_\mu(\mathbf{y})d\mathcal{B} \qquad \text{(b)}$$

$$\text{smooth internal strain} \rightarrow \begin{cases} \dot{\boldsymbol{\chi}}(\mathbf{x},t) = \dfrac{1}{\Omega_\mu}\displaystyle\int_{\mathcal{B}_\mu} (\overline{\boldsymbol{\mathcal{C}}}^{-1} : \boldsymbol{C}_\mu - \mathbf{I}) : \underbrace{\dot{\bar{\boldsymbol{\varepsilon}}}_\mu(\tilde{\mathbf{u}}_\mu)}_{=\mathcal{A}(\mathbf{y},\boldsymbol{\varepsilon}):\dot{\boldsymbol{\varepsilon}}}\, d\mathcal{B} \\ = \underbrace{\dfrac{1}{\Omega_\mu}\displaystyle\int_{\mathcal{B}_\mu} (\overline{\boldsymbol{\mathcal{C}}}^{-1} : \boldsymbol{C}_\mu - \mathbf{I}) : \mathcal{A}(\mathbf{y},\boldsymbol{\varepsilon}(\mathbf{x},t))d\mathcal{B}}_{=\boldsymbol{\mathcal{M}}(\boldsymbol{\varepsilon}(\mathbf{x},t))} : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) \qquad \text{(c)} \\ = \boldsymbol{\mathcal{M}}(\boldsymbol{\varepsilon}(\mathbf{x},t)) : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) \end{cases}$$

$$\begin{cases} \text{non-smooth} \\ \text{(localized)} \\ \text{internal strain} \end{cases} \rightarrow \begin{cases} \dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t) = \dfrac{1}{\ell_\mu}(\overline{\dot{\boldsymbol{\beta}}_\mu \otimes^s \mathbf{n}_\mu})_{\mathcal{S}_\mu} \\ \ell_\mu(\mathbf{x},t) \equiv \dfrac{\Omega_\mu}{L_\mu} \rightarrow \text{internal/characteristic length} \end{cases} \qquad \text{(d)}$$

(30)

In Eq. (30)(a), $\boldsymbol{\mathcal{C}}^{tang}_{hom}$, stands for the macroscale homogenized tangent constitutive operator, which can be pointwise obtained from a standard homogenization procedure, see [16]. In Eq. (30)(c) $\mathbf{I}$ is the fourth order symmetric unit tensor and $\mathcal{A}(\mathbf{y},\boldsymbol{\varepsilon})$ is a strain-concentration like tensor stemming from a standard homogenization of the constitutive behavior at the domain $\mathcal{B}_\mu$.

Equations in Box 2.2 display, in an incremental (rate) format, the macroscopic constitutive model inherited by the macro-scale from the homogenization procedure at the micro-scale. Although that model will not be used for computational purposes,[4] it supplies relevant insights on the properties of the resulting homogenized constitutive model and crucial requirements on the numerical model for the propagating material failure at the macro-scale, i.e.:

- Eq. (30)(a) corresponds to an inelastic constitutive equation in terms of the average elastic constitutive tensor, $\overline{\boldsymbol{\mathcal{C}}}$, at the microscale, see Eq. (30)(b), and some inelastic strains, $\dot{\boldsymbol{\chi}}(\mathbf{x},t)$ and $\boldsymbol{\varepsilon}^{(i)}(\mathbf{x},t)$, which would play the role of internal variables in phenomenological models. The evolution equations for those internal variables are given by Eqs. (30)(c) and (30)(d). Unlike in phenomenological models, their evolution is determined, at every macroscopic sampling point, $\mathbf{x}$, by homogenized values of entities at the corresponding micro-scale failure cell, $\mathcal{B}_\mu(\mathbf{x})$: i.e. $\overline{\boldsymbol{\varepsilon}}_\mu(\tilde{\mathbf{u}}_\mu(\mathbf{y},t),\boldsymbol{\beta}_\mu(\mathbf{y},t))$, the elastic properties at the micro-scale, $\boldsymbol{C}_\mu(\mathbf{y})$ and the specific failure mechanism activated at the micro-scale, $\mathcal{S}_\mu(\mathbf{x})$. This extends to non-smooth problems, some theoretical results already derived for smooth problems, see for instance [40,41].
- A characteristic length $\ell_\mu(\mathbf{x},t)$ emerges naturally in Eq. (30)(d). If $n_{\dim}\in\{2,3\}$ is the dimension of the problem then, ℓ_μ is the ratio of the measure of the failure cell, $\left|\mathcal{B}_\mu\right| = \Omega_\mu = \mathcal{O}(h_\mu^{n_{\dim}})$, and the measure (length/surface) of the activated micro-failure mechanism, $L_\mu = \left|\mathcal{S}_\mu\right| = \mathcal{O}(h_\mu^{(n_{\dim}-1)})$. In consequence ℓ_μ is of the order of the failure-cell size i.e.: $\ell_\mu = \mathcal{O}(h_\mu)$. If we assume scale separation, $h_\mu \ll L$ (see Fig. 1), then ℓ_μ is a very small number in the upper scale $\ell_\mu(\mathbf{x},t) \ll L$ (see Figs. 1 and 5).
- Eqs. (30) can be specified for the domains $\mathcal{B}_{loc}(t)$ and $\mathcal{B}\setminus\mathcal{B}_{loc}(t)$ at the macro scale as:

(1) *Smooth domain* $\mathcal{B}\setminus\mathcal{B}_{loc}(t)$:

The microscale behavior is instantaneously elastic (see Eqs. (4) and (6)) and, therefore, $\dot{\boldsymbol{\beta}}_\mu(\mathbf{y},t) = \mathbf{0}\ \forall\mathbf{y}$. Then, in Eq. (30)(d), $\dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t)=\mathbf{0}$ and Eqs. (30)(a) and (30)(c) read

$$\begin{aligned}\dot{\boldsymbol{\sigma}}(\mathbf{x},t) &= \overline{\boldsymbol{\mathcal{C}}} : [\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) + \dot{\boldsymbol{\chi}}(\mathbf{x},t)] \qquad \text{(a)}\\ \dot{\boldsymbol{\chi}}(\mathbf{x},t) &= \boldsymbol{\mathcal{M}} : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) \qquad \text{(b)}\end{aligned} \tag{31}$$

and, therefore,

$$\dot{\boldsymbol{\sigma}}(\mathbf{x},t) = \underbrace{\overline{\boldsymbol{\mathcal{C}}} : (\mathbf{I}+\boldsymbol{\mathcal{M}})}_{\overline{\boldsymbol{\mathcal{C}}}^{elas}_{hom}} : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) = \overline{\boldsymbol{\mathcal{C}}}^{elas}_{hom} : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) \tag{32}$$

where $\overline{\boldsymbol{\mathcal{C}}}^{elas}_{hom}$ stands for the *homogenized elastic constitutive operator* at the failure cell $\mathcal{B}_\mu$.

Notice, from Eq. (31)(a), that a bounded strain, $\dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) = \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t)$ is recovered at $\mathcal{B}\setminus\mathcal{B}_{loc}(t)$ as,

$$\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) = \underbrace{\overline{\boldsymbol{\mathcal{C}}}^{-1} : \dot{\boldsymbol{\sigma}}(\mathbf{x},t) - \dot{\boldsymbol{\chi}}(\mathbf{x},t)}_{\equiv\dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t)\ (bounded)} = \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) \quad \forall\mathbf{x}\in\mathcal{B}\setminus\mathcal{B}_{loc}(t). \tag{33}$$

(2) *Non-smooth domain* $\mathcal{B}_{loc}(t)$:

Material micro-structure displays some failure mechanism by triggering inelasticity (see Eqs. (5)) at the set of activated cohesive bands, $\mathcal{B}_{\mu,act}(\mathbf{x})$.

Therefore, $\dot{\boldsymbol{\beta}}_\mu(\mathbf{y},t)\neq\mathbf{0}\ \forall\mathbf{y}\in\mathcal{B}_{\mu,act}(\mathbf{x})$ and, from Eqs. (30)(a) and (30)(d),

$$\begin{aligned}\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) &= \underbrace{\overline{\boldsymbol{\mathcal{C}}}^{-1} : \dot{\boldsymbol{\sigma}}(\mathbf{x},t) - \dot{\boldsymbol{\chi}}(\mathbf{x},t)}_{\equiv\dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t)\ (bounded)} + \frac{1}{\ell_\mu}\overline{(\dot{\boldsymbol{\beta}}_\mu\otimes^s\mathbf{n}_\mu)}_{\mathcal{S}_\mu}\\ &= \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) + \frac{1}{\ell_\mu}\overline{(\dot{\boldsymbol{\beta}}_\mu\otimes^s\mathbf{n}_\mu)}_{\mathcal{S}_\mu} \quad \forall\mathbf{x}\in\mathcal{B}_{loc}.\end{aligned} \tag{34}$$

Eqs. (33) and (34) can be summarized for the complete domain $\mathcal{B}$ as:

$$\begin{aligned}\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) &= \underbrace{\overline{\boldsymbol{\mathcal{C}}}^{-1} : \dot{\boldsymbol{\sigma}}(\mathbf{x},t) - \dot{\boldsymbol{\chi}}(\mathbf{x},t)}_{\equiv\dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t)\ (smooth)} + \frac{\kappa_{\mathcal{B}_{loc}}}{\ell_\mu}\overline{(\dot{\boldsymbol{\beta}}_\mu\otimes^s\mathbf{n}_\mu)}_{\mathcal{S}_\mu}\\ &= \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) + \frac{\kappa_{\mathcal{B}_{loc}}}{\ell_\mu}\overline{(\dot{\boldsymbol{\beta}}_\mu\otimes^s\mathbf{n}_\mu)}_{\mathcal{S}_\mu} = \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) + \delta^{\ell_\mu}_{\mathcal{S}}\overline{(\dot{\boldsymbol{\beta}}_\mu\otimes^s\mathbf{n}_\mu)}_{\mathcal{S}_\mu}.\end{aligned} \tag{35}$$

[4] Instead, the homogenized value of the stresses in Eq. (18) is point-wise used to evaluate the current macro-stresses in terms of the corresponding macro-strains.

Remark 2.3. Comparing Eq. (35) with Eq. (2), one realizes that the former defines a ℓ_μ-*regularized discontinuity kinematics of bandwidth* $h = \ell_\mu$. This is a fact of crucial importance for a proper and meaningful modeling of the material failure propagation at the upper scale. In fact, Eq. (35) suggests that the numerical approach for capturing this propagation has to be consistent with the aforementioned kinematics of regularized strong discontinuity and, more specifically, regularized with the characteristic length ℓ_μ.

Remark 2.4. The role of the characteristic length, ℓ_μ, is not only computational, but it has also other relevant physical and mechanical implications. Consideration of such a characteristic length for multi-scale based approaches has been claimed from the materials mechanics community [20] and introduced in a heuristic way in some previous approaches [30].

Remarkably, in the proposed approach, this length, ℓ_μ, is *naturally derived* from the homogenization setting; it depends on both specific data of the problem and local results of the analysis. For instance: (a) the considered macro-scale material point, $\mathbf{x}$, (b) the microscale failure-cell size, h_μ, (c) the morphology at the micro-scale (d) the activated failure mechanism at the failure-cell, $\mathcal{S}_\mu$, which, in turn, depends on (d) the history and evolution of the mechanical variables at the micro-scale.

Remark 2.5. It is emphasized that all the results in this section are *a priori results*, i.e. not specifically imposed neither in the algorithmic nor in the numerical implementation of the resulting model. They stem because of using the *standard continuum computational homogenization procedure* in Section 2.3, on a failure cell in the micro-scale, exhibiting strong discontinuities (see Fig. 2). Therefore, these results are going to be considered *assessment conditions*, not specifically imposed in the simulations, whose fulfillment is going to be checked in the numerical simulation results in Section 4 as a proof of the consistency of those results.

2.5. Time evolution of material failure at the macroscale. Transition from weak to strong discontinuity

Eq. (35) corresponds to a specific format of the weak discontinuity kinematics in Eq. (2) with

$$\dot{\gamma}(\mathbf{x},t) := \overline{(\dot{\boldsymbol{\beta}}_\mu \otimes^s \mathbf{n}_\mu)}_{\mathcal{S}_\mu} = \frac{1}{L_\mu}\int_{\mathcal{S}_\mu} (\dot{\boldsymbol{\beta}}_\mu \otimes \mathbf{n}_\mu)^S d\mathcal{S}. \tag{36}$$

Since $\overline{(\dot{\boldsymbol{\beta}}_\mu \otimes^s \mathbf{n}_\mu)}_{\mathcal{S}_\mu}$ is not necessarily a rank-one tensor,[5] it does not match necessarily the strong discontinuity kinematics in Eq. (3) corresponding to a propagating displacement discontinuity, but it evolves towards this situation along time. The time of the analysis, at which Eq. (36) takes the format

$$\begin{aligned} &\dot{\gamma}(\mathbf{x},t_B) = \overline{(\dot{\boldsymbol{\beta}}_\mu(\mathbf{y},t_B) \otimes^s \mathbf{n}_\mu(\mathbf{y}))}_{\mathcal{S}_\mu} = (\dot{\boldsymbol{\beta}}(\mathbf{x},t_B) \otimes \mathbf{n}(\mathbf{x}))^S = \dot{\xi}(\mathbf{m}\otimes\mathbf{n})^S \neq \mathbf{0} \\ &\dot{\boldsymbol{\beta}} = \dot{\xi}\mathbf{m}; \quad \|\mathbf{m}\| = \|\mathbf{n}\| = 1 \end{aligned} \tag{37}$$

will determine the *bifurcation time* $t_B(\mathbf{x})$. In other words, t_B is the time at which material failure, displaying localized strains in a weak discontinuity format (see Eq. (2)), starts propagating, as a strong discontinuity, at the considered macro-scale point, $\mathbf{x}$. Determination of this time, and the so-called polarization directions $\mathbf{n}(\mathbf{x})$ and $\mathbf{m}(\mathbf{x})$,[6] is done in the context of the well-known discontinuous bifurcation analysis [42,43] by solving the problem:

Problem.

$$\begin{aligned} &\textit{GIVEN } \mathbf{Q}_t(\mathbf{x},\mathbf{n}) = \mathbf{n}\cdot\boldsymbol{\mathcal{C}}^{tang}_{hom}(\mathbf{x},t)\cdot\mathbf{n}; \quad \mathbf{x}\in\mathcal{S};\ t\in[0,T] \\ &\textit{FIND } \text{the first } t\equiv t_B(\mathbf{x})\in[0,T] \quad \text{and all} \quad \mathbf{n}\in\mathbb{I} := \{\boldsymbol{\nu}\in\mathbb{R}^{n_{\dim}} \mid \|\boldsymbol{\nu}\| = 1\} \\ &\textit{FULFILLING}: \quad \mathbf{Q}_{t_B}(\mathbf{x},\mathbf{n})\cdot\mathbf{m} = \mathbf{0} \quad \text{for some } \mathbf{m}\in\mathbb{I} \end{aligned} \tag{38}$$

where $\boldsymbol{\mathcal{C}}^{tang}_{hom}$ is the macroscale (homogenized) tangent constitutive (see Eq. (30)(a)), $[0,T]$ is the time interval of interest and $\mathbf{Q}_t(\mathbf{x},\mathbf{n})$ is the localization tensor.

[5] Since the sum (integral) of rank-one tensors is not a rank-one tensor.

[6] For symmetric $\boldsymbol{\mathcal{C}}^{tang}_{hom}$ the problem is indifferent to the order of the couple $(\mathbf{n},\mathbf{m})$.

A general numerical procedure for solving this problem can be found in [44]. In general Eq. (38) provides the bifurcation time, $t_B(\mathbf{x})$, and two solutions for the normal to the discontinuity path $\mathcal{S}$ (see Fig. 1) i.e.: $\mathbf{n}_1(\mathbf{x}) \equiv \mathbf{n}$, $\mathbf{n}_2(\mathbf{x}) \equiv \mathbf{m}$.

Now, examining Eq. (37), one realizes that the equality is fulfilled under the following circumstances:

(1) The fluctuations jump at the microscale, $\dot{\boldsymbol{\beta}}_\mu$, is *spatially constant* at the failure-cell:

$$\begin{aligned}
&\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},\mathbf{y},t) = \dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t) \quad \forall \mathbf{y} \in \mathcal{B}_\mu \\
&\Rightarrow \begin{cases} \overline{(\dot{\boldsymbol{\beta}}_\mu \otimes^s \mathbf{n}_\mu)}_{\mathcal{S}_\mu} = (\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t) \otimes \overline{\mathbf{n}_\mu(\mathbf{x},\mathbf{y})}_{\mathcal{S}_\mu})^S = \zeta(\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t) \otimes \mathbf{a}(\mathbf{x}))^S \\ \mathbf{a}(\mathbf{x}) = \dfrac{1}{\zeta}\overline{\mathbf{n}_\mu(\mathbf{x},\mathbf{y})}_{\mathcal{S}_\mu}; \quad \|\mathbf{a}\| = 1; \quad \zeta(\mathbf{x}) = \left\|\overline{\mathbf{n}_\mu(\mathbf{x},\mathbf{y})}_{\mathcal{S}_\mu}\right\| = \zeta(\mathbf{n}_\mu) \end{cases}
\end{aligned} \tag{39}$$

where factor $\zeta(\mathbf{x})$ is a measure of the *tortuosity* of the activated micro-scale failure path $\mathcal{S}_\mu$ (for instance, for $\mathcal{S}_\mu$ being a straight line then $\zeta = 1$). A similar expression for this tortuosity can be found in [45].

Replacing Eq. (39) into Eq. (37) yields:

$$\zeta(\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t) \otimes \mathbf{a}(\mathbf{x}))^S = (\dot{\boldsymbol{\beta}}(\mathbf{x},t) \otimes \mathbf{n}(\mathbf{x}))^S \Rightarrow \begin{cases} \dot{\boldsymbol{\beta}}(\mathbf{x},t) = \zeta(\mathbf{n}_\mu)\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t) = \mathcal{F}(\mathbf{n}_\mu, \dot{\boldsymbol{\beta}}_\mu) \\ \mathbf{n}(\mathbf{x}) = \mathbf{a}(\mathbf{x}) = \dfrac{1}{\zeta(\mathbf{n}_\mu)}\overline{(\mathbf{n}_\mu)}_{\mathcal{S}_\mu} = \mathcal{G}(\mathbf{n}_\mu). \end{cases} \tag{40}$$

(2) The activated failure path at the micro-scale, $\mathcal{S}_\mu$, is a straight line (or a plain surface), with spatially constant normal $\mathbf{n}_\mu$:

$$\begin{aligned}
&\mathbf{n}_\mu(\mathbf{x},\mathbf{y}) = \mathbf{n}_\mu(\mathbf{x}) \\
&\Rightarrow \mathbf{n}(\mathbf{x}) = \mathbf{n}_\mu(\mathbf{x}); \qquad \dot{\boldsymbol{\beta}}(\mathbf{x},t) = \overline{\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},\mathbf{y},t)}.
\end{aligned} \tag{41}$$

This is an academic case, since it corresponds to a non-realistic microstructural morphology of the failure-cell ($\mathcal{B}_{\mu,coh}$ defines a straight/plane band). Therefore, it will be discarded because of its lack of physical significance and practical interest.

Remark 2.6. The previous analysis leads to the following statements:

1. The (rate of the) micro-scale displacement-jump function (after bifurcation) is constant across the failure-cell ($\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},\mathbf{y},t) = \dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t), \forall \mathbf{y} \in \mathcal{B}_\mu(\mathbf{x})$) in Eq. (39).
2. The (rate of the) macro-scale displacement-jump function, $\dot{\boldsymbol{\beta}}(\mathbf{x},t)$, (after bifurcation) equals the micro-fluctuation jump, $\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t)$, corrected by the (scalar) tortuosity factor ($\dot{\boldsymbol{\beta}}(\mathbf{x},t) = \zeta(\mathbf{x})\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t)$ in Eq. (40)(a)).
3. After bifurcation, the macroscale normal to the crack-path $\mathbf{n}(\mathbf{x})$ equals the (normalized) average of the normal to the corresponding microscopic failure mechanism $\mathbf{a}(\mathbf{x}) = \frac{1}{\zeta}\overline{\mathbf{n}_\mu(\mathbf{x},\mathbf{y})}_{\mathcal{S}_\mu}$ in Eq. (40)(b).
4. It should also be noticed another subtle aspect of statements 2 and 3 above. In fact, inspection of Eqs. (37) and (39) reveals that the symmetric operator $(\mathbf{n} \otimes \mathbf{m})^S$ makes the results insensitive with respect to an interchange of vectors $\mathbf{n} \rightleftarrows \mathbf{m}$. Therefore, besides the solutions in Eq. (40), solutions

$$\begin{aligned}
&\dot{\boldsymbol{\beta}}^*(\mathbf{x}) = \left\|\dot{\boldsymbol{\beta}}_\mu(\mathbf{x})\right\| \mathbf{a}(\mathbf{x}) = \left\|\dot{\boldsymbol{\beta}}_\mu(\mathbf{x})\right\| \frac{\overline{(\mathbf{n}_\mu)}_{\mathcal{S}_\mu}}{\zeta(\mathbf{n}_\mu)} = \mathcal{F}^*(\dot{\boldsymbol{\beta}}_\mu, \mathbf{n}_\mu) \\
&\mathbf{n}^*(\mathbf{x}) = \frac{1}{\left\|\dot{\boldsymbol{\beta}}_\mu(\mathbf{x})\right\|}\dot{\boldsymbol{\beta}}_\mu(\mathbf{x}) = \mathcal{G}^*(\dot{\boldsymbol{\beta}}_\mu)
\end{aligned} \tag{42}$$

fulfilling

$$(\dot{\boldsymbol{\beta}}^* \otimes \mathbf{n}^*)^S = (\underbrace{\dot{\boldsymbol{\beta}}_\mu(\mathbf{x})}_{=\dot{\boldsymbol{\beta}}} \otimes \underbrace{\frac{\overline{(\mathbf{n}_\mu)}_{\mathcal{S}_\mu}}{\zeta(\mathbf{n}_\mu)}}_{=\mathbf{n}})^S = (\dot{\boldsymbol{\beta}} \otimes \mathbf{n})^S \tag{43}$$

are also admissible. In other words, what is relevant for the resulting macroscopic homogenized constitutive model is not the couple $(\mathbf{n}, \dot{\boldsymbol{\beta}})$ but its symmetric tensor product.

This *indifference*, with respect to vector swap in $(\mathbf{n}\otimes\mathbf{m})^S$, motivates that *a given microscopic failure mechanism* $\{\mathbf{n}_\mu, \dot{\boldsymbol{\beta}}_\mu\}$ *gives rise to two possible combinations* of the macro-scale displacement jump and normal:

$$\begin{aligned}&\{\dot{\boldsymbol{\beta}}, \mathbf{n}\} \equiv \{\mathcal{F}(\mathbf{n}_\mu, \dot{\boldsymbol{\beta}}_\mu), \mathcal{G}(\mathbf{n}_\mu)\}\\ &\{\dot{\boldsymbol{\beta}}^*, \mathbf{n}^*\} \equiv \{\mathcal{F}^*(\dot{\boldsymbol{\beta}}_\mu, \mathbf{n}_\mu), \mathcal{G}^*(\dot{\boldsymbol{\beta}}_\mu)\}\end{aligned} \tag{44}$$

as in Eqs. (40) and (42). This enriches the modeling capabilities of the approach for propagation directions in the macroscale. Indeed, even with a relatively poor description of the failure mechanisms at the microscopic scale (restricted set of $\mathbf{n}_\mu$'s but unrestricted set of $\dot{\boldsymbol{\beta}}_\mu$'s) a much wider set of propagation directions ($\{\mathbf{n} \mid \mathbf{n} = \mathcal{G}(\mathbf{n}_\mu)\} \bigcup \{\mathbf{n}^* \mid \mathbf{n}^* = \mathcal{G}^*(\dot{\boldsymbol{\beta}}_\mu)\}$) can be captured at the macro-scale.

These are *a priori* results, not explicitly imposed in the approach, which shall be corroborated by the numerical simulations in Section 4.

2.6. Energy dissipation. Fracture energy issues

The material failure at the micro-cell $\mathcal{B}_\mu$, translates into propagating failure at the macro-scale, which, in turn, is captured through the regularized strain localization band, $\mathcal{B}_{loc}(t)$ (see Fig. 5). The proposed approach enforces this *bandwidth to be precisely the characteristic length,* $\ell_\mu(\mathbf{x})$ in Eq. (30)(d).

Let us consider the fracture energy, $G^f_\mu(\mathbf{y})$, corresponding to points $\mathbf{y} \in \mathcal{B}_{\mu,act}$ at the microscale, understood as a material property that can be different for every material compound at the microstructure, and the macro-scale fracture energy, $G^f(\mathbf{x})$, obtained *as an output* from the homogenization procedure. According to their definition (expended power per unity of mid-surface of the localization band) those fracture energies can be, respectively computed, in terms of the fracture energy densities[7] (per unit of volume) $g^f(\mathbf{x})$ and $g^f_\mu(\mathbf{x})$:

$$\begin{aligned}&g^f(\mathbf{x}) = \frac{G^f(\mathbf{x})}{\ell_\mu(\mathbf{x})} = \int_0^\infty \boldsymbol{\sigma}(\mathbf{x},t) : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t)dt \quad \forall \mathbf{x} \in \mathcal{B}_{loc} \qquad \text{(a)}\\ &g^f_\mu(\mathbf{y}) = \frac{G^f_\mu(\mathbf{y})}{k} = \int_0^\infty \boldsymbol{\sigma}_\mu(\mathbf{y},t) : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{y},t)dt \quad \forall \mathbf{y} \in \mathcal{B}_{\mu,coh}. \quad \text{(b)}\end{aligned} \tag{45}$$

In this setting, the following lemma holds:

Lemma. *The macroscopic fracture energy,* $G^f(\mathbf{x})$, is the average of the microscopic fracture energy $G^f_\mu(\mathbf{y})$, *along the activated failure mechanism path at the micro-scale,* $\mathcal{S}_\mu(\mathbf{x})$ *(see Fig. 5):*

$$G^f(\mathbf{x}) = \frac{1}{L_\mu(\mathbf{x})} \int_{\mathcal{S}_\mu(\mathbf{x})} G^f_\mu(\mathbf{y})\, d\mathcal{S}_\mu = \overline{G^f_\mu(\mathbf{y})}_{\mathcal{S}_\mu}. \tag{46}$$

Proof. Considering the Hill–Mandel principle in Eq. (17), Eq. (45)(a) reads,

$$\begin{aligned}\frac{G^f(\mathbf{x})}{\ell_\mu(\mathbf{x})} &= \int_0^\infty \boldsymbol{\sigma}(\mathbf{x},t) : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t)dt = \int_0^\infty \left[\frac{1}{\Omega_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y},t) : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{x},\mathbf{y},t) d\mathcal{B}_\mu\right] dt\\ &= \int_0^\infty \left[\frac{1}{\Omega_\mu}\int_{\mathcal{B}_{\mu,act}} \boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y},t) : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{x},\mathbf{y},t) \underbrace{d\mathcal{B}_{\mu,act}}_{=k\cdot d\mathcal{S}_\mu}\right] dt\\ &= \int_{\mathcal{S}_\mu} \left[\frac{k}{\Omega_\mu}\int_{t_B}^\infty \boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y},t) : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{x},\mathbf{y},t)dt\right] d\mathcal{S}_\mu\end{aligned} \tag{47}$$

[7] Strictly speaking, the integration of the macroscopic specific fracture energy, $g^f(\mathbf{x})$ in Eq. (45)(a) should be evaluated along the interval $[t_B, \infty]$, where $t_B(\mathbf{x})$ stands for the macroscopic bifurcation time in Eq. (38). Nevertheless, here the pre-bifurcation dissipated energy at the macro-scale is considered negligible, so that it is integrated along the interval $[0, \infty]$.

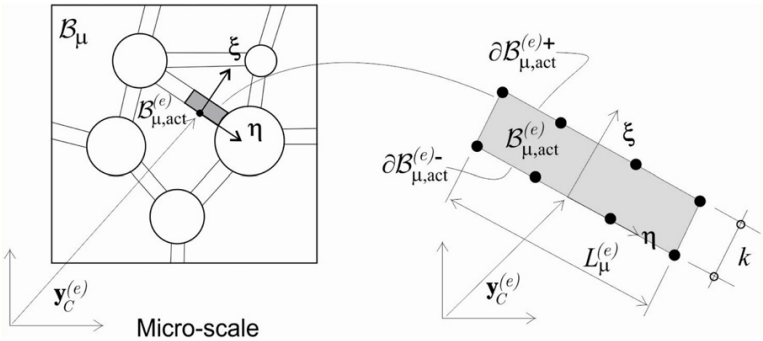


Fig. 6. Multiscale model: finite element discretization at the micro-scale.

where the fact that, in the domain outside the active bands $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,act}$, the material is elastic (with no dissipation), has been taken into account. Now, inserting Eq. (45)(b) into Eq. (47), and considering the definition of the characteristic length, ℓ_μ in Eq. (30)(d), yields

$$\frac{G^f(\mathbf{x})}{\ell_\mu(\mathbf{x})} = \frac{1}{\Omega_\mu}\int_{\mathcal{S}_\mu} \underbrace{\left[k\int_0^\infty \boldsymbol{\sigma}_\mu(\mathbf{x},\mathbf{y},t) : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{x},\mathbf{y},t)dt\right]}_{G^f_\mu(\mathbf{y})} d\mathcal{S}_\mu = \frac{1}{\Omega_\mu}\int_{\mathcal{S}_\mu(\mathbf{x})} G^f_\mu(\mathbf{y})\, d\mathcal{S}_\mu$$
$$= \underbrace{\frac{L_\mu}{\Omega_\mu}}_{\frac{1}{\ell_\mu(\mathbf{x})}} \frac{1}{L_\mu}\int_{\mathcal{S}_\mu(\mathbf{x})} G^f_\mu(\mathbf{y})\, d\mathcal{S}_\mu = \frac{1}{\ell_\mu(\mathbf{x})}\overline{G^f_\mu(\mathbf{y})_{\mathcal{S}_\mu}} \tag{48}$$

which proves the Lemma.

Remark 2.7. For homogeneous fracture energy at the active cohesive bands ($G^f_\mu(\mathbf{y}) = G^f_\mu \; \forall \mathbf{y} \in \mathcal{B}_{\mu,act}$) the lemma, consistently, translates into *conservation of the fracture energy along the scales* i.e.:

$$G^f(\mathbf{x}) = G^f_\mu. \tag{49}$$

The previous results rely crucially on the ability to induce a localization band of bandwidth $\ell_\mu(\mathbf{x})$ at the macro-scale. This issue is tackled in Appendix B.

3. Propagation of material failure

3.1. Propagation of material failure at the micro-scale. Failure cell finite element model

A standard finite element model is adopted for the numerical simulation of the RVE/failure-cell response. One, quadrilateral,[8] finite element is used across the thickness of the cohesive bands $\mathcal{B}_{\mu,coh}$, as shown in Fig. 6. The interpolation order along the coordinate η may be arbitrary. However, it must be linear along the coordinate ξ in order to capture the function $\mathcal{H}_{\mathcal{B}_{\mu,k}}(\eta)$ defined in Eq. (8)(b).

Propagation of material failure along the cohesive bands takes place, in a natural fashion, as inelasticity is triggered in one element, $\mathcal{B}^{(e)}_{\mu,coh}$, and it propagates across other elements of $\mathcal{B}_{\mu,coh}$. The set of all elements, e, of the cohesive bands experiencing inelastic loading at the center point $\mathbf{y}^{(e)}_C$ of the element, see Fig. 6, determines, at every time, t, the failure mechanism at the micro-scale, i.e.:

$$\mathcal{B}_{\mu,act}(t) := \left\{\bigcup \mathcal{B}^{(e)}_{\mu,coh};\; \dot{r}_\mu(\mathbf{y}^{(e)}_C, t) > 0\right\} \tag{50}$$

[8] However, there is no intrinsic restriction for using other options, like triangular elements.

where the instantaneous inelasticity status (de-cohesion) is evaluated in terms of a positive evolution of the internal variable $\dot{r}_\mu > 0$ (see Box 2.1).
From Eq. (50), one can determine the measure (length/surface) of the activated failure mechanism in Eq. (30)(d) and the corresponding characteristic length as

$$\begin{cases} L_\mu(\mathbf{x},t) = \displaystyle\sum_{\forall \mathcal{B}^{(e)}_{\mu,act} \subset \mathcal{B}_{\mu,act}(\mathbf{x},t)} L^{(e)}_\mu & \text{(a)} \\ \ell_\mu(\mathbf{x},t) \equiv \dfrac{\Omega_\mu}{L_\mu(\mathbf{x},t)} < h^{(e)}(\mathbf{x}) & \text{(b)} \end{cases} \tag{51}$$

where $L^{(e)}_\mu$ stands for the individual length of the active element, $\mathcal{B}^{(e)}_{\mu,act}$, in the set of active cohesive bands $\mathcal{B}_{\mu,act}(t)$.

Remark 3.1. In practice, dependence on pseudo-time, t, of entities $\mathcal{B}_{\mu,act}(t)$, $L_\mu(\mathbf{x},t)$ and $\ell_\mu(\mathbf{x},t)$ in Eq. (51) only holds for the time interval $[t_Y(\mathbf{x}), t_B(\mathbf{x})]$, where $t_Y(\mathbf{x})$ stands for the time where material inelasticity starts at the failure cell, and $t_B(\mathbf{x})$ is the macroscopic bifurcation time in Eq. (38). At time $t_B(\mathbf{x})$, the failure mechanism at the cell is assumed fully consolidated and $\ell_\mu(\mathbf{x},t)$ in Eq. (51)(b) is made constant in time and redefined as:

$$\ell_\mu(\mathbf{x},t) := \ell_\mu(\mathbf{x}) = \ell_\mu(\mathbf{x}, t_B(\mathbf{x})) \quad \forall t \geq t_B(\mathbf{x}). \tag{52}$$

Remark 3.2. In Eq. (51)(b), limitation $\ell_\mu(\mathbf{x},t) < h^{(e)}(\mathbf{x})$ resolves, the singularity for the case that no damage occurs at the micro-scale (i.e. $L_\mu = 0 \rightarrow \ell_\mu = \infty$ in the elastic case, when ℓ_μ does not play any role). On the other hand, it ensures a positive value for the integration weights in the numerical integration procedure involved in the strain injection method described in Appendix B (see Table 3). In any case, this limitation, which can be removed if other crack-propagation techniques are used at the macro-scale, is not a strong one since, roughly speaking, it establishes that the failure-cell size, h_μ, has to be smaller than the macro-scale finite element size ($h_\mu(\mathbf{x}) \simeq \ell_\mu(\mathbf{x},t) < h^{(e)}(\mathbf{x})$). This restriction is, normally, largely exceeded by the scale separation requirement.

3.1.1. *Unstable failure modes. Imposition of constant fluctuation jumps at cohesive bands*

Inspection of equations in Box 2.2 reveals that the model might exhibit some instability. Indeed, the structure of the inelastic strain in Eq. (30)(c), allows situations displaying:

$$\begin{aligned} &\dot{\boldsymbol{\beta}}_\mu(\mathbf{y}) \neq \mathbf{0} \quad \text{for some } \mathbf{y} \in \mathcal{S}_\mu && \text{(a)} \\ &\dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t) = \frac{1}{\ell_\mu}(\overline{\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},\mathbf{y}) \otimes^s \mathbf{n}_\mu(\mathbf{x},\mathbf{y})})_{\mathcal{S}_\mu} = \mathbf{0} && \text{(b)} \end{aligned} \tag{53}$$

which can give rise to some instabilities in the microscopic failure mechanisms. In Appendix A, this issue is studied and a remedy to preclude this instabilities, based on imposing an internal restriction in the fluctuations on same of the cohesive bands, is presented.

3.2. *Propagation of material-failure at the macroscale. Crack-path-field and strain injection techniques in multi-scale problems*

A difficult issue in computational modeling of material failure is the appropriate capture of its onset and propagation. Material failure propagation algorithms aim at answering the following two questions: (1) *when* does failure trigger at a given material point?, and, (2) *how* does it propagates?

As commented in Section 3.1, at the lower scale (failure-cell) where the morphology and position of candidate propagation mechanisms (set $\mathcal{B}_{\mu,coh}$) are predefined, this is a minor issue; both questions are automatically answered in terms of when and what cohesive-band elements enter into the softening regime according to Eq. (50). However, at the upper scale, there is no predefinition of the failure path and, in principle, any material point may fail and propagate in any direction.

To answer the above questions, in this work we have extended, to the multiscale case, a procedure for modeling onset and propagation of discontinuities recently developed by the authors for one-scale problems [31]. The proposed methodology is based on the use of two specific techniques:

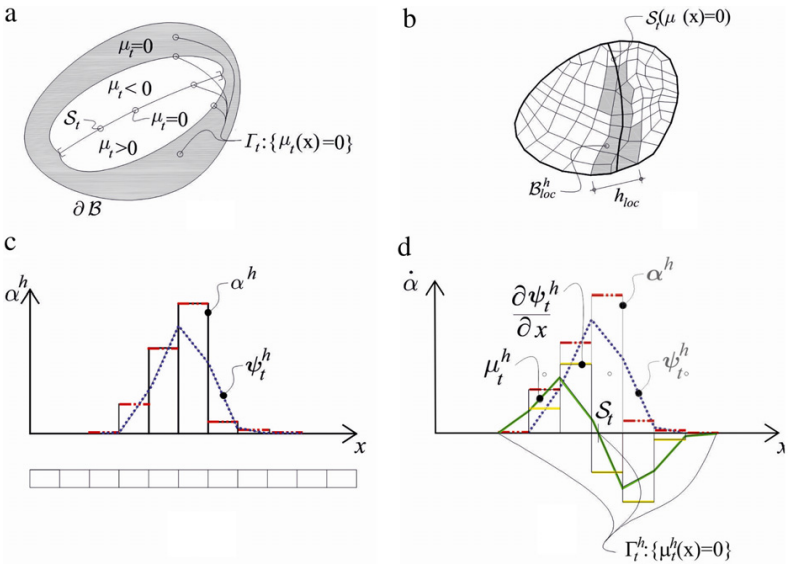


Fig. 7. Crack-path-field technique. (a) and (b) evolving crack-path, $\mathcal{S}_t$, (c) and (d) 1D sketch of the algorithm to obtain the crack path field from the smooth derivative of the localized variable α (see also Box 3.1).

(1) *Crack-path-field techniques*, consisting of the identification of the trace, during time and across the domain $\mathcal{B}$, of an evolving strong discontinuity represented by a strain localized field, in the subset $\mathcal{B}_{sd} \subset \mathcal{B}$. The so-called crack path field, $\mu(\mathbf{x}, t)$, is obtained from a selected localized scalar variable, $\alpha(\mathbf{x}, t)$, and its zero level set identifies, at every time t of the analysis, the domain, Γ_t, that contains the crack-path, $\mathcal{S}_t$. It is computed as:

$$\mathcal{S}_t \subset \Gamma_t := \left\{ \mathbf{x} \in \mathcal{B}_{loc};\ \mu(\mathbf{x}, t) = \frac{\partial \alpha}{\partial e} \equiv \nabla \alpha \cdot \hat{\mathbf{e}}(\mathbf{x}, t) = 0 \right\} \tag{54}$$

which identifies the crack path as the locus of the e-directional maximum of the field $\alpha(\mathbf{x}, t)$ (see Fig. 7). The optimum value for the directional unit vector, $\hat{\mathbf{e}}(\mathbf{x}, t)$, in Eq. (54) should be the one orthogonal to the crack path, Γ_t, which, in turn is the unknown of the problem. From the authors' experience, a convenient, approximate, value for $\hat{\mathbf{e}}(\mathbf{x}, t)$, is extracted from the gradient of a displacement field scalar measure,[9] $a(\mathbf{x}, t)$, i.e.:

$$\begin{aligned} &\mathbf{e}(\mathbf{x}, t) = \boldsymbol{\nabla} a(\mathbf{x}, t) \\ &a(\mathbf{x}, t) = \sum_{i=1}^{n_{\text{dim}}} \boldsymbol{\nabla} u_i(\mathbf{x}, t); \quad u_i(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t) \cdot \hat{\mathbf{e}}_i \\ &\hat{\mathbf{e}}(\mathbf{x}, t) = \frac{\mathbf{e}(\mathbf{x}, t)}{\|\mathbf{e}(\mathbf{x}, t)\|}. \end{aligned} \tag{55}$$

For the present multiscale approach, the localized variable $\alpha(\mathbf{x}, t)$ in Eq. (54) has to be specifically imported from the microstructure. Here, it is proposed the following definition in terms of the damage variable at the microscale, r_μ in Box 2.1, i.e.:

$$\begin{cases} \dot{\alpha}(\mathbf{x}, t) = \ell_\mu(\mathbf{x}, t) \dfrac{1}{\Omega_\mu} \displaystyle\int_{\Omega_\mu} \dot{r}_\mu(\mathbf{x}, \mathbf{y}, t)\, d\Omega = \frac{1}{L_\mu(\mathbf{x}, t)} \int_{\Omega_\mu} \dot{r}_\mu(\mathbf{x}, \mathbf{y}, t)\, d\Omega & \text{(a)} \\ \alpha(\mathbf{x}, 0) = 0 & \text{(b)} \end{cases} \tag{56}$$

where ℓ_μ is the characteristic length, defined in Eq. (51)(b).

[9] This approximately identifies $\mathbf{e}(\mathbf{x}, t)$ as the direction orthogonal the localization propagation.

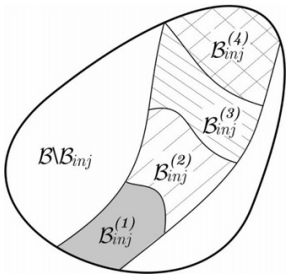


Fig. 8. Strain injection: the injection domains, $\mathcal{B}_{inj}^{(i)}(t)$, are injected an assumed (rate of) strain, $\dot{\varepsilon}_{inj}^{(i)}(\mathbf{x},t)$, in the specific time interval $[]_{inj}^{(i)} \subset [0,T]$.

Box 3.1.
Crack-path-field problem

Problem.

GIVEN:

$\alpha(\mathbf{x},t) : \mathcal{B} \times [0,T] \to \mathbb{R}^+$

$\hat{\mathbf{e}}(\mathbf{x},t) : \mathcal{B} \times [0,T] \to \mathbb{I} := \{\mathbf{v} \in \mathbb{R}^{n_{\text{dim}}} |\ \|\mathbf{v}\| = 1\}$

$$\mathcal{V} := \left\{ \eta(\mathbf{x}) = \sum_{1}^{n_{node}} N_i(\mathbf{x})\eta_i;\ \eta \in H^1(\mathcal{B}) \right\}$$

$$\mathcal{V}_0 := \left\{ \eta(\mathbf{x}) = \sum_{1}^{n_{node}} N_i(\mathbf{x})\,\eta_i; \eta \in H^1(\mathcal{B});\ \eta|_{\partial\mathcal{B}} = 0 \right\}$$

FIND:

(a) $\psi_t(\mathbf{x}) \in \mathcal{V}$ *fulfilling* $\displaystyle\int_{\mathcal{B}^h} \overline{\psi}(\psi_t - \alpha(\mathbf{x},t))d\mathcal{B} = 0 \quad \forall\overline{\psi} \in \mathcal{V}_0$

(b) $\mu_t(\mathbf{x}) \in \mathcal{V}$ *fulfilling* $\displaystyle\int_{\mathcal{B}^h} \overline{\mu} \left(\mu_t - \underbrace{\frac{\partial\psi_t}{\partial e}}_{\nabla\psi_t\cdot\hat{\mathbf{e}}} \right) d\mathcal{B} = 0 \quad \forall\overline{\mu} \in \mathcal{V}_0$ (57)

(c) $\Gamma_t := \{\mathbf{x} \in \mathcal{B};\ \mu_t(\mathbf{x}) = 0\} \to$ *zero level set of* $\mu_t(\mathbf{x})$

The corresponding algorithm, based on smoothing the variables $\alpha(\mathbf{x},t)$ and $\frac{\partial\alpha}{\partial e}$ in Eqs. (54) and (56), is presented in Box 3.1.

(2) *Strain injection techniques*: consisting of the insertion in specific, time-dependent, sub-domains $\mathcal{B}_{inj}^{(i)}(t) \subset \mathcal{B}$, $\mathcal{B}_{inj}(t) = \bigcup_{i=1}^{n_{inj}} \mathcal{B}_{inj}^{(i)}(t)$ (see Fig. 8), of goal-oriented strain fields. By using strain injection techniques a number of different strain rate patterns, $\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) \equiv \dot{\boldsymbol{\varepsilon}}_t(\mathbf{x})$, can be inserted at the specific domains, $\mathcal{B}_{inj}^{(i)}(t)$, using classical two-field, $\dot{\boldsymbol{\varepsilon}}_t(\mathbf{x})/\boldsymbol{\nabla}^s\dot{\boldsymbol{u}}_t(\mathbf{x})$ mixed finite element strategies [37]. The format of the considered problem in terms of rates (see Box 3.2), allows keeping the resulting time-varying problem smooth in time, in spite of the, possible, sudden changes of the injected rate-of-strain formats. More details on the technique can be found in a work of the authors [31].

Box 3.2.
General strain-injection variational problem (in rate form)

Problem.

$$
\begin{aligned}
&\textit{GIVEN}:\\
&\mathcal{V}_t := \{\boldsymbol{\eta}(\mathbf{x},t) \in H^1(\mathcal{B}) \times [0,T];\ \boldsymbol{\eta}(\mathbf{x},t)|_{\partial_u \mathcal{B}} = \dot{\mathbf{u}}^*(\mathbf{x},t)\}\\
&\mathcal{V}_0 := \{\boldsymbol{\eta}(\mathbf{x}) \in H^1(\mathcal{B});\ \boldsymbol{\eta}(\mathbf{x})|_{\partial_u \mathcal{B}} = \mathbf{0}\}\\
&\mathcal{E} := \{\boldsymbol{\mu}(\mathbf{x},t) \in L^2(\mathcal{B})^{n_{\text{dim}} \times n_{\text{dim}}} \times [0,T]\}\\
&\textit{FIND}:\\
&\dot{\mathbf{u}}_t(\mathbf{x}) \equiv \dot{\mathbf{u}}(\mathbf{x},t) : \mathcal{B} \times [0,T] \to \mathbb{R}^{n_{\text{dim}}}; \qquad \dot{\mathbf{u}}_t \in \mathcal{V}_t\\
&\dot{\boldsymbol{\varepsilon}}_t(\mathbf{x}) \equiv \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) : \mathcal{B} \times [0,T] \to S^{n_{\text{dim}} \times n_{\text{dim}}}; \quad \dot{\boldsymbol{\varepsilon}}_t \in \mathcal{E}
\end{aligned}
\tag{58}
$$

where

$$
\dot{\boldsymbol{\varepsilon}}_t(\mathbf{x}) = \begin{cases} \dot{\boldsymbol{\varepsilon}}_{inj}^{(i)}(\mathbf{x},t) & \forall \mathbf{x} \in \mathcal{B}_{inj}^{(i)}(t), i \in \{1,\ldots,n_{inj}\}\ t \in []_{inj}^{(i)} \\ \boldsymbol{\nabla}^s \dot{\mathbf{u}}_t(\mathbf{x}) & \forall \mathbf{x} \in \mathcal{B} \setminus \mathcal{B}_{inj}(t) \end{cases}
\tag{59}
$$

FULFILLING:

$$
\begin{aligned}
&\int_{\mathcal{B}} \boldsymbol{\nabla}^s \boldsymbol{\eta} : \dot{\boldsymbol{\Sigma}}(\dot{\boldsymbol{\varepsilon}}_t)\, d\mathcal{B} - W^{ext}(\boldsymbol{\eta}, \dot{\mathbf{b}}, \dot{\mathbf{t}}^*) = 0 \qquad \text{(a)}\\
&\int_{\mathcal{B}} \boldsymbol{\mu} : (\dot{\boldsymbol{\varepsilon}}_t - \boldsymbol{\nabla}^s \dot{\mathbf{u}}_t)\, d\mathcal{B} = 0 \qquad \text{(b)}\\
&\forall \boldsymbol{\eta} \in \mathcal{V}_0 \quad \forall \boldsymbol{\mu} \in \mathcal{E}
\end{aligned}
\tag{60}
$$

where, in Eq. (59) $[]_{inj}^{(i)}$ stands for the time interval during which the ith injection is done.

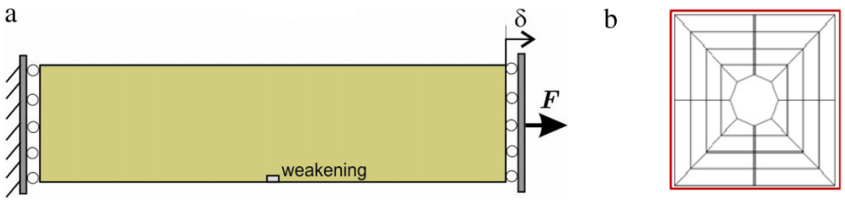


Fig. 9. (DNS) vs. (FE^2) analysis. Tensile stretched strip. (a) Structural scale description. (b) Unit-cell with a central void, enriched with a single vertical cohesive band at the center and discretized by quadrilateral elements.

For the purposes of this work, two different injections are carried out: (a) *weak discontinuity* injection and (b) *strong discontinuity* injection. Their essential features for the present multiscale approach are described in detail in Appendix B.

Remark 3.3. In the context of the strain-injection technique, the localization domain, denoted as $\mathcal{B}_{loc}$ in the proposed multi-scale approach (see Section 2.1 and Fig. 1), becomes the strain-injection domain, $\mathcal{B}_{inj} \equiv \mathcal{B}_{loc}$ (see Fig. 8).

4. Representative simulations

4.1. Consistency of the proposed approach: direct numerical simulation (DNS) versus multi-scale (FE^2) analyses

For the purposes of comparing, the ability of the proposed multi-scale approach to reproduce direct numerical simulation results, the following problems are considered.

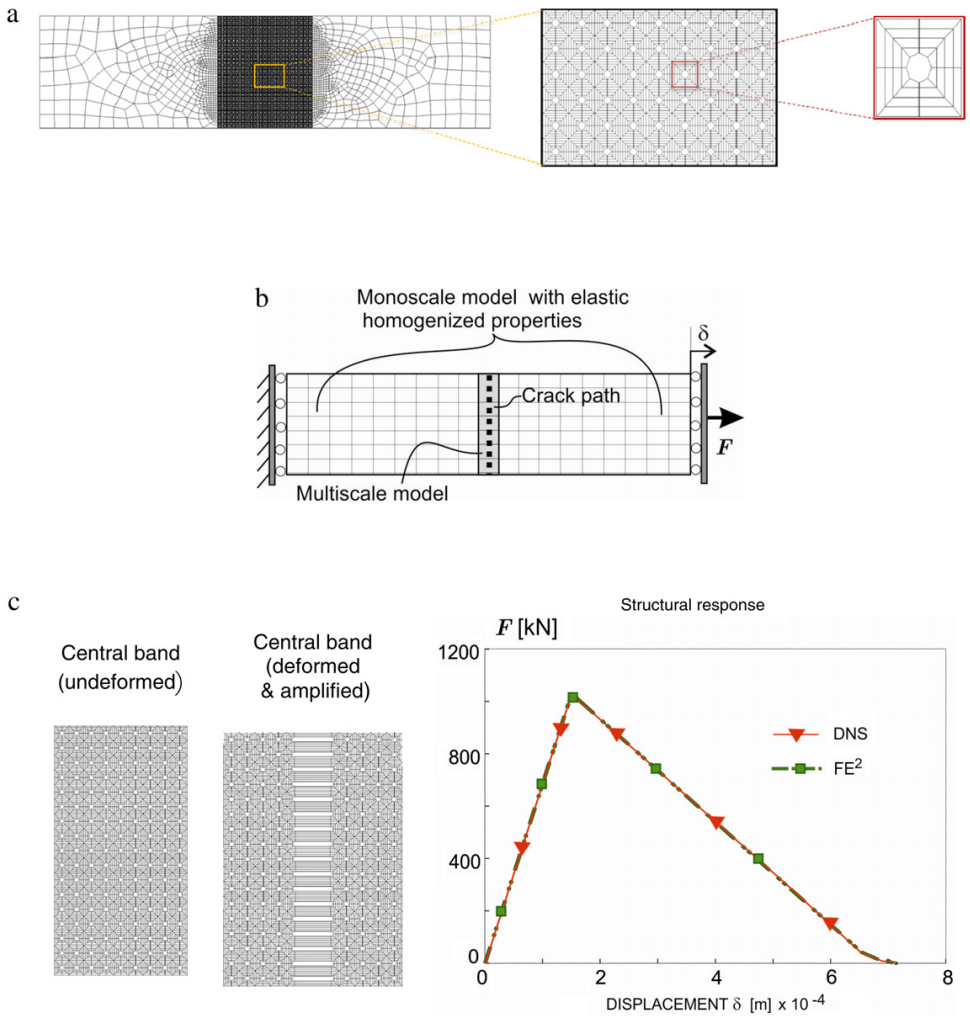


Fig. 10. DNS vs. FE^2 analysis in the tensile stretched strip. (a) Finite element discretization for the DNS analysis. (b) Finite element discretization for the FE^2 analysis and (c) qualitative and quantitative results in both analyzes.

4.1.1. Strip under uniaxial stretching

Let us consider the 2D strip under uniaxial tensile stretching (in plane strain) depicted in Fig. 9(a), loaded by imposing a horizontal homogeneous displacement, δ, at the right edge, while the left edge stays fixed. The length and height of the specimen are, respectively 2 m and 0.5 m.

The material is assumed to exhibit a simple mesostructure consisting of a periodic distribution of voids embedded into a homogeneous elastic matrix. The failure mechanism is represented by vertical cohesive bands connecting the voids. The unit cell of this mesostructure is shown in Fig. 9(b).

The considered material properties are given in Table 1. In order to break the homogeneity of the specimen, that vertical cohesive band placed at the central-lower part of the specimen is weakened to trigger the failure at that point.

In Fig. 10 the considered finite element mesh for the DNS (made of bi-linear quadrilaterals) is shown. A central band of the specimen is finely discretized to capture the considered microstructure. The remaining of the specimen is discretized in a much coarser mesh, whose elements are considered elastic along the whole analysis, so they are

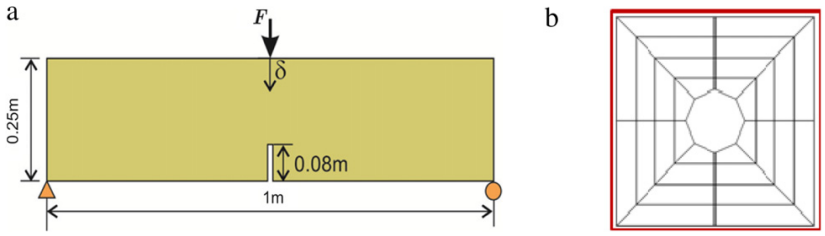


Fig. 11. (DNS) vs. (FE^2) analysis. Three-point bending test. (a) Structural (macroscale) description of the problem. (b) RVE/failure-cell with a central opening enriched with a vertical cohesive band and discretized by quadrilateral elements.

Table 1
Material properties for the matrix and cohesive bands at the micro-scale.

	Young's modulus (GPa)	Poisson ratio	Yield stress (MPa)	Fracture energy (N/m)
Matrix	30	0.15	Elastic	–
Cohesive bands	30	0.15	3	1000

equipped with the homogenized elastic properties stemming from the homogenization of the RVE in Fig. 10(a). For the FE^2 analysis the coarse mesh of quadrilaterals in Fig. 10(b), is used. Again, only the central row of elements are amenable to become inelastic, through a multiscale modeling, whereas the remaining elements are considered elastic with homogenized elastic properties.

The theoretical solution of this problem consists of a central vertical crack, which opens homogeneously (constant crack opening), starting from the weakened point, and resulting in the classical two-branch structural response (see Fig. 10(c)). That solution was exactly captured by both the DNS and the FE^2 solutions, assessing the proposed approach in this case.

4.1.2. *Three-point beam bending test*

In order to check the ability of the proposed approach to reproduce multi-scale fracture problems with a *propagating crack* at the macro-scale, the classical three point beam bending test, with a central notch, is now considered for a DNS vs. FE^2 comparison.

The considered structural problem and the mesoscale morphology (the same than in the previous case) are displayed in Fig. 11. The considered material properties are the ones displayed in Table 1.

Again, the finite element mesh for the DNS analysis, the discretization for the FE^2 analysis and the comparisons for the structural responses are presented in Fig. 12. It can be noticed the ability of the proposed multi-scale approach to pass DNS vs. FE^2 benchmarks, also in cases of non-homogeneous (variable crack opening) propagating fracture.

4.2. *Objectivity analysis of the proposed multiscale model*

Objectivity is here understood as the insensitivity of the numerical solutions with respect to: (a) changes in size and bias of the macroscale finite element mesh and, (b) changes of the size and shape of the failure-cell adopted for representing the material micro/mesostructure.

4.2.1. *Objectivity with respect to the macroscale mesh size and bias*

For the strip under uniaxial stretching in Section 4.1.1, now solved in the multi-scale context, the three meshes depicted in Fig. 13 are considered. For computational saving reasons, the meshes are only modified at the central zone of the strip, i.e. the region where the crack propagates. Meshes (a), (b) and (c) have quadrilateral elements with sizes: $h^e \simeq 0.065$ m, 0.035 m and 0.024 m respectively. In addition, multiscale modeling is restricted to the central zone, so that the remaining of the specimen is modeled with the one-scale (elastic) homogenized material.

The mesostructure is assumed composed of a matrix, in plane strain state, with voids of circular shape with arbitrary diameter and space distributions. In Fig. 13(a), the adopted failure cell is displayed.

In order to trigger the crack initiation in the central zone of the strip, one element in every mesh is perturbed (in red in Fig. 13), by decreasing in 25% the ultimate stress at the corresponding microscale cohesive bands.

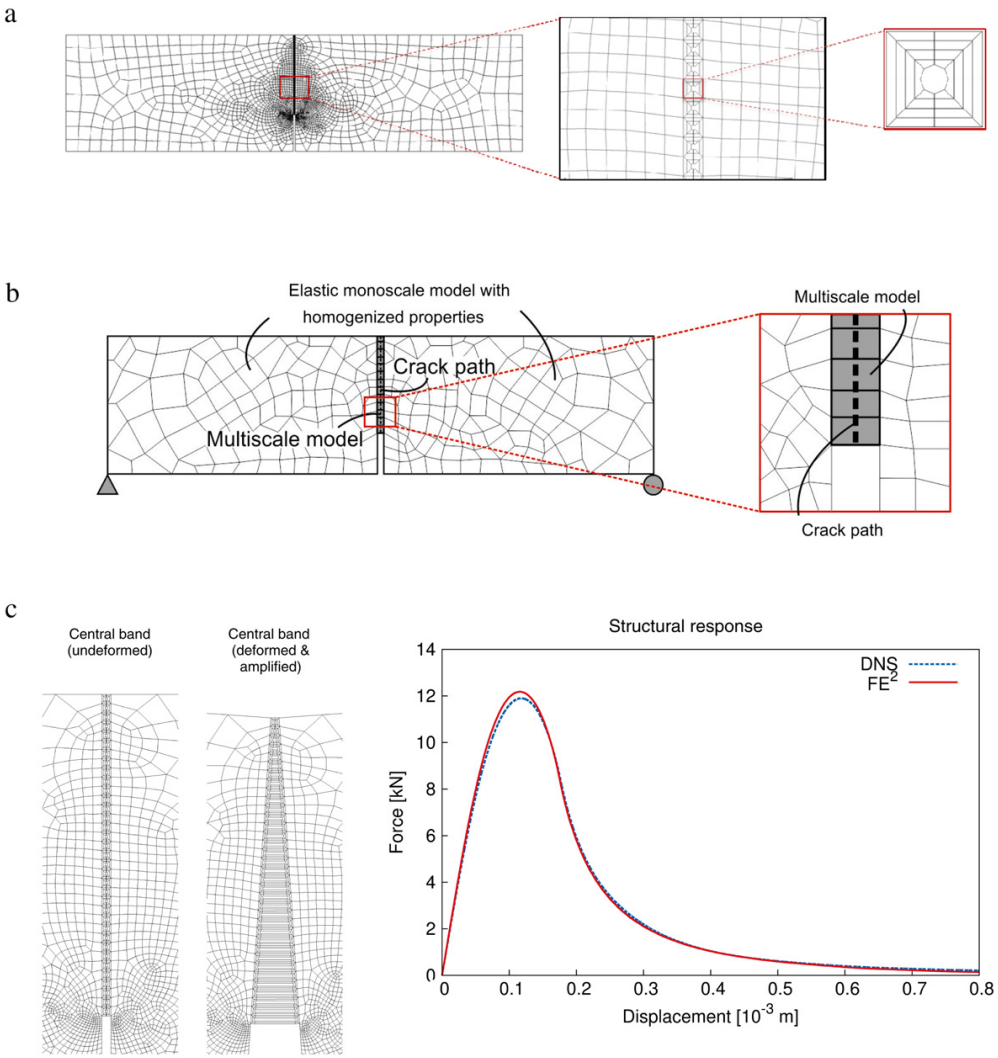


Fig. 12. DNS vs. FE^2 analysis for the three-point bending test. (a) Finite element discretization for the DNS analysis. (b) Macro-scale finite element discretization for the FE^2 analysis (central part) and (c) qualitative and quantitative results for both analyzes.

In Fig. 14 the obtained numerical results are displayed. Fig. 14(a) plots the total load $\boldsymbol{F}$ vs. the horizontal displacement δ for the three meshes. Fig. 14(b) displays the resulting crack path at the end of the analysis for every mesh.

The insensitivity of the structural responses in Fig. 14(a), with respect to meshes of different size and bias, proves the mesh objectivity of the model.

4.2.2. Objectivity with respect to shape and size of the failure-cell

A basic issue in multiscale material modeling is that the response, displayed by cells representing the microscopic morphology of the material, should be independent of changes in shape and size of these cells, provided they are large enough to characterize the material statistically.

In order to check this specific issue using the proposed approach, the test sketched in Fig. 15 is done. A set of different cells is adopted for characterizing the material response. They are constructed from the repetition, in the horizontal and vertical directions, of a basic cell (see Fig. 15(b)).

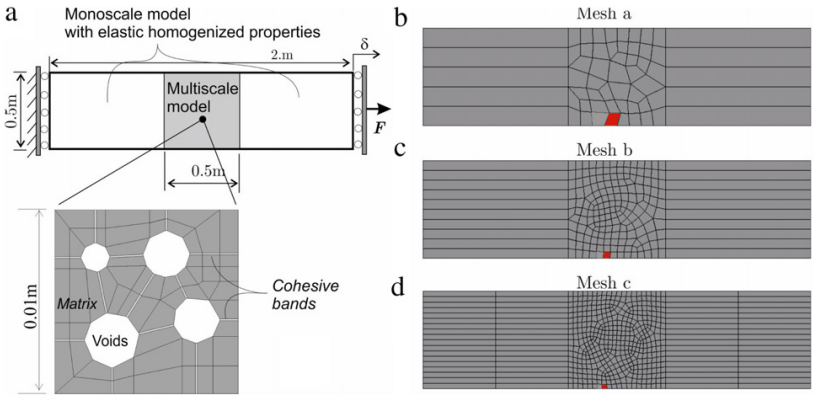


Fig. 13. Objectivity with respect to changes of the finite element mesh at the macro-scale. (a) Multiscale model and failure-cell; (b) coarse discretization (58FE, $h^e \simeq 0.065$ m); (c) medium discretization (192FE, $h^e \simeq 0.035$ m); (d) fine discretization (504 FE, $h^e \simeq 0.024$ m). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

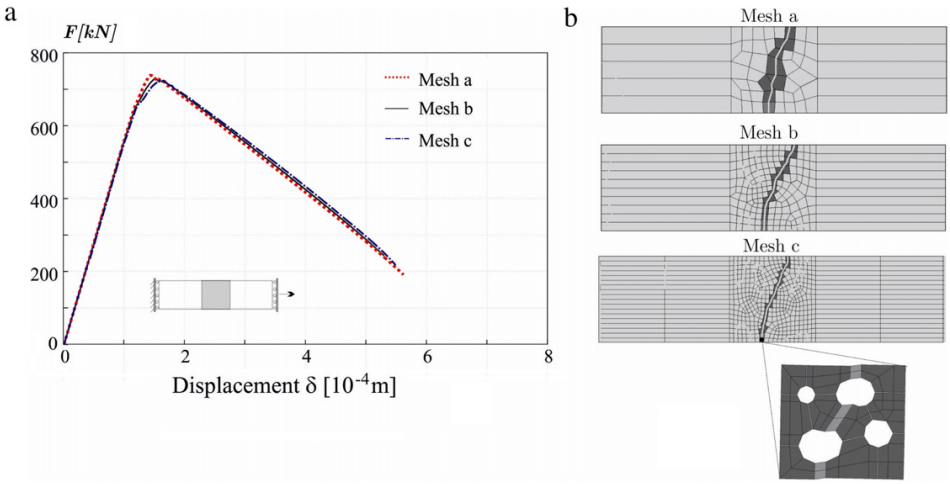


Fig. 14. Objectivity with respect to changes of the finite element mesh at the macro-scale: (a) reaction force vs. end-displacement; (b) crack path (for meshes (a), (b) and (c)) and deformed mesoscopic cell displaying the macroscopic failure mechanisms (activated cohesive bands).

The macroscale structural problem and the material properties are the same than in Section 4.2.1: a strip undergoing tensile loading with multiscale modeling restricted to the central zone (see Fig. 15(a)).

The mesostructure consists of a matrix with a periodic array of cylindrical voids, arranged in a squared pattern of size $h = 0.001$ m, see Fig. 15(a). In order to represent this mesostructure, the unit mesoscopic cell, depicted in Fig. 15(b) and denoted (1×1), is used first. Then, two alternative cells, obtained by repeating cell (1×1) in the horizontal and vertical directions, denoted as (2×1) and (2×2) respectively in Fig. 15(a), are additionally used for the multiscale analysis. Under the chosen loading conditions and microstructure failure cell, the microscale failure mode consists, in the three cases, of the activation of a single vertical column of cohesive bands (see Fig. 15(b)), which translates in a macroscale vertical crack at the central part of the strip (see Fig. 15(a)).

Fig. 15(c) displays the structural response obtained with the three considered mesoscopic cells, (1×1), (2×1) and (2×2), in terms of the load vs. horizontal displacement response. The almost indistinguishable results for the three cases assess the objectivity of the response. In the achievement of these results, it is again emphasized the key role played by the regularization of the kinematics in Eq. (35), in terms of the microscopic characteristic length ℓ_μ (in

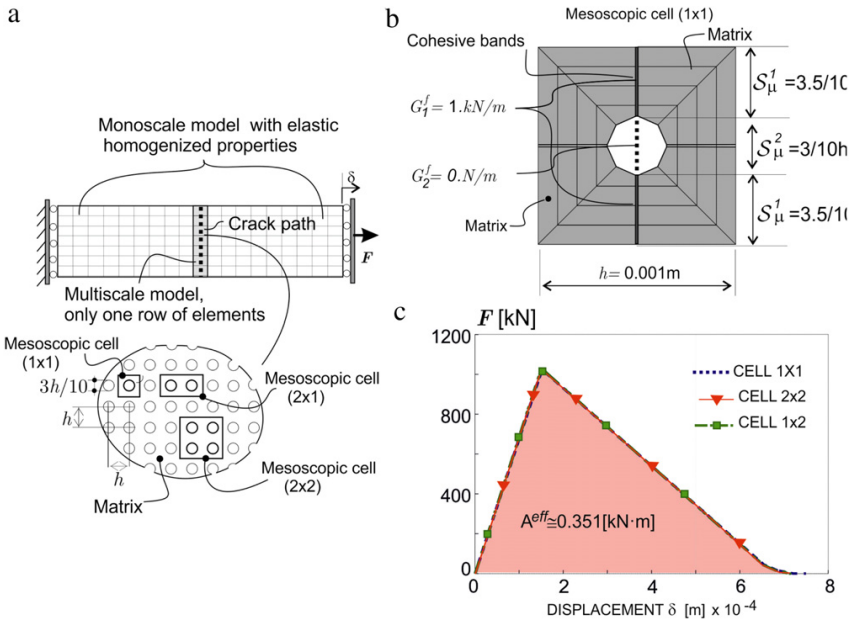


Fig. 15. Objectivity with respect to changes in the failure-cell. (a) Test description; (b) microcell (1 × 1); (c) Structural responses (load vs. horizontal displacement curves) for mesoscopic failure cells: 1 × 1, 2 × 1 and 2 × 2.

Eq. (51)(b)). Indeed, this length takes different values in each case, i.e.: $\ell_\mu = h$ for microcell (1 × 1), and $\ell_\mu = 2h$ for microcells (2 × 1) and (2 × 2).

4.2.3. Fracture energy

Next, the value of the macroscale fracture energy, $G^f(\mathbf{x})$, obtained from the analyzes, is compared with the theoretical (analytical) one in Eq. (46). In order to induce a straight macroscale crack-path with an exactly measurable length, the strip is solved with a vertical structured mesh. As shown in Fig. 15(a), only one row of elements at the central zone of the strip is endowed with a multiscale model. Micro-cell (1 × 1) is taken and a vertical macroscopic crack develops. The material parameters are the same than in the previous case.

Taking into account that the inner void in the failure cell can be considered as a cohesive band with a null fracture energy and length $meas(\mathcal{S}^{void}_{\mu,coh}) = 3/10h$, the *average value of the cohesive fracture energy along the developed microscopic failure mechanism* is evaluated (see Eq. (46) and Fig. 15(b)) as:

$$G^f_{analytical} \equiv \overline{G^f_\mu(\mathbf{y})}_{\mathcal{S}_\mu} = \frac{7/10h}{h} G^f_\mu = 700\ \mathrm{N/m} \tag{61}$$

this supplying the analytical value of the overall (macroscopic) fracture energy, $G^f_{analytical}$. In addition, the numerically obtained value of the overall fracture energy, $G^f_{numerical}$, can be extracted from the area, A^{eff}, under the structural force–displacement curve in Fig. 15(c) i.e.:

$$G^f_{numerical} = \frac{A^{eff}}{length(\mathcal{S}) \times thickness} = \frac{351\ \mathrm{N\ m}}{0.5\ \mathrm{m} \times 1\ \mathrm{m}} = 702\ \mathrm{N/m} \tag{62}$$

which matches the theoretical value in Eq. (61).

4.3. Multiscale model assessment: Nooru-Mohamed test

The experimental test on concrete specimens reported by Nooru-Mohamed [46], and sketched in Fig. 16, is now used to assess the ability of the proposed approach to model *propagating material failure at two scales*. The adopted

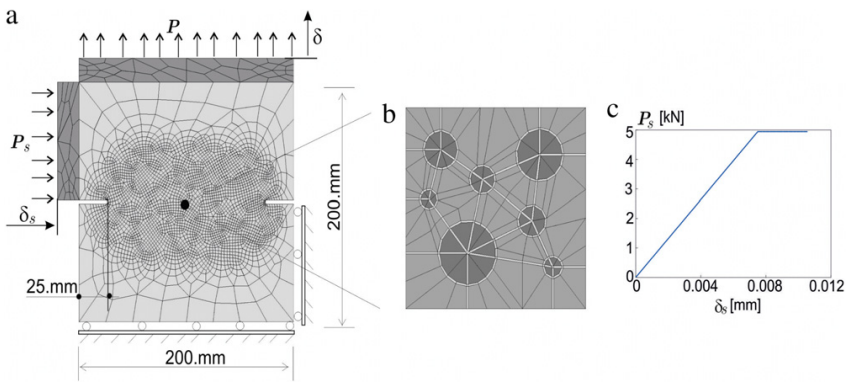


Fig. 16. Nooru-Mohamed test (a) mesh of the specimen with 3464 quadrilateral finite elements; (b) mesoscopic cell and (c) applied shear force, $\boldsymbol{P}_S$, history as a function of the horizontal displacement of the left edge of the specimen δ_s.

Table 2
Nooru-Mohamed test. Material properties.

	Young's modulus (GPa)	Poisson ratio	Ultimate stress (MPa)	Fracture energy (N/m)
Matrix	21	0.15	Elastic	–
Cohesive bands	21	0.15	3	20
Aggregates	100	0.15	Elastic	–

failure cell is presented in Fig. 16(b). For computational cost reasons, the cell crudely represents the morphology of concrete as a matrix (mortar) with some inclusions (aggregates) of different sizes. Therefore, no attempt has been done to represent the statistically distribution of heterogeneities and failure characteristics, at the mesoscopic level, observed in standard concretes. The interest is restricted to assess the ability of the numerical model to simulate a rather complex structural behavior, involving propagating material failure in two scales, and to compare the results with the wide set of experiences available for this benchmark using *single scale phenomenological approaches*.

4.3.1. Geometrical features and loading paths

The specimen, of size 200 mm $\times$ 200 mm and 50 mm of thickness, with two notches (25 mm depth and 0.125 depth/with ratio) is depicted in light gray in Fig. 16(a). The dark gray part of the structure is made of steel. It is considered infinitely rigid and used to apply the different loads in a distributed form.

Loading consists of a shear force applied in a nearly elastic regime, until reaching the value $\boldsymbol{P}_S = 5\ (\mathrm{kN})$ and remaining constant afterwards. Subsequently, an increasing vertical uniformly distributed load $\boldsymbol{P}$ is superposed (see Fig. 15(a)–(c)). During the first loading stage (shear force application), the left and top steel parts are constrained to remain vertical and horizontal, respectively.

4.3.2. Material properties description

The mesoscopic failure cell, depicted in Fig. 16(b), crudely represents a concrete-like material at the mesoscale (matrix/aggregates). Light gray regions represent cohesive/failure bands, medium gray zone stands for the cement-like matrix and dark gray regions represents aggregates (the cohesive bands into the aggregates are inhibited to failure). Material properties of the cell are summarized in Table 2.

4.3.3. Crack propagation result

In order to analyze the ability of the proposed approach for capturing the crack evolution, attention is first focused on the macroscopic scale. In Fig. 17, the iso-displacement contours at the end of the simulation depict highly localized strain zones representing two evolving cracks, which initiate at the notch roots. A slight crack curvature can be observed at the bottom crack, and roughly, both cracks can be viewed as straight surfaces with an inclination of about 10°. In spite of the commented crude representation of the microstructure, the obtained macroscale crack pattern fairly coincides with the results reported in the experiment [46].

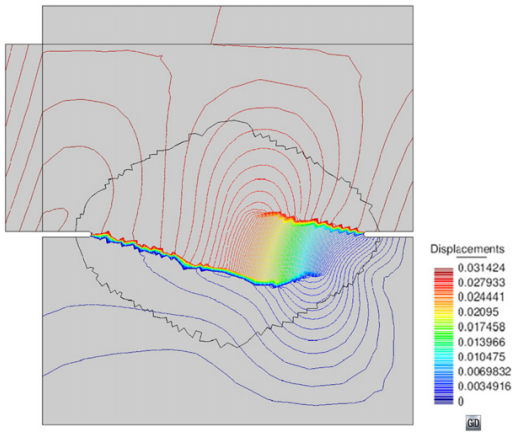


Fig. 17. Nooru-Mohamed test. Iso-lines of the macroscopic vertical displacements. Grouping of iso-lines indicates a jump in the displacement field and, therefore, signals the crack path.

The series of snapshots in Fig. 18 show the evolution of the strain injected embedded weak discontinuity region (finite elements in dark gray), and the finite elements with injected embedded strong discontinuities (in red). The weak discontinuity region spreads across a wide zone, forming a bulb-like zone at the tip of each crack endowed with highly flexible elements, suitable to capture localized strains and their propagation. On this basis, the crack-path-field technique in Box 3.1, supplies the crack-path, as the zero level set of the crack path field μ_t, which, in turn, defines the proper position (in every Γ_t-crossed element) of the injected strong discontinuity mode.

Fig. 19 displays additional results of the numerical solution. The crack path obtained at the end of analysis is shown in Fig. 19(a), which is compared with the experimental one. The structural response in terms of vertical loads vs. vertical displacement of the top specimen is shown in Fig. 19(b).

Classical effects of a propagating crack on the structural response are displayed by the softening character of the structural response curve in spite of the very crude representation of the material through the failure cell.

Next, attention is focused on the behavior of the mesoscopic cell at the singular Gauss point, $\mathbf{x}_{sing}^{(e)}$ (the one capturing the mesoscopic failure, see Fig. 27, and Table 3 in Appendix B) for different elements at the micro-scale. Fig. 20 depicts a zoom of the left notch, showing the crack-path (line in blue) and the normal vector, $\mathbf{n}^{(e)}$, used for embedding the strong discontinuity kinematics in each finite element (see Fig. 29, in Appendix B). A number of elements have been selected to analyze the mesoscopic results. For every selected element, the corresponding deformed failure cell, at the end of analysis, is shown. There, it can be checked that those elements crossed by the crack path display a clear mesoscopic failure mode characterized by a (regularized) displacement jump, in a set of cohesive elements, which is constant for all of them as predicted by the theoretical results (see Remark 2.6). The one down-left the notch, though nearby the crack path, is not crossed by it and, as expected, displays a smooth deformation mode.

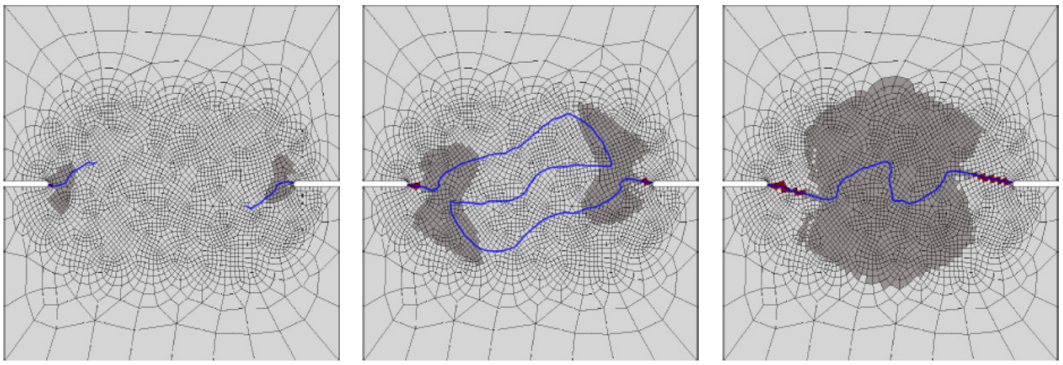
4.3.4. Macro-/meso-scale relationship

The effects of distribution and richness of possible failure mechanisms at the mesoscopic cell, and their effects, on the crack patterns at the macro-scale and on the structural response, are now examined.

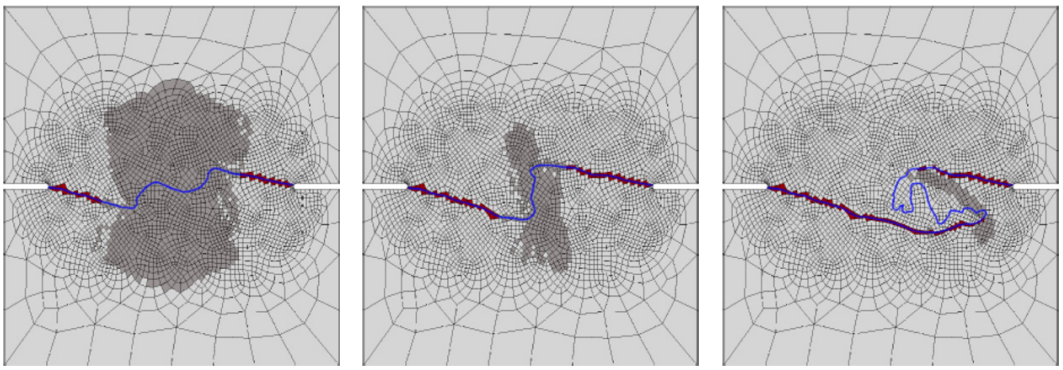
For this purpose, the original cell in Fig. 16(b) is considered, together with two additional failure cells (see Fig. 21). They keep the original material morphology, in terms of matrix and inclusions, but considerably reduce the number of possible failure mechanisms (by locking the inelastic behavior of appropriated cohesive bands).
In case of Fig. 21(b) the only possible failure mechanism is almost horizontal, whereas in Fig. 21(c), it is oblique.

Fig. 22 displays the macroscopic crack pattern obtained with each of these cells.

Fig. 23 displays the structural responses in terms of load vs. vertical displacement, provided by the three cells in Fig. 21. As it can be expected, locking of failure modes due to removal of cohesive bands at the mesoscopic scale leads to the following response: the poorer is the mesoscopic cell description the higher is the structural strength.



(a) $\delta = 0.011$ mm. (b) $\delta = 0.014$ mm. (c) $\delta = 0.018$ mm.



(d) $\delta = 0.019$ mm. (e) $\delta = 0.020$ mm. (f) $\delta = 0.025$ mm.

Fig. 18. Nooru-Mohamed test. Evolution of injected weak discontinuity region (in dark gray), zero-level set of the crack path field (lines in blue) and finite elements injected with strong discontinuity (in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

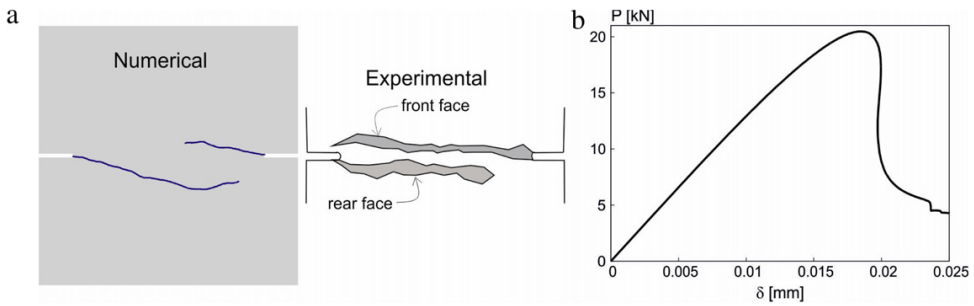


Fig. 19. Nooru-Mohamed test. Characterization of the structural response. (a) Numerical and experimental crack path; (b) Axial force P as a function of the axial displacement δ.

5. Concluding remarks

Along this work, a new approach to two-scale modeling of material failure, based on computational homogenization (FE^2), has been presented. The specific features of the approach rely on two main aspects: (a) the adopted

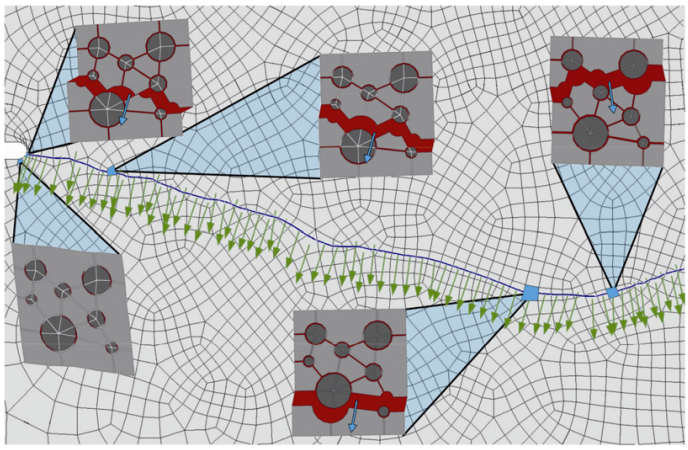


Fig. 20. Nooru-Mohamed test. Deformed (amplified) failure cells, for several elements at the macroscale at the end of the analysis, displaying the activated failure mechanism (in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

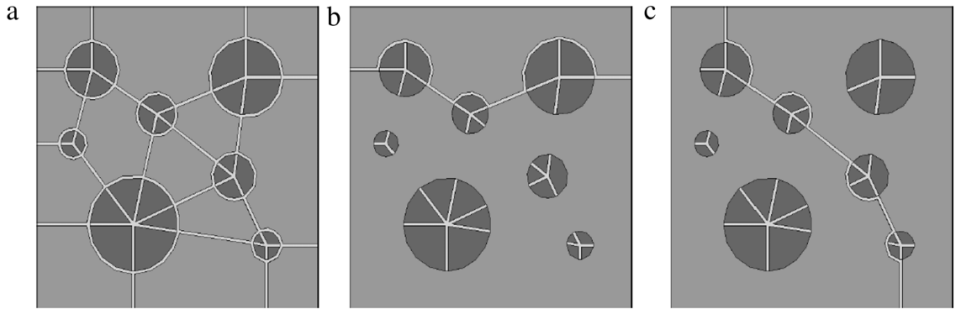


Fig. 21. Nooru-Mohamed test. Mesoscale cells with different cohesive-band patterns. (a) Original failure cell; (b) Failure cell allowing only a quasi-horizontal failure mechanism; (c) Failure cell allowing an oblique failure mechanism.

continuum setting for representation of the fracture, based on the Continuum Strong Discontinuity Approach (CSDA) and developed by the authors in previous works [35], and, (b) a standard (continuum) format of the computational homogenization procedure.

The resulting framework is endowed with the following properties:

(1) Minimally invasive nature with regard to procedures well established in the literature on multiscale modeling of materials.[10] In fact,
- in terms of the *computational homogenization procedure*: the proposed approach displays no substantial difference with respect to the ones used for smooth (continuous) problems. A RVE/failure cell is defined at the lower (micro/meso scale) from which stresses and strains are homogenized in a classical continuum format to obtain a, point-wise defined, stress–strain constitutive model at the macro-scale.
- In terms of *material failure propagation*: existing algorithms for mono-scale crack propagation modeling can be easily extended to the multi-scale case. In this sense, the authors' experience in extending the crack-path-field and strain injection techniques [31] previously developed for the mono-scale case to the present multiscale case, should be extensible to other families of crack-propagation techniques.

[10] In this work, the modern framework of *variational multi-scale formulation* [16,17] has been considered as the best reference.

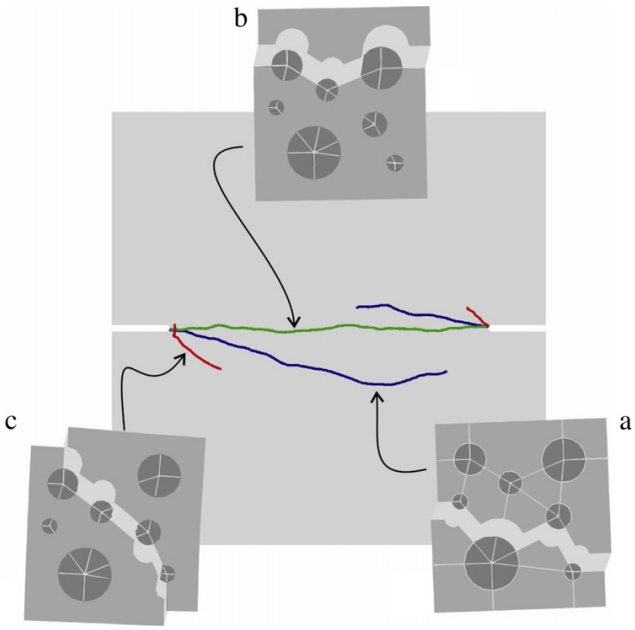


Fig. 22. Nooru-Mohamed test. Crack paths simulated according with the adopted mesoscopic cell.

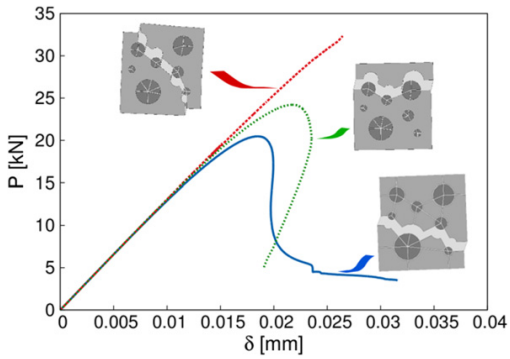


Fig. 23. Nooru-Mohamed test. Structural responses for different meso-cells.

(2) Consistency of the homogenized results. This has been checked by comparison, for a number of representative cases, of the (FE^2) results and the ones obtained by direct numerical simulation (DNS).
(3) Objectivity character (independence) of the results in terms of:
   - finite element mesh size and bias, at the macro-scale
   - failure-cell (size and shape) at the micro/meso-scale.

A key issue in this achievements is the use of a characteristic length, which naturally emerges from theoretical considerations on the homogenized constitutive model, to regularize the weak/strong discontinuity kinematics at the macro-scale in the context of the CSDA.

The performance of the proposed approach has been assessed by a number of examples. They go beyond the, simple, homogenization-based assessment (focusing only on the obtained homogenized macro-scale constitutive equation), but towards a much more ambitious goal: modeling the fracture onset and propagation at the macro-scale

based on the, simpler, in the physics, but more complex in the morphology, micro-scale representations, and the corresponding structural action–response.

However, the authors are aware that realistic multiscale representations of material failure require the use of more sophisticated microscopic morphologies, allowing complex failure mechanisms at the micro-scale [47]. In this sense, realistic multiscale analysis of material failure, including the extension to 3D cases, face a great challenge: the enormous involved computational cost. It is well known that the algorithmic complexity of multiscale analysis leads to the so-called *tyranny of the scales* [48], which makes computations unaffordable, even by resorting to intensive computation procedures (parallel computing), for relatively crude representations of the material morphology and coarse finite element meshes at the macro-scale. A possible remedy for this drawback could stem from the development and use of high performance reduced order modeling (HPROM) techniques for multiscale problems. In a recent work [49] the authors have developed some specific and efficient algorithms for the smooth (continuous) case. Extension to the, much more challenging, non-smooth cases, like multiscale material failure, is a subject of ongoing research.

Finally, some extensions of the material representation at the micro-scale, beyond the one presented here, seem to fit trivially into the proposed approach at the only cost of the modification of some specific ingredients of the model. For instance:

- Inclusion of non-linear hardening behavior, before the onset of material failure can be trivially included in the considered damage model.
- Consideration of other families of constitutive behavior, like plasticity, rate dependence etc., can be simply done by including these effects in the basic constitutive model.
- The use of cohesive bands to model failure at the micro-scale is neither a compulsory ingredient of the approach. Using other propagating crack models at the micro-scale, either based on continuum strain-localization methods (Continuum Strong Discontinuity Approach, non-local models or gradient-regularized models) or discrete methods (based on cohesive interfaces equipped with traction–separation laws) where the crack onset and propagation is not predefined (just as it is done here at the macro-scale) also fits in the approach at the cost of some additional sophistication.

These and other issues, concerning the extension of the proposed approach, are subject of ongoing research and they will be addressed in future works.

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement n. 320815, Advanced Grant Project COMP-DES-MAT (Advanced tools for computational design of engineering materials).

The funding received from the Spanish Ministry of Economy and Competitiveness through the National Research Plan 2014 project: MAT2014-60919-R is also gratefully acknowledged.

Appendix A. Removal of spurious failure modes

When the fundamental localized solution is perturbed, in terms of the micro-fluctuations in Eq. (22)(a), by increments

$$\dot{\tilde{\mathbf{u}}}^{*}_{\mu}(\mathbf{y}) = \mathcal{H}_{\mathcal{B}_{\mu,coh}}(\mathbf{y})\,\dot{\boldsymbol{\beta}}^{*}_{\mu}(\mathbf{y}); \quad \dot{\tilde{\mathbf{u}}}_{\mu}(\mathbf{y}) \in \boldsymbol{\mathcal{V}}_{\mu} \tag{63}$$

(see Eq. (16)), $\dot{\boldsymbol{\beta}}^{*}_{\mu}(\mathbf{y},t)$ fulfilling

$$\begin{aligned} &\dot{\boldsymbol{\beta}}_{\mu}(\mathbf{y}) \neq \mathbf{0} \quad \text{for some } \mathbf{y} \in \mathcal{S}_{\mu} && \text{(a)} \\ &\dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t) = \frac{1}{\ell_{\mu}}\overline{(\dot{\boldsymbol{\beta}}_{\mu}(\mathbf{x},\mathbf{y}) \otimes^{s} \mathbf{n}_{\mu}(\mathbf{x},\mathbf{y}))}_{\mathcal{S}_{\mu}} = \mathbf{0} && \text{(b)} \end{aligned} \tag{64}$$

a bifurcation in the solution space $\boldsymbol{\sigma}$–$\boldsymbol{\varepsilon}$, may appear. Indeed, in this case, the fundamental branch is characterized by the rate equation (see Eq. (30)(a)):

$$\dot{\boldsymbol{\sigma}}^{(1)}(\mathbf{x},t) = \overline{\boldsymbol{\mathcal{C}}} : [\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t) + \dot{\boldsymbol{\chi}}(\mathbf{x},t) - \dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t)] = \boldsymbol{\mathcal{C}}^{tang}_{hom}(\boldsymbol{\varepsilon}) : \dot{\boldsymbol{\varepsilon}}(\mathbf{x},t). \tag{65}$$

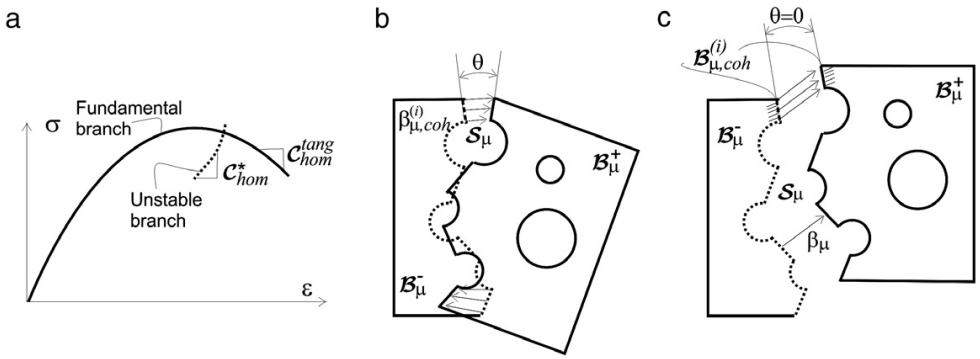


Fig. 24. (a) Instability of the homogenized, **σ–ε**, solution, (b) unstable failure mechanism at the micro-scale, and (c) stabilized failure mechanism, restraining rotation.

On the other hand, the perturbed solution, $\dot{\boldsymbol{\sigma}}^{(2)}(\mathbf{x},t)$, is characterized by Eqs. (64); therefore, $\dot{\boldsymbol{\varepsilon}}^{(i)}(\mathbf{x},t)=\mathbf{0}$ so that substitution in Eq. (30)(a) yields:

$$\dot{\boldsymbol{\sigma}}^{(2)}(\mathbf{x},t)=\overline{\boldsymbol{\mathcal{C}}}:(\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t)+\dot{\boldsymbol{\chi}}^{*}(\mathbf{x}))=\underbrace{\overline{\boldsymbol{\mathcal{C}}}:[\mathbf{I}+\mathcal{M}(\boldsymbol{\varepsilon})]}_{\boldsymbol{\mathcal{C}}^{*}_{hom}}:\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t)=\boldsymbol{\mathcal{C}}^{*}_{hom}:\dot{\boldsymbol{\varepsilon}}(\mathbf{x},t). \tag{66}$$

Therefore, the fundamental and perturbed solutions in Eqs. (65) and (66) are characterized by different tangent moduli, $\boldsymbol{\mathcal{C}}^{tang}_{hom}\neq\boldsymbol{\mathcal{C}}^{*}_{hom}$ (see Fig. 24(a)).

Typically, this situation arises when the perturbation, $\dot{\tilde{\mathbf{u}}}^{*}_{\mu}(\mathbf{y})$ in Eq. (63), consists of a rigid body motion of $\mathcal{B}^{+}_{\mu}$ with respect to $\mathcal{B}^{-}_{\mu}$ (see Fig. 24(b)) i.e.:

$$\dot{\tilde{\mathbf{u}}}^{*}_{\mu}(\mathbf{y})=\mathcal{H}_{\mathcal{B}_{\mu,k}}(\mathbf{a}+\boldsymbol{\Omega}(\boldsymbol{\theta})\cdot\mathbf{y})\in\mathcal{V}_{\mu} \tag{67}$$

where $\mathbf{a}(t)$ stands for a, spatially constant, shift and $\boldsymbol{\Omega}(\boldsymbol{\theta}(t))$ is a, skew-symmetric, rotation tensor with axial vector $\boldsymbol{\theta}(t)$. From Eqs. (16) and (67):

$$\begin{aligned}\int_{\Gamma_{\mu}}(\dot{\tilde{\mathbf{u}}}^{*}_{\mu}\otimes\boldsymbol{\nu}_{\mu})^{S}d\Gamma&=\int_{\Omega_{\mu}}\boldsymbol{\nabla}\otimes^{S}\dot{\tilde{\mathbf{u}}}^{*}_{\mu}(\mathbf{y},t)d\Omega_{\mu}\\&=\int_{\Omega_{\mu}}\mathcal{H}_{\mathcal{B}_{\mu,k}}\underbrace{(\boldsymbol{\Omega}+\boldsymbol{\Omega}^{T})}_{=\mathbf{0}}d\Omega_{\mu}+\int_{\mathcal{S}_{\mu}(\mathbf{x})}\mathbf{n}_{\mu}(\mathbf{y})\otimes^{S}(\mathbf{a}+\boldsymbol{\Omega}(\boldsymbol{\theta})\cdot\mathbf{y})\,d\mathcal{S}_{\mu}=0.\end{aligned} \tag{68}$$

Eq. (68) is a homogeneous system of linear equations in the unknowns $\{\mathbf{a},\boldsymbol{\theta}\}$, which can be written as:

$$\begin{aligned}&\underbrace{\int_{\mathcal{S}_{\mu}(\mathbf{x})}\mathbf{n}_{\mu}(\mathbf{y})d\mathcal{S}_{\mu}}_{\mathbf{r}(\mathcal{S}_{\mu})}\otimes^{S}\mathbf{a}-\underbrace{\int_{\mathcal{S}_{\mu}(\mathbf{x})}(\mathbf{n}_{\mu}(\mathbf{y})\otimes^{S}\mathbf{y})d\mathcal{S}_{\mu}}_{\mathbf{L}(\mathcal{S}_{\mu})}\times\boldsymbol{\theta}=\mathbf{0}\qquad\text{(a)}\\&\rightarrow\mathbf{r}(\mathcal{S}_{\mu})\otimes^{S}\mathbf{a}+\mathbf{L}(\mathcal{S}_{\mu})\times\boldsymbol{\theta}=\mathbf{0}\qquad\text{(b)}\end{aligned} \tag{69}$$

which can be rephrased as:

$$[\mathbf{G}(\mathcal{S}_{\mu})]\cdot[\mathbf{a},\boldsymbol{\theta}]^{T}=\{\mathbf{0}\} \tag{70}$$

where $[\mathbf{G}(\mathcal{S}_{\mu})]$ is the corresponding matrix of coefficients, depending on the corresponding activated failure mechanism $\mathcal{S}_{\mu}$. If $[\mathbf{G}(\mathcal{S}_{\mu})]$ is rank deficient, Eq. (70) has non-trivial solutions, $\mathbf{a}^{*},\boldsymbol{\theta}^{*}$, each one characterizing an unstable mode $\dot{\tilde{\mathbf{u}}}^{*}_{\mu}(\mathbf{y})$ in Eq. (67). Inspection of Eq. (69) reveals that if the rotation angle is imposed to be zero ($\boldsymbol{\theta}=\mathbf{0}$)

then, the equation becomes

$$\begin{cases} \mathbf{r}(\mathcal{S}_\mu) \otimes^S \mathbf{a} = \mathbf{0} \\ \mathbf{r}(\mathcal{S}_\mu) = \displaystyle\int_{\mathcal{S}_\mu(\mathbf{x})} \mathbf{n}_\mu(\mathbf{y}) d\mathcal{S}_\mu \neq \mathbf{0} \end{cases} \Rightarrow \mathbf{a} = \mathbf{0} \tag{71}$$

and system (70) possesses only the trivial solution $\mathbf{a}^* = \mathbf{0}, \boldsymbol{\theta}^* = \mathbf{0}$. This result is exploited in the proposed technique to remove those modes described next.

Consistently with the fundamental result in Remark 2.6 ($\dot{\boldsymbol{\beta}}_\mu(\mathbf{x},\mathbf{y},t) = \dot{\boldsymbol{\beta}}_\mu(\mathbf{x},t),\ \forall t \geq t_B$) we impose on a selected set, $\mathcal{B}_{\mu,restr} = \bigcup_{i=1}^{n_{rest}} \mathcal{B}^{(i)}_{\mu,restr}$, $\mathcal{B}_{\mu,restr} \subset \mathcal{B}_{\mu,coh}$, of n_{restr} cohesive bands (see Fig. 2) the following restriction:

$$\dot{\boldsymbol{\beta}}^{(i)}_\mu(\eta(\mathbf{y}),t)\Big|_{\mathbf{y}\in\mathcal{B}^{(i)}_{\mu,restr}\subset\mathcal{B}_{\mu,restr}} \equiv [\![\dot{\tilde{\mathbf{u}}}_\mu(\xi(\mathbf{y}),\eta(\mathbf{y}),t)]\!]^+_- = \dot{\boldsymbol{\beta}}^{(i)}_\mu(t) \quad i = \{1,..,n_{restr}\};\ \forall t \geq t_B. \tag{72}$$

In Eq. (72) notation $[\![(\cdot)(\xi,\eta)]\!]^+_- \equiv (\cdot)(\xi,\eta)|_{\xi=k} - (\cdot)(\xi,\eta)|_{\xi=0}$ stands for the *apparent jump* between both sides of the cohesive band, see Fig. 2. Restriction in Eq. (72) translates into *clamping* (precluding rotation but allowing a constant fluctuation jump) domain $\mathcal{B}^+_\mu$ with respect to domain $\mathcal{B}^-_\mu$, by means of the restriction on the cohesive band $\mathcal{B}^{(i)}_{\mu,coh}$ (see Fig. 24(c)). This is enough to *remove* the instability modes *for any failure mechanism containing the cohesive band* $\mathcal{B}^{(i)}_{\mu,coh}$.[11]

Moreover, for practical purposes, and due to the small bandwidth k of the bands, Eq. (65) is imposed at the beginning of the analysis, and in total (non-rate) form, without substantial consequences in the results, i.e.:

$$\boldsymbol{\beta}^{(i)}_\mu(\eta(\mathbf{y}),t)\Big|_{\mathbf{y}\in\mathcal{B}^{(i)}_{\mu,coh}\subset\mathcal{B}_{\mu,restr}} \equiv [\![\tilde{\mathbf{u}}_\mu(\xi(\mathbf{y}),\eta(\mathbf{y}),t)]\!]^+_- = \boldsymbol{\beta}^{(i)}_\mu(t) \quad i = \{1,..,n_{restr}\};\ \forall t. \tag{73}$$

In this context, it can be readily proven that if any activated failure mechanism contents, at least, one of the cohesive bands of set $\mathcal{B}_{\mu,restr}$ in Eq. (73), than Eq. (53)(b) has only the trivial solution $\dot{\boldsymbol{\beta}}_\mu(\mathbf{y}) = 0$ and the instability does not appear. From the authors' experience, to prevent instabilities for any possible failure mechanism it is sufficient to include in the set $\mathcal{B}_{\mu,restr}$ *all cohesive bands intersecting the boundary of the failure cell* $\mathcal{B}_\mu$.[12]

As for the implementation of Eq. (73), the restriction can be *strongly imposed*, as an internal constraint in the elemental nodes, or *weakly imposed* via a penalty formulation. From the authors' experience, results are the same for both cases, though the second method turns out to be less code-invasive than the first one.

Appendix B. Strain injection techniques in multiscale problems

B.1. Weak discontinuity injection. Constant stress/discontinuous strain mode

B.1.1. Dipole generalized function

Let us denote $\chi_{\mathcal{S}}(\mathbf{x})$ as the *dipole generalized function* (generalized derivative of the Dirac's delta function) fulfilling

$$\int_{\mathcal{B}} \chi_{\mathcal{S}}\, \phi(\mathbf{x}) d\mathcal{B} = \int_{\mathcal{S}} [\![\phi]\!] d\mathcal{S} \tag{74}$$

for any sufficiently regular function, $\phi(\mathbf{x})$ exhibiting a jump, $[\![\phi]\!]$, across the discontinuity path, $\mathcal{S}$. Let us also define a two-parameter (h,k) sequence, $\chi^{h,k}_{\mathcal{S}}(\mathbf{x})$, in a band of thickness ℓ as:

$$\chi^{h,k}_{\mathcal{S}}(\mathbf{x}) = \begin{cases} -\dfrac{1}{h-k} & \mathbf{x} \in \mathcal{B}_{reg} \\ \dfrac{1}{k} & \mathbf{x} \in \mathcal{B}_{sing} \\ 0 & \text{otherwise} \end{cases} \tag{75}$$

[11] If quadrilateral finite elements are used for modeling the cohesive bands, Eq. (72) is equivalent to impose the fluctuation jump $[\![\dot{\tilde{\mathbf{u}}}_\mu(\xi,\eta)(\mathbf{y}_j,t)]\!]^+_-$ to be the same for all nodes, j, placed in the η direction of the selected element (see Fig. 6).

[12] The fact that any failure mechanism intersects the boundary of the failure cell is exploited for this statement.

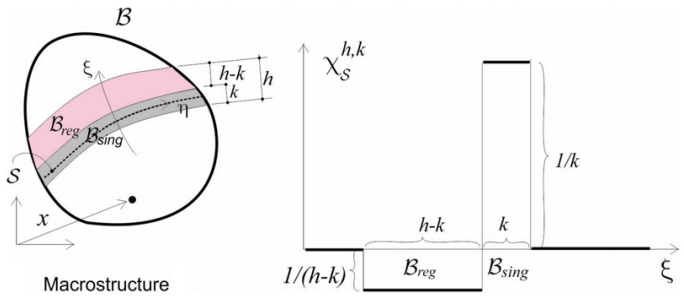


Fig. 25. Regularized dipole function $\chi_{\mathcal{S}}^{h,k}$.

which is sketched, for the 2D case, in Fig. 25. It can be readily checked that $\chi_{\mathcal{S}}^{h,k}(\mathbf{x})$ fulfills Eq. (74) in the limit $h \to 0$, $k \to 0$, so that it can be considered *a regularized sequence converging to the dipole function* in Eq. (74).

A very important property of the regularized dipole function $\chi_{\mathcal{S}}^{h,k}(\mathbf{x})$, to be exploited for subsequent purposes, is (see Fig. 25),

$$\int_{\mathcal{B}} \chi_{\mathcal{S}}^{h,k} d\mathcal{B} = 0. \tag{76}$$

B.1.2. Injection of a constant-stress/discontinuous-strain mode in quadrilateral elements

Let us consider the microscopic instantaneous density of dissipation, $\mathcal{D}_\mu(\mathbf{y},t)$, at the micro-scale:

$$\mathcal{D}_\mu(\mathbf{y},t) := \boldsymbol{\sigma}_\mu : \dot{\boldsymbol{\varepsilon}}_\mu - \dot{\psi}_\mu \geq 0 \tag{77}$$

where $\psi_\mu(\mathbf{y},t)$ stands for the microscopic free energy density of the corresponding constitutive model. From Eq. (77), and taking into account the Hill–Mandel equation (17), the *macroscopic instantaneous density of dissipation*, $\mathcal{D}(\mathbf{x},t)$, can be written as:

$$\begin{aligned} &\mathcal{D}(\mathbf{x},t) := \frac{1}{\Omega_\mu}\int_{\Omega_\mu} \mathcal{D}_\mu(\mathbf{y},t)\, dV = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \dot{\psi} \geq 0 \quad \text{(a)} \\ &\psi(\mathbf{x},t) := \frac{1}{\Omega_\mu}\int_{\Omega_\mu} \psi_\mu(\mathbf{y},t)\, dV \qquad\qquad\qquad\quad \text{(b)} \end{aligned} \tag{78}$$

where $\psi(\mathbf{x},t)$ is defined, through Eq. (78)(b), as the averaged (macroscopic) density of free energy, ψ_μ, at the microscale.

Then, considering a finite element discretization of the domain, $\mathcal{B}$, in quadrilateral elements, $\mathcal{B}^{(e)}$; $\mathcal{B} = \bigcup_{e=1}^{e=n_{elem}} \mathcal{B}^{(e)}$, the *injection domain*, $\mathcal{B}_{inj}$, is defined as

$$\mathcal{B}_{inj}(t) := \left\{ \bigcup_e \mathcal{B}^{(e)};\ \mathcal{D}(\mathbf{x}_c^{(e)},t) > 0 \right\} \tag{79}$$

where $\mathbf{x}_c^{(e)}$ stands for the barycenter of element (e). Eq. (79) defines $\mathcal{B}_{inj}(t)$ as the locus, at time t, of the elements of the macroscale experiencing inelastic behavior at the corresponding microscale level (evaluated at their barycenter $\mathbf{x}_c^{(e)}$). A subset of the injection domain, the *weak discontinuity domain* $\mathcal{B}_{wd}(t)$ is now defined as:

$$\mathcal{B}_{wd}(t) := \left\{ \bigcup_e \mathcal{B}^{(e)};\ \mathcal{B}^{(e)} \subset \mathcal{B}_{inj}(t);\ t \leq t_B(\mathbf{x}_c^{(e)}) \right\}. \tag{80}$$

Eq. (80) characterizes $\mathcal{B}_{wd}(t)$ as the set of elements of the injection domain $\mathcal{B}_{inj}(t)$ whose barycenter has not yet bifurcated, according to the definition of the bifurcation time, t_B, in Eq. (38).

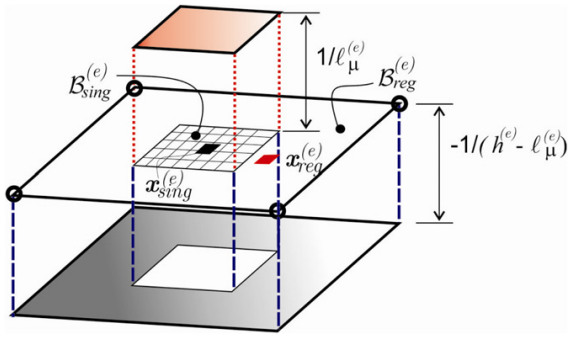


Fig. 26. Injected weak discontinuity mode. Elemental regularized dipole function $\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}$.

The points of $\mathcal{B}_{wd}$ are now endowed a weak discontinuity kinematics, by resorting to the assumed enhanced strain concept [50,51]. The rate of strain, injected in rate form at the typical element, e, with n_{node} nodes, is the following:

$$\dot{\boldsymbol{\varepsilon}}^{(e)}(\mathbf{x},t) \equiv \dot{\boldsymbol{\varepsilon}}_t^{(e)}(\mathbf{x}) = \underbrace{\sum_{i=1}^{i=n_{node}} \boldsymbol{\nabla} \mathrm{N}_i(\mathbf{x}) \otimes \dot{\mathbf{d}}_i(t)}_{(regular)} + \underbrace{\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x})\, \dot{\boldsymbol{\gamma}}^{(e)}(t)}_{(singular)} \quad \forall \mathcal{B}^{(e)} \subset \mathcal{B}_{wd}(t) \tag{81}$$

where N_i are the standard shape functions, $\dot{\mathbf{d}}_i(t)$, the nodal displacements, and $\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}$ is the element "e" counterpart of the regularized dipole-function in Eq. (75), for $k \equiv \ell_\mu^{(e)} = \ell_\mu(\mathbf{x}_c^{(e)})$, fulfilling (see Fig. 26):

$$\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x}) = \begin{cases} -\dfrac{1}{h^{(e)} - \ell_\mu^{(e)}} & \mathbf{x} \in \mathcal{B}_{reg}^{(e)} \\ \dfrac{1}{\ell_\mu^{(e)}} & \mathbf{x} \in \mathcal{B}_{sing}^{(e)} \\ 0 & \text{otherwise} \end{cases} \tag{82}$$

$$h^{(e)} = \frac{\Omega^{(e)}}{L^{(e)}}; \quad \Omega^{(e)} = meas(\mathcal{B}^{(e)}); \quad L^{(e)} = meas(\mathcal{S}^{(e)})$$

where $h^{(e)}$ is a characteristic element width, defined in terms of the measure (area/volume) of the element, $\Omega^{(e)}$, and the measure (length/area), $L^{(e)}$, of the element crack path $\mathcal{S}^{(e)}$. Notice that, by construction, the element dipole function, $\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}$, defined in Eq. (82) fulfills, at element level, the condition in Eq. (76), i.e.:

$$\int_{\mathcal{B}^{(e)}} \chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}\, d\mathcal{B} = 0. \tag{83}$$

The second term in Eq. (81) injects, by means of the mode $\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}$, a discontinuous strain field, intensified by the elemental strains $\dot{\boldsymbol{\gamma}}^{(e)}(t)$.

Notice that Eq. (81) can be rewritten as

$$\begin{cases} \dot{\boldsymbol{\varepsilon}}_t^{(e)}(\mathbf{x}) = \underbrace{\sum_{i=1}^{i=n_{node}} \boldsymbol{\nabla} \mathrm{N}_i(\mathbf{x}) \otimes^s \dot{\mathbf{d}}_i(t)}_{(compatible\ strains)} + \underbrace{\dot{\tilde{\boldsymbol{\gamma}}}_t^{(e)}(\mathbf{x})}_{(enhanced\ strains)} & \forall \mathcal{B}^{(e)} \subset \mathcal{B}_{wd}(t) \\ \dot{\tilde{\boldsymbol{\gamma}}}_t^{(e)} \equiv \dot{\tilde{\boldsymbol{\gamma}}}^{(e)}(\mathbf{x},t) = \chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x})\, \dot{\boldsymbol{\gamma}}_t^{(e)} \end{cases} \tag{84}$$

where the enhanced strains, $\dot{\tilde{\boldsymbol{\gamma}}}_t^{(e)}(\mathbf{x})$, fulfill (in view of Eq. (76)) the condition:

Table 3

Sampling values for the regularized dipole function, $\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}$, Dirac's delta function $\delta_{\mathcal{S}}^{\ell_\mu^{(e)}}$, and integration weights at the injection sampling points (s.p.).

Sampling point	Sampling value ($\chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}$)	Sampling value ($\delta_{\mathcal{S}}^{\ell_\mu^{(e)}}$)	Weight
Regular s.p.: $\mathbf{x}_{reg}^{(e)}$	$\begin{cases} -1/(h^{(e)}-\ell_\mu^{(e)}) \\ h^{(e)} = \dfrac{\Omega^{(e)}}{L^{(e)}} \end{cases}$	0	$\Omega^{(e)} - \ell_\mu^{(e)}\, L^{(e)}$
Singular s.p.: $\mathbf{x}_{sing}^{(e)}$	$1/\ell_\mu^{(e)}$	$1/\ell_\mu^{(e)}$	$\ell_\mu^{(e)}\, L^{(e)}$

$$\int_{\mathcal{B}^{(e)}} \dot{\tilde{\boldsymbol{\gamma}}}_t^{(e)}(\mathbf{x}) d\mathscr{B} = \mathbf{0} \quad \forall \mathcal{B}^{(e)} \tag{85}$$

this defining a specific class of assumed–enhanced strain elements [50]. Notice also the role of the characteristic length $\ell_\mu^{(e)} \equiv \ell_\mu(\mathbf{x}_c^{(e)})$, imported from the lower scale at the element barycenter, in the definition in (82) and (84), in accordance with Remark 2.4.

A subsequent refinement of the weak discontinuity injection in Eq. (81) consists of the sub-integration, at the element barycenter, of the first term of the right-hand-side of Eqs. (84). This can be interpreted as a *complementary injection of an element-wise constant strain field* for describing the compatible-strain part in Eq. (74). The resulting mixed finite-element problem in Box 3.2 can be solved in closed form, resulting [44]:

$$\begin{aligned} \dot{\boldsymbol{\varepsilon}}_t(\mathbf{x}) &= \sum_{\forall \mathcal{B}^{(e)} \in \mathcal{B}_{wd}(t)} \phi^{(e)} \Big(\underbrace{\dot{\bar{\boldsymbol{\varepsilon}}}^{(e)}(t)}_{(\textit{regular, constant, strains})} + \underbrace{\dot{\tilde{\boldsymbol{\gamma}}}^{(e)}(\mathbf{x},t)}_{(\textit{singular, enhanced strains})} \Big) \\ \dot{\bar{\boldsymbol{\varepsilon}}}^{(e)}(t) &= \sum_{i=1}^{i=n_{node}} \boldsymbol{\nabla} \mathrm{N}_i(\mathbf{x}_c^e) \otimes^s \dot{\mathbf{d}}_i(t) \\ \dot{\tilde{\boldsymbol{\gamma}}}^{(e)}(\mathbf{x},t) &= \chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x})\, \dot{\tilde{\boldsymbol{\gamma}}}_t^{(e)} \end{aligned} \tag{86}$$

where $\phi^{(e)}(\mathbf{x})$ is the characteristic function of element e ($\phi^{(e)}(\mathbf{x}) = 1\ \forall \mathbf{x} \in \mathcal{B}^{(e)}$, $\phi^{(e)}(\mathbf{x},t) = 0\ \forall \mathbf{x} \notin \mathcal{B}^{(e)}$).

It is well known that, for the considered quadrilateral elements, sub-integration of the compatible strains, may translate into violation of the inf–sup conditions [37] and the subsequent appearance of zero-energy modes polluting the solution (hourglass modes). However, it has to be emphasized that, in Eq. (86) *reduced integration is restricted to just a portion of the whole domain* (the weak discontinuity domain $\mathcal{B}_{wd}(t) \subset \mathcal{B}$ defined in Eq. (80)). This fact dramatically changes, in a favorable sense, the stability properties of the resulting formulation. *The possible hourglass modes do not appear whenever a sufficient portion of the domain is fully integrated* [31].

The resulting injection procedure is summarized in Box A1. It provides a finite element formulation, highly sensitive to propagation of localized strain fields. In spite that the actual kinematics of a strong discontinuity, fulfilling Eq. (3), is not included, and, therefore, some degree of stress locking could still appear [52] the resulting element provides reliable information for the crack path field problem in Box 3.1. Therefore, it is used as a first ingredient of the proposed crack-propagation algorithm.

B.1.3. Stress sampling and integration rule

The standard (four) Gauss quadrature rule, corresponding to full integration of two-dimensional quadrilaterals, is complemented with two additional sampling points (placed at the center of the element), see Figs. 26 and 27. They are termed the *singular sampling point*, denoted $\mathbf{x}_{sing}^{(e)}$, and *the regular sampling point*, denoted $\mathbf{x}_{reg}^{(e)}$ and sampling the rest of the element. Entities to be sampled (typically the stresses) are then additionally stored at these additional sampling points as it is done for regular sampling points.

Therefore, for the weak-discontinuity injected element, numerical integration (typically, evaluation of the incremental internal forces in terms of the stresses), is based on those two *injection-sampling points*, by defining the weights indicated in Table 3.

Box A1.
Weak-discontinuity injection variational problem (in rate form)

Problem.

GIVEN:

$$\hat{\mathcal{V}}_t := \left\{ \boldsymbol{\eta}_t(\mathbf{x}) = \sum_{i=1}^{n_{node}} N_i(\mathbf{x})\boldsymbol{\eta}_i \in [H^1(\mathcal{B})]^{n_{\dim}};\ \boldsymbol{\eta}_t(\mathbf{x})|_{\partial_u \mathcal{B}} = \dot{\mathbf{u}}^*(\mathbf{x},t) \right\}$$
$$\hat{\mathcal{V}}_0 := \{ \boldsymbol{\eta}(\mathbf{x}) \in H^1(\mathcal{B});\ \boldsymbol{\eta}(\mathbf{x})|_{\partial_u \mathcal{B}} = \mathbf{0} \}$$
$$\hat{\mathcal{E}} := \left\{ \delta\boldsymbol{\mu} = \sum_{e=1}^{n_{elem}} \phi^{(e)}(\mathbf{x})\, \delta\boldsymbol{\mu}^{(e)};\ \delta\boldsymbol{\mu}^{(e)} \in \mathbb{S}^{n_{\dim}\times n_{\dim}} \right\} \tag{87}$$
$$\tilde{\Gamma} := \left\{ \delta\tilde{\boldsymbol{\gamma}} = \sum_{e=1}^{n_{elem}} \chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x})\ \delta\tilde{\boldsymbol{\gamma}}^{(e)};\ \delta\tilde{\boldsymbol{\gamma}}^{(e)} \in \mathbb{S}^{n_{\dim}\times n_{\dim}} \right\}$$

FIND

$$\dot{\mathbf{u}}_t(\mathbf{x}) \equiv \dot{\mathbf{u}}(\mathbf{x},t) : \mathcal{B}\times[0,T] \to \mathbb{R}^{n_{\dim}};\quad \dot{\mathbf{u}}_t \in \hat{\mathcal{V}}_t$$
$$\dot{\bar{\boldsymbol{\varepsilon}}}_t(\mathbf{x}) \equiv \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) : \mathcal{B}\times[0,T] \to \mathbb{S}^{n_{\dim}\times n_{\dim}};\quad \dot{\bar{\boldsymbol{\varepsilon}}}_t \in \hat{\mathcal{E}}$$
$$\dot{\tilde{\boldsymbol{\gamma}}}_t(\mathbf{x}) \equiv \dot{\tilde{\boldsymbol{\gamma}}}(\mathbf{x},t) : \mathcal{B}\times[0,T] \to \mathbb{S}^{n_{\dim}\times n_{\dim}};\quad \dot{\tilde{\boldsymbol{\gamma}}}_t \in \tilde{\Gamma}$$

FULFILLING:

$$\int_{\mathcal{B}} \boldsymbol{\nabla}^s\boldsymbol{\eta} : \dot{\boldsymbol{\sigma}}_t\, d\mathcal{B} - W^{ext}(\boldsymbol{\eta}, \dot{\mathbf{b}}, \dot{\mathbf{t}}^*) = 0 \quad \forall\boldsymbol{\eta} \in \hat{\mathcal{V}}_0 \tag{a}$$

$$\begin{cases} \displaystyle\int_{\mathcal{B}} \delta\boldsymbol{\mu} : (\dot{\bar{\boldsymbol{\varepsilon}}}_t - \boldsymbol{\nabla}^S\dot{\mathbf{u}}_t(\mathbf{x}))\, d\mathcal{B} = 0 \quad \forall\delta\boldsymbol{\mu} \in \hat{\mathcal{E}} \\ \dot{\boldsymbol{\varepsilon}}_t(\mathbf{x},t) = \dot{\bar{\boldsymbol{\varepsilon}}}_t(\mathbf{x}) + \dot{\tilde{\boldsymbol{\gamma}}}_t(\mathbf{x}) \\ \begin{cases} \dot{\bar{\boldsymbol{\varepsilon}}}_t(\mathbf{x}) = \displaystyle\sum_{e=1}^{n_{elem}} \phi^{(e)}(\mathbf{x})\, \dot{\bar{\boldsymbol{\varepsilon}}}_t^{(e)} \rightarrow \textit{assumed (regular) strain} \\ \dot{\tilde{\boldsymbol{\gamma}}}_t(\mathbf{x}) = \displaystyle\sum_{e=1}^{n_{elem}} \chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x})\, \dot{\tilde{\boldsymbol{\gamma}}}_t^{(e)} \in \tilde{\Gamma} \rightarrow \textit{enhanced (singular) strain} \end{cases} \end{cases} \tag{b}$$

$$\begin{cases} \displaystyle\int_{\mathcal{B}} \delta\tilde{\boldsymbol{\gamma}} : \dot{\boldsymbol{\sigma}}_t d\mathcal{B} = \sum_{e=1}^{n_{elem}} \delta\tilde{\boldsymbol{\gamma}}^{(e)} \int_{\mathcal{B}^{(e)}} \chi_{\mathcal{S}}^{h^{(e)},\ell_\mu^{(e)}}(\mathbf{x}) : \dot{\boldsymbol{\sigma}}_t^{(e)} d\mathcal{B} \\ = \displaystyle\sum_{e=1}^{n_{elem}} \int_{\mathcal{S}^{(e)}} \delta\boldsymbol{\gamma}^{(e)} : [\![\dot{\boldsymbol{\sigma}}_t]\!]_{\mathcal{S}^{(e)}}\, d\Gamma = 0 \quad \forall\delta\boldsymbol{\gamma}^{(e)} \in \mathbb{S}^{n_{\dim}\times n_{\dim}} \end{cases} \tag{c}$$

$$\dot{\boldsymbol{\sigma}}(\mathbf{x},t) \equiv \dot{\boldsymbol{\sigma}}_t(\mathbf{x}) = \dot{\boldsymbol{\Sigma}}(\dot{\boldsymbol{\varepsilon}}_t(\mathbf{x})) \rightarrow \textit{constitutive equation} \tag{d}$$

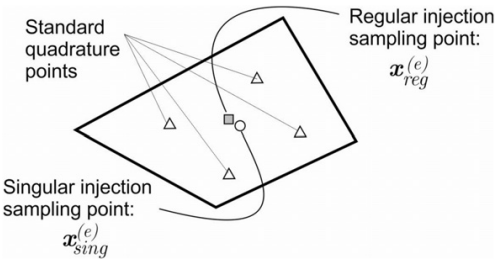


Fig. 27. Sampling points involved in the numerical integration.

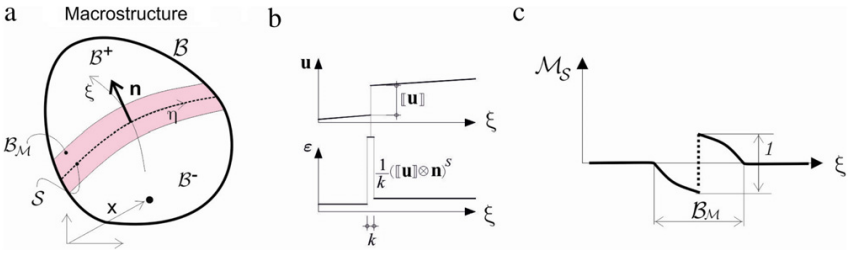


Fig. 28. (a) Body exhibiting a strong discontinuity. (b) Strong discontinuity kinematics (c) Unit jump function.

Replacement of the integration rule, in Table 3, into Eq. (88)(b)–(c) yields the element-wise equation

$$\begin{cases} \displaystyle\int_{\mathcal{S}^{(e)}} \delta\gamma^{(e)} : [\![\dot{\boldsymbol{\sigma}}_t]\!]^{(e)}\, d\Gamma = 0 \Rightarrow & \text{(a)} \\ [\![\dot{\boldsymbol{\sigma}}_t]\!]^{(e)} \equiv \dot{\boldsymbol{\sigma}}_t(\mathbf{x}^{(e)}_{\ell_\mu}) - \dot{\boldsymbol{\sigma}}_t(\mathbf{x}^{(e)}_h) = \mathbf{0} \Rightarrow & \text{(b)} \\ \begin{cases} \dot{\boldsymbol{\Sigma}}_{sing}\left(\dot{\bar{\boldsymbol{\varepsilon}}}^{(e)}_t + \dfrac{\dot{\tilde{\gamma}}^{(e)}_t}{\ell^{(e)}_\mu}\right) = \dot{\boldsymbol{\Sigma}}_{reg}\left(\dot{\bar{\boldsymbol{\varepsilon}}}^{(e)}_t - \dfrac{\dot{\tilde{\gamma}}^{(e)}_t}{h^{(e)} - \ell^{(e)}_\mu}\right) \\ \dot{\bar{\boldsymbol{\varepsilon}}}^{(e)}_t = \displaystyle\sum_{i=1}^{i=4} \boldsymbol{\nabla} \mathrm{N}_i(\mathbf{x}^e_c) \otimes^s \dot{\mathbf{d}}_i(t) \end{cases} & \text{(c)} \end{cases} \tag{89}$$

where $\boldsymbol{\Sigma}_{sing}$ and $\boldsymbol{\Sigma}_{reg}$ stand for the stress-evaluation via the continuum constitutive model at the regular and singular sampling points, respectively, as stated by the continuum strong discontinuity approach. Typically, $\boldsymbol{\Sigma}_{sing}$ is made inelastic with strain softening (the softening modulus is regularized with the characteristic length, ℓ_μ), and $\boldsymbol{\Sigma}_{reg}$ is enforced to be instantaneously elastic (for both loading and unloading).

Eq. (89)(b) states that the (incremental) stress field is constant all over the element, in spite of the imposed discontinuous strain field in Fig. 26. In addition, Eq. (89)(c) supplies an elementary-uncoupled additional equation for solving the strain jump, $\dot{\tilde{\gamma}}^{(e)}_t$, in terms of the regular elemental strain, $\dot{\bar{\boldsymbol{\varepsilon}}}^{(e)}_t = \sum_{i=1}^{i=4} \boldsymbol{\nabla} \mathrm{N}_i(\mathbf{x}^e_c) \otimes^s \dot{\mathbf{d}}_i(t)$. Therefore, the additional degrees of freedom, $\dot{\tilde{\gamma}}^{(e)}_t$, corresponding to the enhanced strain can be condensed at element level following standard procedures [50,53,54].

B.2. Injection of a propagating strong discontinuity

In a second stage, a propagating strong discontinuity is injected in the following domain:

$$\mathcal{B}_{sd}(t) := \left\{ \bigcup_e \mathcal{B}^{(e)};\ \mathcal{B}^{(e)} \subset \mathcal{B}_{inj}(t);\ t > t_B(\mathbf{x}^{(e)}_c) \right\}. \tag{90}$$

Comparing Eqs. (90), (79) and (80) one realizes that *the strong discontinuity injection takes place as soon as a weak-discontinuity-injected element bifurcates* according to the problem in Eq. (37). In other words, as soon as the homogenized strain field at the upper scale is compatible with a strong discontinuity kinematics (see Eqs. (37) and (40)).

Let us now consider the continuum body, $\mathcal{B}$, (see Fig. 28(a)) split by the strong discontinuity path $\mathcal{S}$ into two parts, $\mathcal{B}^+$ and $\mathcal{B}^-$. Across $\mathcal{S}$, the rate of displacement field, $\dot{\mathbf{u}}(\mathbf{x})$, experiences a jump $[\![\dot{\mathbf{u}}]\!] = \dot{\mathbf{u}}|_{\mathbf{x}\in(\partial\mathcal{B}^+\cap\mathcal{S})} - \dot{\mathbf{u}}|_{\mathbf{x}\in(\partial\mathcal{B}^-\cap\mathcal{S})}$.

The kinematic description of the displacement field in $\mathcal{B}$, (see Fig. 28(a)) reads:

$$\dot{\mathbf{u}} = \bar{\dot{\mathbf{u}}} + \mathcal{H}_{\mathcal{S}}\, [\![\dot{\mathbf{u}}]\!] \tag{91}$$

where $\bar{\dot{\mathbf{u}}}$ stands for the smooth part of the displacement field, $\dot{\boldsymbol{\beta}} \equiv [\![\dot{\mathbf{u}}]\!] = \dot{\mathbf{u}}|_{\mathbf{x}\in(\partial\mathcal{B}^+\cap\mathcal{S})} - \dot{\mathbf{u}}|_{\mathbf{x}\in(\partial\mathcal{B}^-\cap\mathcal{S})}$ stands for the displacement jump and $\mathcal{H}_{\mathcal{S}}$ is the Heaviside (step) function, shifted to $\mathcal{S}$. Due to computational reasons, related to the imposition of the essential boundary conditions, it is convenient to re-formulate Eq. (91) as the following equivalent

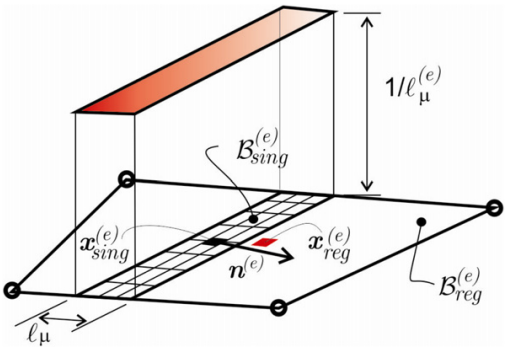


Fig. 29. Injected strong discontinuity mode. Elemental regularized Dirac's delta function $\delta_{\mathcal{S}}^{\ell_\mu^{(e)}}$.

expression, see [36] for further details,

$$\dot{\mathbf{u}} = \underbrace{\hat{\dot{\mathbf{u}}} - \varphi[\![\dot{\mathbf{u}}]\!]}_{\tilde{\dot{\mathbf{u}}}} + \mathcal{H}_{\mathcal{S}}[\![\dot{\mathbf{u}}]\!] = \hat{\dot{\mathbf{u}}} + \underbrace{(\mathcal{H}_{\mathcal{S}} - \varphi)}_{\mathcal{M}_{\mathcal{S}}}\dot{\boldsymbol{\beta}} \tag{92}$$

where $\hat{\dot{\mathbf{u}}}$ fulfills the Dirichlet boundary conditions of the problem, φ is a continuous, in principle arbitrary, function fulfilling:

$$\varphi(\mathbf{x}) = \begin{cases} 0 & \forall \mathbf{x} \in (\mathcal{B} \setminus \mathcal{B}_{\mathcal{M}})^{-} \\ 1 & \forall \mathbf{x} \in (\mathcal{B} \setminus \mathcal{B}_{\mathcal{M}})^{+}. \end{cases} \tag{93}$$

In Eq. (93) $\mathcal{M}_{\mathcal{S}}(\mathbf{x}) = \mathcal{H}_{\mathcal{S}} - \varphi(\mathbf{x})$ is the *unit jump function*, whose support is $\mathcal{B}_{\mathcal{M}}$ and exhibits a unit jump across $\mathcal{S}$, see Fig. 28. In Eq. (93), $\varphi(\mathbf{x})$ is the so-called *indicatrix function* [31]. The (infinitesimal) strain field corresponding to Eq. (92) reads:

$$\dot{\varepsilon} = \boldsymbol{\nabla}^S \dot{\mathbf{u}} = \underbrace{\boldsymbol{\nabla}^S \hat{\dot{\mathbf{u}}} + (\mathcal{M}_{\mathcal{S}} \otimes \boldsymbol{\nabla}\dot{\boldsymbol{\beta}})^S - (\boldsymbol{\nabla}\varphi \otimes \dot{\boldsymbol{\beta}})^S}_{\hat{\dot{\boldsymbol{\varepsilon}}}\ (regular)} + \underbrace{\delta_{\mathcal{S}}(\mathbf{n} \otimes \dot{\boldsymbol{\beta}})^S}_{(singular)} = \hat{\dot{\varepsilon}} + \delta_{\mathcal{S}} \left(\mathbf{n} \otimes \dot{\boldsymbol{\beta}}\right)^S. \tag{94}$$

In the present multiscale context, the proposed injection procedure consists of the incremental injection of *the elemental strong discontinuity mode*,

$$\begin{aligned} \dot{\boldsymbol{\varepsilon}}_t^{(e)} &= \underbrace{\hat{\dot{\boldsymbol{\varepsilon}}}^{(e)}(t)}_{\substack{(regular,\ constant,\\ strains)}} + \underbrace{\dot{\tilde{\gamma}}^{(e)}(\mathbf{x}, t)}_{\substack{(singular,\ enhanced\\ strains)}} \quad \forall \mathcal{B}^{(e)} \in \mathcal{B}_{sd}(t) \\ \hat{\dot{\boldsymbol{\varepsilon}}}^{(e)}(t) &= \sum_{i=1}^{i=4} \boldsymbol{\nabla} \mathrm{N}_i(\mathbf{x}_c^e) \otimes^s \dot{\mathbf{d}}_i(t) \\ \dot{\tilde{\gamma}}_t^{(e)} &\equiv \dot{\tilde{\gamma}}^{(e)}(\mathbf{x}, t) = \delta_{\mathcal{S}}^{\ell_\mu^{(e)}} (\dot{\boldsymbol{\beta}}^{(e)} \otimes^S \mathbf{n}^{(e)})(\mathbf{x}) \end{aligned} \tag{95}$$

in terms of the $\ell_\mu^{(e)}$-regularized Dirac's delta function, $\delta_{\mathcal{S}}^{\ell_\mu^{(e)}}$, displayed in Fig. 29, and Table 3, fulfilling

$$\delta_{\mathcal{S}}^{\ell_\mu^{(e)}} = \begin{cases} \dfrac{1}{\ell_\mu^{(e)}} & \mathbf{x} \in \mathcal{B}_{sing}^{(e)} \\ 0 & \mathbf{x} \in \mathcal{B}_{reg}^{(e)}. \end{cases} \tag{96}$$

Notice the directional, and placement-dependent, character of the injected mode in Eqs. (95)–(96) and Fig. 29, in contrast with the *isotropic* weak-discontinuity mode in Eqs. (86) and Fig. 26. Precise information of this placement, ensuring the spatial continuity of the injected crack-path is a fundamental requirement in the method. The direction

Box A2.
Strong-discontinuity injection variational problem (in rate form)

Problem.

$$
\begin{aligned}
&\textit{GIVEN}:\\
&\hat{\mathcal{V}}_t := \left\{ \boldsymbol{\eta}_t(\mathbf{x}) = \sum_{i=1}^{n_{node}} N_i(\mathbf{x})\, \boldsymbol{\eta}_i \in [H^1(\mathcal{B})]^{n_{\dim}};\ \boldsymbol{\eta}_t(\mathbf{x})|_{\partial_u \mathcal{B}} = \dot{\mathbf{u}}^*(\mathbf{x},t) \right\}\\
&\hat{\mathcal{V}}_0 := \{ \boldsymbol{\eta}(\mathbf{x}) \in H^1(\mathcal{B});\ \boldsymbol{\eta}(\mathbf{x})|_{\partial_u \mathcal{B}} = \mathbf{0} \}\\
&\hat{\mathcal{E}} := \left\{ \delta\boldsymbol{\mu} = \sum_{e=1}^{n_{elem}} \phi^{(e)}(\mathbf{x})\, \delta\boldsymbol{\mu}^{(e)};\ \delta\boldsymbol{\mu}^{(e)} \in \mathbb{S}^{n_{\dim}\times n_{\dim}} \right\}\\
&\tilde{\Gamma} := \left\{ \delta\tilde{\boldsymbol{\gamma}} = \sum_{e=1}^{n_{elem}} \delta_{\mathcal{S}}^{\ell_\mu^{(e)}}(\mathbf{x})\, (\delta\boldsymbol{\beta}^{(e)} \otimes^S \mathbf{n}^{(e)});\ \delta\boldsymbol{\beta}^{(e)} \in \mathbb{R}^{n_{\dim}} \right\}\\
&\textit{FIND}\\
&\dot{\mathbf{u}}_t(\mathbf{x}) \equiv \dot{\mathbf{u}}(\mathbf{x},t) : \mathcal{B}\times[0,T] \to \mathbb{R}^{n_{\dim}};\quad \dot{\mathbf{u}}_t \in \hat{\mathcal{V}}\\
&\dot{\bar{\boldsymbol{\varepsilon}}}_t(\mathbf{x}) \equiv \dot{\bar{\boldsymbol{\varepsilon}}}(\mathbf{x},t) : \mathcal{B}\times[0,T] \to \mathbb{S}^{n_{\dim}\times n_{\dim}};\quad \dot{\bar{\boldsymbol{\varepsilon}}}_t \in \hat{\mathcal{E}}\\
&\dot{\boldsymbol{\beta}}_t(\mathbf{x}) \equiv \dot{\boldsymbol{\beta}}(\mathbf{x},t) : \mathcal{B}\times[0,T] \to \mathbb{R}^{n_{\dim}};\quad \dot{\boldsymbol{\beta}}_t \in \mathbb{R}^{n_{\dim}}
\end{aligned}
\tag{97}
$$

FULFILLING:

$$
\begin{aligned}
&\int_{\mathcal{B}} \boldsymbol{\nabla}^s \boldsymbol{\eta} : \dot{\boldsymbol{\sigma}}_t\, d\mathcal{B} - W^{ext}(\boldsymbol{\eta}, \dot{\mathbf{b}}, \dot{\mathbf{t}}^*) = 0 \quad \forall \boldsymbol{\eta} \in \hat{\mathcal{V}}_0 \qquad \text{(a)}\\
&\begin{cases}
\displaystyle\int_{\mathcal{B}} \delta\boldsymbol{\mu} : [\dot{\bar{\boldsymbol{\varepsilon}}}_t - \nabla^S \dot{\mathbf{u}}(\mathbf{x}) - (\nabla\varphi(\mathbf{x}) \otimes^s \dot{\boldsymbol{\beta}}_t(\mathbf{x}))]\, d\mathcal{B} = 0 \quad \forall \delta\boldsymbol{\mu} \in \hat{\mathcal{E}}\\
\dot{\boldsymbol{\varepsilon}}_t(\mathbf{x},t) = \dot{\bar{\boldsymbol{\varepsilon}}}_t(\mathbf{x}) + \dot{\tilde{\boldsymbol{\gamma}}}_t(\mathbf{x})\\
\quad \begin{cases}
\dot{\bar{\boldsymbol{\varepsilon}}}_t(\mathbf{x}) = \displaystyle\sum_{\forall \mathcal{B}^{(e)} \subset \mathcal{B}_{sd}(t)} \phi^{(e)}(\mathbf{x})\, \dot{\bar{\boldsymbol{\varepsilon}}}_t^{(e)} \rightarrow \textit{assumed (regular) strain} \qquad \text{(b)}\\
\dot{\tilde{\boldsymbol{\gamma}}}_t(\mathbf{x}) = \displaystyle\sum_{\forall \mathcal{B}^{(e)} \subset \mathcal{B}_{sd}(t)} \delta_{\mathcal{S}}^{\ell_\mu^{(e)}} (\dot{\boldsymbol{\beta}}_t^{(e)} \otimes^S \mathbf{n}^{(e)})(\mathbf{x}) \in \tilde{\Gamma} \rightarrow \textit{assumed (singular) strain}
\end{cases}
\end{cases}\\
&\begin{cases}
\displaystyle\int_{\mathcal{B}} (\delta\dot{\boldsymbol{\beta}} \otimes^S \mathbf{n}^{(e)}) : \chi_{\mathcal{S}}^{h,\ell_\mu} \dot{\boldsymbol{\sigma}}_t d\mathcal{B} = \sum_{e=1}^{n_{elem}} (\delta\boldsymbol{\beta}^{(e)} \otimes^S \mathbf{n}^{(e)}) : \int_{\mathcal{B}^{(e)}} \chi_{\mathcal{S}}^{h,\ell_\mu^{(e)}}(\mathbf{x}) \dot{\boldsymbol{\sigma}}_t^{(e)} d\mathcal{B}\\
= \displaystyle\sum_{e=1}^{n_{elem}} \int_{\mathcal{S}^{(e)}} \delta\boldsymbol{\beta}^{(e)}\, [\![\dot{\boldsymbol{\sigma}}_t \cdot \mathbf{n}^{(e)}]\!]_{\mathcal{S}^{(e)}}\, d\Gamma = 0 \quad \forall \delta\boldsymbol{\beta}^{(e)} \in \mathbb{R}^{n_{\dim}}
\end{cases}\\
&\dot{\boldsymbol{\sigma}}(\mathbf{x},t) \equiv \dot{\boldsymbol{\sigma}}_t(\mathbf{x}) = \dot{\boldsymbol{\Sigma}}(\dot{\boldsymbol{\varepsilon}}_t(\mathbf{x})) \rightarrow \textit{constitutive equation}
\end{aligned}
\tag{98}
$$

of the element normal, $\mathbf{n}^{(e)}$, is provided by the solution, $\mathbf{n}$, of the discontinuous bifurcation problem in Eq. (38), evaluated at the center of the element $\mathbf{x}_c^{(e)}$ and at the bifurcation time $t_B(\mathbf{x}_c^{(e)})$. The resulting injection procedure is summarized in Box A2. The integration rule is the one displayed in Table 3 and Fig. 27.

B.3. Space and time integration

As commented above, injection of weak-discontinuity and strong-discontinuity modes, in Sections 1 and 2, requires, in principle, specific integration rules in space, i.e.: a standard four-point Gauss quadrature rule, $\mathbf{x}_i$, $i = (1,2,3,4)$, in $\mathcal{B} \setminus \mathcal{B}_{inj}(t)$, and the two additional sampling/injection points, $\mathbf{x}_{sing}^{(e)}$ and $\mathbf{x}_{reg}^{(e)}$, for injected elements, $\mathcal{B}^{(e)} \subset \mathcal{B}_{wd}(t)$ and $\mathcal{B}^{(e)} \subset \mathcal{B}_{sd}(t)$ so that $\mathcal{B}_{inj} = \mathcal{B}_{wd} \cup \mathcal{B}_{sd}$ (see Table 3, Figs. 25 and 26).

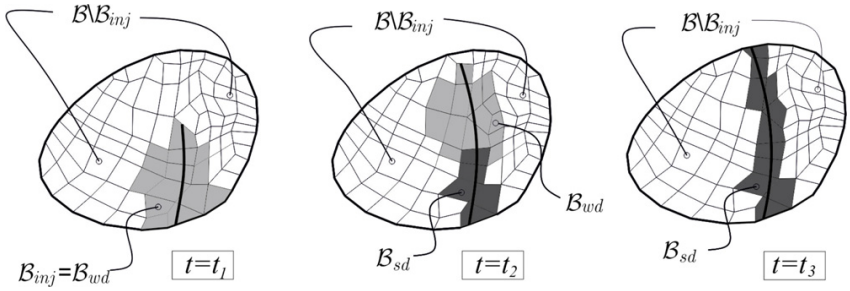


Fig. 30. Evolution of the injection domains for three typical stages of the discontinuity propagation ($t_1 < t_2 < t_3$).

This domain-specific integration rule can become cumbersome in two senses:

(1) Domains $\mathcal{B} \setminus \mathcal{B}_{inj}(t)$, $\mathcal{B}_{wd}(t)$ and $\mathcal{B}_{sd}(t)$ change with time (see Fig. 30). This poses, in principle, some additional problems on the time-integration of the resulting rate of the mechanical balance of forces.
(2) The implementation of those specific integration rules in a standard finite element code becomes code-invasive.

These flaws can be readily overcome by the following procedure. After standard manipulations, the discrete (finite element) version of problem in Eqs. (60)(a), (88)(a) and (98)(a) reads:

$$\begin{aligned}\dot{\boldsymbol{R}}_{mech}(\dot{\mathbf{d}}(t)) = &\int_{\mathcal{B}\backslash\mathcal{B}_{inj}(t)} \mathbf{B}^T(\mathbf{x}) \cdot \{\dot{\boldsymbol{\sigma}}(\mathbf{x}, \dot{\mathbf{d}}(t))\} d\mathcal{B} \\ &+ \int_{\mathcal{B}_{wd}(t)} \mathbf{B}^T(\mathbf{x}) \cdot \{\dot{\boldsymbol{\sigma}}(\mathbf{x}, \dot{\mathbf{d}}(t))\} d\mathcal{B} + \int_{\mathcal{B}_{sd}(t)} \mathbf{B}^T(\mathbf{x}) \cdot \{\dot{\boldsymbol{\sigma}}(\mathbf{x}, \dot{\mathbf{d}}(t))\} d\mathcal{B} - \dot{\mathbf{F}}^{ext}(t) = \mathbf{0}\end{aligned} \tag{99}$$

where, $\boldsymbol{R}_{mech}(\mathbf{d}(t))$ stands for the mechanical residue (unbalanced forces), $\mathbf{d}(t)$ is the vector of nodal displacements, $\mathbf{B}(\mathbf{x})$ stands for the classical deformation matrix, $\{\boldsymbol{\sigma}^h(\mathbf{x}, \mathbf{d}(t))\}$ are the homogenized stresses (in Voigt's notation) computed in accordance with the corresponding strain injections and $\mathbf{F}^{ext}$ are the external forces. Notice that, dependence on time of the integration domains, $\mathcal{B}_{inj}(t)$, $\mathcal{B}_{wd}(t)$ and $\mathcal{B}_{sd}(t)$, in Eq. (99), makes time integration of the residue equation ($\dot{\boldsymbol{R}}(t)$; $t \in [0, t_{n+1}] \rightarrow \boldsymbol{R}(t_{n+1}) = \mathbf{0}$) a sensitive issue.

However, this issue can be easily solved by appropriately rephrasing the integral kernels in Eq. (99). In [31] it is proven that, by defining some *specific stress entities* (the so called *effective stresses,* $\tilde{\boldsymbol{\sigma}}(\mathbf{x}_i, t)$) at the standard Gauss points, the spatial integration in Eq. (99) *can be rephrased to* a standard four Gauss points integration rule in *the whole integration domain* $\mathcal{B}$ (and, therefore, not time dependent) i.e.

$$\dot{\boldsymbol{R}}_{mech}(t) \equiv \dot{\boldsymbol{R}}_{mech,t} = \int_{\mathcal{B}} \mathbf{B}^T(\mathbf{x}) \cdot \{\dot{\tilde{\boldsymbol{\sigma}}}(\mathbf{x}, \dot{\mathbf{d}}(t))\} d\mathcal{B} - \dot{\mathbf{F}}^{ext}(t) = \mathbf{0}. \tag{100}$$

Now, Eq. (100) can be exactly integrated[13] along the time interval $[0, t_{n+1}]$ as:

$$\begin{aligned}\boldsymbol{R}_{mech}(\mathbf{d}_{n+1}) &= \int_{\mathcal{B}} \mathbf{B}^T(\mathbf{x}) \cdot \{\tilde{\boldsymbol{\sigma}}_{n+1}(\mathbf{x}, \mathbf{d}_{n+1})\} d\mathcal{B} - \mathbf{F}^{ext}(t) \\ &= \sum_{\forall \mathcal{B}^{(e)} \subset \mathcal{B}} \int_{\mathcal{B}^{(e)}} \mathbf{B}^{(e)^T}(\mathbf{x}) \cdot \{\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}(\mathbf{x})\} d\mathcal{B} - \mathbf{F}^{ext}_{n+1} = \mathbf{0}\end{aligned} \tag{101}$$

which returns the classical format for the unbalanced residual forces at the current time t_{n+1}. The specific rules for updating the effective stresses, $\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}$ appearing in Eq. (101), in terms of the homogenized stresses, $\boldsymbol{\sigma}^{(e)}_{n+1}$ at all sampling points, are given in Box A3.

[13] Assuming zero initial balanced forces.

Box A3.
Effective stress evaluation

DATA: $\mu_{n+1}(\mathbf{x})$, $\Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x})$, $\Delta\boldsymbol{\gamma}^{(e)}_{n+1}$, $\Delta\beta^{(e)}_{n+1}$, $\tilde{\boldsymbol{\sigma}}^{(e)}_{n}(\mathbf{x}_i)$, $\boldsymbol{\sigma}^{(e)}_{n}(\mathbf{x}_i)$, $\mathcal{B}_{wd}(t_{n+1})$, $\mathcal{B}_{sd}(t_{n+1})$
OUTPUT: $\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}(\mathbf{x}_i)$, $\boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}_i)$

(1) *Compute the elemental indicatrix function*:

$$\mu_{n+1}(\mathbf{x}) \rightarrow \varphi^{(e)}_{n+1}(\mathbf{x}) = \sum_{i=1}^{n_{mode}} N_i(\mathbf{x})\, \varphi^{(e)}_i(\mu_{n+1})$$

(2) *Compute strains*, $\boldsymbol{\varepsilon}^{(e)}_{n+1}$, *at all sampling points*

$$\boldsymbol{\varepsilon}^{(e)}_{n+1}(\mathbf{x}_i) = \boldsymbol{\varepsilon}^{(e)}_{n}(\mathbf{x}_i) + \Delta\boldsymbol{\varepsilon}^{(e)}_{n}(\mathbf{x}_i) \quad \mathbf{x}_i \equiv \mathbf{x}_{G_i}\ (i=1,...,4),\ \mathbf{x}^{(e)}_{reg},\ \mathbf{x}^{(e)}_{sing}$$

$$\Delta\boldsymbol{\varepsilon}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{reg}) = \begin{cases} \nabla^S \Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x}^{(e)}_{reg}) & \mathcal{B}^{(e)} \in \mathcal{B} \setminus (\mathcal{B}_{wd} \cup \mathcal{B}_{sd})(t_{n+1}) \\ \nabla^S \Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x}^{(e)}_{reg}) + \chi_{\mathcal{S}}^{h^{(e)},\ell^{(e)}_{\mu}}(\mathbf{x}^{(e)}_{reg})\, \Delta\boldsymbol{\gamma}^{(e)}_{n+1} & \mathcal{B}^{(e)} \in \mathcal{B}_{wd}(t_{n+1}) \\ \nabla^S \Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x}^{(e)}_{reg}) - (\boldsymbol{\nabla}\varphi^{(e)}_{n+1}(\mathbf{x}^{(e)}_{reg}) \otimes \Delta\boldsymbol{\beta}^{(e)}_{n+1})^S & \mathcal{B}^{(e)} \in \mathcal{B}_{sd}(t_{n+1}) \end{cases}$$

$$\Delta\boldsymbol{\varepsilon}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{sing}) = \begin{cases} \nabla^S \Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x}^{(e)}_{sing}) & \mathcal{B}^{(e)} \in \mathcal{B} \setminus (\mathcal{B}_{wd} \cup \mathcal{B}_{sd})(t_{n+1}) \\ \nabla^S \Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x}^{(e)}_{sing}) + \chi_{\mathcal{S}}^{h^{(e)},\ell^{(e)}_{\mu}}(\mathbf{x}^{(e)}_{sing})\, \Delta\boldsymbol{\gamma}^{(e)}_{n+1} & \mathcal{B}^{(e)} \in \mathcal{B}_{wd}(t_{n+1}) \\ \nabla^S \Delta\hat{\mathbf{u}}_{n+1}(\mathbf{x}^{(e)}_{sing}) + \left[\dfrac{\mathbf{n}^{(e)}}{\ell^{(e)}_{\mu}} - (\boldsymbol{\nabla}\varphi^{(e)}_{n+1}(\mathbf{x}^{(e)}_{sing})\right] \otimes \Delta\boldsymbol{\beta}^{(e)}_{n+1})^S & \mathcal{B}^{(e)} \in \mathcal{B}_{sd}(t_{n+1}) \end{cases}$$

(3) *Compute regular stresses*, $\boldsymbol{\sigma}^{(e)}_{n+1}$, *at all sampling points*

$$\boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}_i) = \boldsymbol{\Sigma}[\boldsymbol{\varepsilon}^{(e)}_{n+1}(\mathbf{x}_i)], \quad \mathbf{x}_i \equiv \mathbf{x}_{G_i}\ (i=1,...,4)$$
$$\boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{reg}) = \boldsymbol{\Sigma}_{reg}[\boldsymbol{\varepsilon}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{reg})]$$
$$\boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{sing}) = \boldsymbol{\Sigma}_{sing}[\boldsymbol{\varepsilon}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{sing})]$$

(4) *Update effective stresses*, $\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}$, *at standard (Gauss) sampling points* $\rightarrow \mathbf{x}_i \equiv \mathbf{x}_{G_i}\ (i=1,...,4)$

$$\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}(\mathbf{x}_i) = \tilde{\boldsymbol{\sigma}}^{(e)}_{n}(\mathbf{x}_i) + \Delta\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}(\mathbf{x}_i)$$

$$\Delta\tilde{\boldsymbol{\sigma}}^{(e)}_{n+1}(\mathbf{x}_i) = \begin{cases} \boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}_i) - \boldsymbol{\sigma}^{(e)}_{n}(\mathbf{x}_i) & \mathcal{B}^{(e)} \subset \mathcal{B} \setminus (\mathcal{B}_{wd} \cup \mathcal{B}_{sd})(t_{n+1}) \\ \boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{reg}) - \boldsymbol{\sigma}^{(e)}_{n}(\mathbf{x}^{(e)}_{reg}) & \mathcal{B}^{(e)} \subset \mathcal{B}_{wd}(t_{n+1}) \\ \begin{cases} \xi[\boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{sing}) - \boldsymbol{\sigma}^{(e)}_{n}(\mathbf{x}^{(e)}_{sing})] \\ \quad + (1-\xi)[\boldsymbol{\sigma}^{(e)}_{n+1}(\mathbf{x}^{(e)}_{reg}) - \boldsymbol{\sigma}^{(e)}_{n}(\mathbf{x}^{(e)}_{sing})] \\ \xi = \dfrac{\ell^{(e)}_{\mu}}{h^{(e)}} \end{cases} & \mathcal{B}^{(e)} \subset \mathcal{B}_{sd}(t_{n+1}) \end{cases}$$

B.4. Staggered resolution of the coupled propagation–injection problem

The crack-path-field problem defined in Box 3.1 is stated in terms of $\alpha(\mathbf{x},t)$ (the localized strain-like internal variable, which depends directly on the solution, nodal displacements, $\mathbf{d}_{n+1}$, of the non-linear mechanical problem in Eq. (101)). On the other hand, this mechanical problem also depends on the crack path, $\mu(\mathbf{x},t)$, obtained from the crack path field problem in Box 3.1 (see Fig. 31).

Thus, both problems are coupled, and two sets of discretized, in time and space, equations can be written in terms of the corresponding residues:

$$\begin{aligned} &\boldsymbol{R}_{mech}(\mathbf{d}_{n+1}, \boldsymbol{\mu}_{n+1}) = \mathbf{0}, \quad \text{(a)} \\ &\boldsymbol{R}_{prop}(\mathbf{d}_{n+1}, \boldsymbol{\mu}_{n+1}) = \mathbf{0}, \quad \text{(b)} \end{aligned} \tag{102}$$

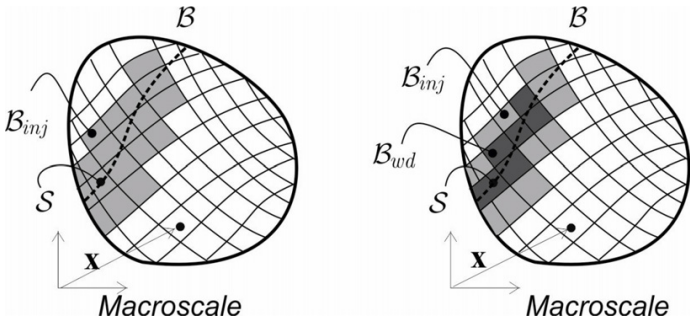


Fig. 31. Coupled propagation (crack-path field) and mechanical (strain injection) problems.

where $\boldsymbol{R}_{prop}$ stands for the residue of the crack propagation problem[14] in Box 3.1 and $\boldsymbol{\mu}_{n+1}$ is the vector of nodal values of the crack-path field $\mu(\mathbf{x}, t_{n+1})$.

As for the strain injection procedure, it can be noticed that, at the initial stages of the non-linear loading process $\mathcal{B}_{inj} = \varnothing$, no discontinuity is injected and, therefore, the problem in Eq. (102) is uncoupled in the sense propagation-problem $\rightarrow$ mechanical-problem. In this case Eq. (102)(a) can be first directly solved for $\mathbf{d}_{n+1}$ and then, Eq. (102)(b) can be solved for $\boldsymbol{\mu}_{n+1}$:

$$\begin{aligned} &\boldsymbol{R}_{mech}(\mathbf{d}_{n+1}) = \mathbf{0} \rightarrow \mathbf{d}_{n+1} && \text{(a)} \\ &\boldsymbol{R}_{prop}(\mathbf{d}_{n+1}, \boldsymbol{\mu}_{n+1}) = \boldsymbol{R}^{*}_{prop}(\boldsymbol{\mu}_{n+1}) = \mathbf{0} \rightarrow \boldsymbol{\mu}_{n+1}. && \text{(b)} \end{aligned} \tag{103}$$

As soon as at least one element belongs to $\mathcal{B}_{inj}$, the problem becomes coupled in both senses.

However, numerical experience shows that this coupling is weak. In fact, the dependence of the mechanical problem on the variable $\boldsymbol{\mu}_t(\mathbf{x})$ is only through the instantaneous position of the discontinuity path, $\mathcal{S}_t$, to determine the indicatrix function φ and the sets $\mathcal{B}_{inj}$ (see Fig. 31), so that the coupling between $\mathbf{d}_{n+1}$ and $\boldsymbol{\mu}_{n+1}$ in the term $\boldsymbol{R}_{mech}(\mathbf{d}_{n+1}, \boldsymbol{\mu}_{n+1})$ can be considered weak. This suggests the use of a staggered simplified procedure, replacing $\boldsymbol{\mu}_{n+1}$ by $\boldsymbol{\mu}_n$ in Eq. (102)(a), leading to the set of uncoupled equations

$$\begin{aligned} &\boldsymbol{R}_{mech}(\mathbf{d}_{n+1}, \boldsymbol{\mu}_n) = \boldsymbol{R}^{*}_{mech}(\mathbf{d}_{n+1}) = \mathbf{0} \rightarrow \mathbf{d}_{n+1} && \text{(a)} \\ &\boldsymbol{R}_{prop}(\mathbf{d}_{n+1}, \boldsymbol{\mu}_{n+1}) = \boldsymbol{R}^{*}_{prop}(\boldsymbol{\mu}_{n+1}) = \mathbf{0} \rightarrow \boldsymbol{\mu}_{n+1}. && \text{(b)} \end{aligned} \tag{104}$$

The staggered resolution of the coupled problem that leads to the uncoupled equations (104) allows envisaging the crack-path-field problem in Box 3.1 as *a post-processing procedure* (typically a double smoothing) of the localizing variable $\alpha_n(\mathbf{x})$ in the mechanical problem. *The crack propagation problem can then be interpreted as a local* (element-wise based) *tracking algorithm that can be straightforwardly, implemented in a finite element code in a non-invasive manner.*

References

[1] R. Hill, Elastic properties of reinforced solids: some theoretical principles, J. Mech. Phys. Solids 11 (5) (1963) 357–372.
[2] P. Germain, The method of virtual power in continuum mechanics. Part 2: Microstructure, SIAM J. Appl. Math. 25 (3) (1973) 556–575.
[3] E. Sanchez-Palencia, Comportements local et macroscopique d'un type de milieux physiques heterogenes, Internat. J. Engrg. Sci. 12 (4) (1974) 331–351.
[4] E. Sanchez-Palencia, Non-Homogeneous Media and Vibration Theory, 1980.
[5] J. Michel, H. Moulinec, et al., Effective properties of composite materials with periodic microstructure: a computational approach, Comput. Methods Appl. Mech. Engrg. 172 (1–4) (1999) 109–143.
[6] S. Nemat-Nasser, Averaging theorems in finite deformation plasticity, Mech. Mater. 31 (8) (1999) 493–523.
[7] K. Terada, M. Hori, et al., Simulation of the multi-scale convergence in computational homogenization approaches, Internat. J. Solids Struct. 37 (16) (2000) 2285–2311.

[14] Which, unlike the mechanical problem, can be proven to be a linear problem.

[8] C. Miehe, Strain-driven homogenization of inelastic microstructures and composites based on an incremental variational formulation, Internat. J. Numer. Methods Engrg. 55 (11) (2002) 1285–1322.
[9] J. Dolbow, M.A. Khaleel, et al. Multiscale mathematics initiative: a roadmap. Pacific Northwest National Laboratory, 2004.
[10] C. Miehe, C. Bayreuther, On multiscale FE analyses of heterogeneous structures: from homogenization to multigrid solvers, Internat. J. Numer. Methods Engrg. 71 (10) (2007) 1135–1180.
[11] F. Vernerey, W.K. Liu, et al., Multi-scale micromorphic theory for hierarchical materials, J. Mech. Phys. Solids 55 (12) (2007) 2603–2651.
[12] Z. Yuan, J. Fish, Towards realization of computational homogenization in practice1, Internat. J. Numer. Methods Engrg. 73 (2008) 361–380.
[13] Y. Efendiev, T.Y. Hou, Multiscale Finite Element Methods: Theory and Applications, Springer Verlag, 2009.
[14] J. Fish, Multiscale Methods: Bridging the Scales in Science and Engineering, Oxford Univ. Pr., 2009.
[15] M.G.D. Geers, V. Kouznetsova, et al., Multi-scale computational homogenization: Trends and challenges, J. Comput. Appl. Math. 234 (7) (2010) 2175–2182.
[16] E.A. de Souza Neto, R.A. Feijóo, Variational foundations of multi-scale constitutive models of solids: small and large strain kinematical formulation. LNCC Research & Development Report 16, 2006.
[17] E.A. de Souza Neto, R.A. Feijoo, On the equivalence between spatial and material volume averaging of stress in large strain multi-scale solid constitutive models, Mech. Mater. 40 (10) (2008) 803–811.
[18] J. Oliver, A.E. Huespe, Theoretical and computational issues in modelling material failure in strong discontinuity scenarios, Comput. Methods Appl. Mech. Engrg. 193 (27–29) (2004) 2987–3014.
[19] I. Gitman, H. Askes, et al., Representative volume: Existence and size determination, Eng. Fract. Mech. 74 (16) (2007) 2518–2534.
[20] Z.P. Bazant, Can multiscale-multiphysics methods predict softening damage and structural failure? Int. J. Multiscale Comput. Eng. 8 (1) (2010) 61–67.
[21] V. Kouznetsova, W. Brekelmans, et al., An approach to micro–macro modeling of heterogeneous materials, Comput. Mech. 27 (1) (2001) 37–48.
[22] V. Kouznetsova, M. Geers, et al., Multi-scale second-order computational homogenization of multi-phase materials: a nested finite element solution strategy, Comput. Methods Appl. Mech. Engrg. 193 (48–51) (2004) 5525–5550.
[23] T. Belytschko, J.H. Song, Coarse-graining of multiscale crack propagation, Internat. J. Numer. Methods Engrg. 81 (5) (2010) 537–563.
[24] V.P. Nguyen, O. Lloberas-Valls, et al., On the existence of representative volumes for softening quasi-brittle materials—a failure zone averaging scheme, Comput. Methods Appl. Mech. Engrg. (2010).
[25] E.W.C. Coenen, V.G. Kouznetsova, et al., A multi-scale approach to bridge microscale damage and macroscale failure: a nested computational homogenization-localization framework, Int. J. Fract. 178 (1–2) (2012) 157–178.
[26] P.J. Sánchez, P.J. Blanco, et al., Failure-oriented multi-scale variational formulation: Micro-structures with nucleation and evolution of softening bands, Comput. Methods Appl. Mech. Engrg. 257 (2013) 221–247.
[27] S. Toro, P.J. Sanchez, et al., A two-scale failure model for heterogeneous materials: numerical implementation based on the finite element method, Internat. J. Numer. Methods Engrg. 97 (5) (2014) 313–351.
[28] J. Oliver, Modelling strong discontinuities in solid mechanics via strain softening constitutive equations. Part 1: Fundamentals, Internat. J. Numer. Methods Engrg. 39 (21) (1996) 3575–3600.
[29] J. Oliver, A.E. Huespe, Continuum approach to material failure in strong discontinuity settings, Comput. Methods Appl. Mech. Engrg. 193 (30–32) (2004) 3195–3220.
[30] J.F. Unger, An FE2-X-1 approach for multiscale localization phenomena, J. Mech. Phys. Solids 61 (4) (2013) 928–948.
[31] J. Oliver, I.F. Dias, et al., Crack-path field and strain-injection techniques in computational modeling of propagating material failure, Comput. Methods Appl. Mech. Engrg. 274 (2014) 289–348.
[32] J. Oliver, A.E. Huespe, et al., Stability and robustness issues in numerical modeling of material failure with the strong discontinuity approach, Comput. Methods Appl. Mech. Engrg. 195 (2006) 7093–7114.
[33] J. Oliver, Continuum modelling of strong discontinuities in solid mechanics using damage models, Comput. Mech. 17 (1–2) (1995) 49–61.
[34] R. Faria, J. Oliver, et al., A strain-based plastic viscous-damage model for massive concrete structures, Internat. J. Solids Struct. 35 (14) (1998) 1533–1558.
[35] J. Oliver, A.E. Huespe, et al., From continuum mechanics to fracture mechanics: The strong discontinuity approach, Eng. Fract. Mech. 69 (2) (2002) 113–136.
[36] J. Oliver, On the discrete constitutive models induced by strong discontinuity kinematics and continuum constitutive equations, Internat. J. Solids Struct. 37 (48–50) (2000) 7207–7229.
[37] O.C. Zienkiewicz, R.L. Taylor, The Finite Element Method. Volumes 1 and 2, McGraw-Hill, 1994.
[38] T. Belytschko, S. Loehnert, et al., Multiscale aggregating discontinuities: a method for circumventing loss of material stability, Internat. J. Numer. Methods Engrg. 73 (6) (2008) 869–894.
[39] I. Stakgold, Green's Functions and Boundary Value Problems, Wiley, New York, 1998.
[40] J.C. Michel, P. Suquet, Nonuniform transformation field analysis, Internat. J. Solids Struct. 40 (25) (2003) 6937–6955.
[41] J.C. Michel, P. Suquet, Computational analysis of nonlinear composite structures using the nonuniform transformation field analysis, Comput. Methods Appl. Mech. Engrg. 193 (48–51) (2004) 5477–5502.
[42] K. Willam, N. Sobh, Bifurcation analysis of tangential material operators, in: G.N. Pande, J. Middleton (Eds.), Transient/Dynamic Analysis and Constitutive Laws for Engineering Materials. Vol. 2, Martinus-Nijhoff Publishers, 1987, pp. 1–13.
[43] N.S. Ottosen, K. Runesson, Properties of discontinuous bifurcation solutions in elasto-plasticity, Internat. J. Solids Struct. 27 (4) (1991) 401–421.
[44] J. Oliver, A.E. Huespe, et al., On the numerical resolution of the discontinuous material bifurcation problem, Internat. J. Numer. Methods Engrg. 83 (6) (2010) 786–804.
[45] P.J. Blanco, P.J. Sánchez, et al., Variational foundations and generalized unified theory of RVE-based multiscale models, Arch. Comput. Methods Eng. (2014) 1–63.

[46] N.B. Nooru-Mohamed, Mixed-mode fracture of concrete: an experimental approach y, Delft University of Technolog, 1992.
[47] E. Roubin, A. Vallade, et al., Multi-scale failure of heterogeneous materials: A double kinematics enhancement for Embedded Finite Element Method, Internat. J. Solids Struct. 52 (0) (2015) 180–196.
[48] C. Oskay, J. Fish, Eigendeformation-based reduced order homogenization for failure analysis of heterogeneous materials, Comput. Methods Appl. Mech. Engrg. 196 (7) (2007) 1216–1243.
[49] J.A. Hernández, J. Oliver, A.E. Huespe, M.A. Caicedo, J.C. Cante, High-performance model reduction techniques in computational multiscale homogenization, Comput. Methods Appl. Mech. Engrg. 276 (2014) 149–189.
[50] J.C. Simo, M.S. Rifai, A class of mixed assumed strain methods and the method of incompatible modes, Internat. J. Numer. Methods Engrg. 29 (8) (1990) 1595–1638.
[51] J. Simo, J. Oliver, A new approach to the analysis and simulation of strong discontinuities. Fracture and Damage in Quasi-brittle Structures, $ &$ FN Spon, 1994.
[52] J. Oliver, A.E. Huespe, et al., Strain localization, strong discontinuities and material fracture: Matches and mismatches, Comput. Methods Appl. Mech. Engrg. 241–244 (2012) 323–336.
[53] J. Oliver, Modelling strong discontinuities in solid mechanics via strain softening constitutive equations. Part 2. Numerical simulation, Internat. J. Numer. Methods Engrg. 39 (21) (1996) 3601–3623.
[54] J.E. Dolbow, A. Devan, Enrichment of enhanced assumed strain approximations for representing strong discontinuities: addressing volumetric incompressibility and the discontinuous patch test, Internat. J. Numer. Methods Engrg. 59 (1) (2004) 47–67.

B.3 Paper #3

Title: Reduced Order Modeling strategies for Computational Multiscale Fracture.
Authors:

- **J. Oliver**: Professor of Continuum Mechanics and Structural analysis at the Escola Tecnica Superior d'Enginyers de Camins, Canals i Ports (Civil Engineering School) of the Universitat Politècnica de Catalunya (Technical University of Catalonia BarcelonaTech). Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **A. E. Huespe**: Professor of Mechanics at the Faculty of Chemical Engineering, Dept. of Materials, National University of Litoral, Santa Fe, Argentina. Independent researcher of Conicet at CIMEC (Centro de Investigaciones en Mecánica Computacional), National University of Litoral (UNL).
- **J. A. Hernández**: Assistant Professor of Structural Engineering and Strength of Materials at the School of Industrial and Aeronautic Engineering of Terrassa, of the Technical University of Catalonia. Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **E. Roubin**: Maître de conférence at the 3SR (Sols, Solides, Structures et Risques) and the IUT DGGC in Grenoble.

Journal of Computer Methods in Applied Mechanics and Engineering
Editors: Thomas J.R. Hughes, J. Tinsley Oden, Manolis Papadrakakis
ISSN: 0045-7825
Elsevier Editors

Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 313 (2017) 560–595
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Reduced order modeling strategies for computational multiscale fracture

J. Oliver[a,b,*], M. Caicedo[a], A.E. Huespe[a,c], J.A. Hernández[b,d], E. Roubin[e]

[a] *E.T.S d'Enginyers de Camins, Canals i Ports, Technical University of Catalonia (BarcelonaTech), Campus Nord UPC, Mòdul C-1, c/ Jordi Girona 1-3, 08034, Barcelona, Spain*
[b] *CIMNE – Centre Internacional de Metodes Numerics en Enginyeria, Campus Nord UPC, Mòdul C-1, c/ Jordi Girona 1-3, 08034, Barcelona, Spain*
[c] *CIMEC-UNL-CONICET, Güemes 3450, Santa Fe, Argentina*
[d] *Escola Superior d'Enginyeries Industrial, Aerospacial i Audiovisual de Terrassa, Technical University of Catalonia (BarcelonaTech), C/Colom 11, 08222, Terrassa, Spain*
[e] *Laboratoire 3SR, Université Grenoble Alpes (UGA), CNRS Domaine Universitaire, 38000 Grenoble Cedex, France*

Received 6 May 2016; received in revised form 23 August 2016; accepted 27 September 2016
Available online 14 October 2016

Highlights

- The RVE homogenization is stated in terms of a saddle-point problem.
- RVE unknowns are unconventionally rephrased in terms of the fluctuation strains.
- Separate Reduced Order Models are used in the RVE elastic and inelastic domains.
- HPROM combines Reduced Order Modeling and Reduced Optimal Quadrature techniques.
- Results exhibit outstanding speedups for multi-scale fracture problems.

Abstract

The paper proposes some new computational strategies for affordably solving multiscale fracture problems through a FE^2 approach. To take into account the mechanical effects induced by fracture at the microstructure level the Representative Volume Element (RVE), assumed constituted by an elastic matrix and inclusions, is endowed with a large set of cohesive softening bands providing a good representation of the possible microstructure crack paths. The RVE response is then homogenized in accordance with a model previously developed by the authors and upscaled to the macro-scale level as a continuum stress–strain constitutive equation, which is then used in a conventional framework of a finite element modeling of propagating fracture.

For reduced order modeling (ROM) purposes, the RVE boundary value problem is first formulated in displacement fluctuations and used, via the Proper Orthogonal Decomposition (POD), to find a low-dimension space for solving the reduced problem. A *domain separation strategy* is proposed as a first technique for model order reduction: unconventionally, the low-dimension space is spanned by a basis in terms of fluctuating strains, as primitive kinematic variables, instead of the conventional formulation in terms of displacement fluctuations. The RVE spatial domain is then decomposed into a *regular domain* (made of the matrix and the inclusions) and *a singular domain* (constituted by cohesive bands), the required RVE boundary conditions are rephrased in terms

* Corresponding author at: CIMNE – Centre Internacional de Metodes Numerics en Enginyeria, Campus Nord UPC, Mòdul C-1, c/ Jordi Girona 1-3, 08034, Barcelona, Spain.
E-mail address: xavier.oliver@upc.edu (J. Oliver).

http://dx.doi.org/10.1016/j.cma.2016.09.039

of strains and imposed via Lagrange multipliers in the corresponding variational problem. Specific low-dimensional strain basis is then derived, independently for each domain, via the POD of the corresponding strain snapshots.

Next step consists of developing a hyper-reduced model (HPROM). It is based on a second proposed technique, the *Reduced Optimal Quadrature* (ROQ) which, again unconventionally, is determined through optimization of the numerical integration of the primitive saddle-point problem arising from the RVE problem, rather than its derived variational equations, and substitutes the conventional Gauss quadrature. The ROQ utilizes a very reduced number of, optimally placed, sampling points, the corresponding weights and placements being evaluated through a greedy algorithm. The resulting low-dimensional and reduced-quadrature variational problem translates into very relevant savings on the computational cost and high computational speed-ups.

Particular attention is additionally given to numerical tests and performance evaluations of the new hyper-reduced methodology, by "a-priori" and "a-posteriori" error assessments. Moreover, for the purposes of validation of the present techniques, a real structural problem exhibiting propagating fracture at two-scales is modeled on the basis of the strain injection-based multiscale approach previously developed by the authors. The performance of the proposed strategy, in terms of speed-up vs. error, is deeply analyzed and reported.

Keywords: Reduced Order Modeling (ROM); Hyper-Reduced Order Modeling (HPROM); Multiscale fracture models; Computational homogenization; Reduced Optimal Quadrature (ROQ)

Acronyms

BVP: Boundary Value Problem
CSDA: Continuum-Strong Discontinuity Approach
EFEM: Embedded Finite Element Methodology
FE^2: Two-scale (macro and micro or meso) model, where both scales of analysis are represented by finite element approaches
HF: High Fidelity model
HPROM: HyPer-Reduced Order Modeling
OQN: Optimal Quadrature Number
POD: Proper Orthogonal Decomposition
ROM: Reduced Order Modeling
ROQ: Reduced Optimal Quadrature
RVE: Representative Volume Element
SVD: Singular Value Decomposition
VBVP: Variational Boundary Value Problem

Symbols

Number of dimensions of vectors or vectorial spaces

n_σ: Number of components of the stress and strain (or micro-stress and micro-strain) vectors described with Voigt's notation, typically: for strains and stresses in plane states: $n_\sigma = 4$, and $\boldsymbol{\varepsilon} = [\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, \varepsilon_{xy}]^T$ or $\boldsymbol{\sigma} = [\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}]^T$.

n_ε: Dimension of the reduced micro-strain fluctuation space (number of basis of $\boldsymbol{\Psi}$ spanning the full space of reduced micro-strains).

$n^I_{\varepsilon,reg}$: Dimension of the reduced micro-strain fluctuation space $\boldsymbol{\Psi}^I_{reg}$ associated with the inelastic snapshots and Gauss points in the regular domain ($\mathcal{B}_{\mu,\text{reg}}$).

$n^I_{\varepsilon,coh}$: Dimension of the reduced micro-strain fluctuation space $\boldsymbol{\Psi}^I_{coh}$ associated with the inelastic snapshots and Gauss points in the domain of cohesive bands ($\mathcal{B}_{\mu,\text{coh}}$).

n_φ: Dimension of the reduced micro internal energy space (number of basis of $\boldsymbol{\Phi}$ spanning the full space of reduced micro-internal energy φ_μ).

Number of quadrature points

N_g: Number of Gauss points associated with the original HF finite element mesh.
$N_{g,reg}$: Number of Gauss points of the HF model in the micro-cell regular domain ($\mathcal{B}_{\mu,\text{reg}}$).
$N_{g,coh}$: Number of Gauss points of the HF model in the micro-cell domain of the cohesive bands ($\mathcal{B}_{\mu,\text{coh}}$).
N_r: Number of quadrature points adopted for the ROQ scheme.

Other symbols

p_{snp}: Number of micro-strain snapshots taken during the off-line sampling process constituting the micro-strain snapshot matrix [$\boldsymbol{X}$].
q_{snp}: Number of micro-energy snapshots taken during the off-line sampling process constituting the micro-energy snapshot matrix [$\boldsymbol{X}_\phi$].

1. Introduction

Multiscale modeling is expected to become a key approach to enable the next wave of design paradigms for engineering materials and structures. Indeed, it appears as an excellent potential setting to account for the physical links between the different lower scale components, within the material (grains, particles, defects, inclusions, etc.), and the overall large scale properties. However, this modeling concept, coined by the scientific community some decades ago, is not yet part of the routinely engineering analysis and design methods. Quoting from a report by a group of experts to the US National Science Foundation [1]

> "... In recent years, a large and growing body of literature in physics, chemistry, biology, and engineering has focused on various methods to fit together simulation models of two or more scales, and this has led to the development of various multi-level modeling approaches..... To date, however, progress on multiscale modeling has been agonizingly slow. Only a series of major breakthroughs will help us establish a general mathematical and computational framework for handling multiscale events and reveal to us the commonalities and limitations of existing methods ...."

Certainly, and focusing on the specific case of computational homogenization-based multiscale techniques, they can hardly be applied beyond some simple and academic purposes. Reasons for this arise from the multiplicative character of the algorithmic complexity,[1] and the corresponding computational cost, for hierarchical micro/macro computations in multiscale analyses. A representative example of this challenge is the computational two-scale modeling of structural fracturing materials, where a detailed geometrical description of the material morphology and failure propagation is necessary at the low scale, in conjunction with a very fine representation of the crack distribution and propagation at the upper scale.

On the other hand, model order reduction (data compression) techniques have become an intensive research field in the computational mechanics community, because of the increasing interest on computational modeling of complex phenomena in large scale multiphysics problems.

This work deals with a combination of both subjects by focusing on the reduced order modeling of computational multiscale modeling of fracturing materials. Several works on fracture multiscale modeling using the Representative Volume Element (RVE) notion have been published in the recent years, a (not complete) list is [2–7]. In this work, the multiscale framework for numerical modeling of structural fracture in heterogeneous quasi-brittle materials, described in [8], which overcomes some of the classically encountered difficulties for this type of approaches, is utilized as start point. In short, it can be defined as a FE^2 approach (see [9]), where the Continuum-Strong Discontinuity Approach (CSDA) is used at the macroscale. Only a brief overview of this technique is presented in Appendix A.

Contrarily, emphasis is given in this work to the RVE modeling, i.e. the finite element approach at the micro (or meso) scale of the problem. The objective here is to develop a Hyper-Reduced Order Model (HPROM) of the

[1] Here understood as the number of numerical operations to be performed in the considered algorithm.

RVE finite element model. The most strong and motivating argument for developing a reduced order model of a non-linear RVE problem, within a FE^2 framework, stems from the following reason: during the online computation of a multiscale problem, the RVE problem has to be solved a very large number of times.[2] This relevant feature must be taken into account for selecting an adequate HPROM. Moreover, since the parametric space of the RVE driving forces is small,[3] techniques based on off-line micro-cell sampling, such as the one proposed in this work, seem to be feasible.

Development of reduced models for non-homogeneous materials has been tackled in numerous previous contributions, such as [10], where the proposed reduction techniques are based on Fourier's transforms, or [11], where a reduced model is applied to homogenization analysis of hyperelastic solids subjected to finite strains. Also, the work in [12] developed a hyper-reduced model of a monoscale analysis considering non-linear material behavior. However, the existing literature barely considers reduced order modeling of non-smooth problems, as is the case of fracture, where discontinuous displacements occur. The multiscale case, when fracture is also considered at the smaller scale of the problem, makes the task even much heavier. Indeed, only very few contributions have been presented in the literature about this topic, see for example: [13], which follows an eigendeformation-based methodology, or [14,15] that resort to global–local approaches.

In this work, a number of techniques are combined to optimize the HPROM performance of FE^2 multiscale modeling algorithms for multiscale propagating fracture. They are:

1. A *domain separation strategy*. The RVE is split into the *regular domain* (made of the elastic matrix and possible inclusions) and the *singular domain* (the cohesive bands exhibiting a softening cohesive behavior), which are designed to provide a good enough representation of the microscopic fracture and of its effects on the homogenized material behavior [8]. The distinct constitutive behavior of both domains suggests a specific ROM strategy for each of them in order to obtain a reduction strategy keeping available the information on the mechanical variables in every specific sub-domain. Therefore, selection of the ROM low-dimensional projection space is made independently for each of these domains.
2. In combination with the previous strategy, the ROM boundary problem for the RVE is formulated in an unconventional manner i.e.: *in terms of the strain fluctuations rather than in terms of the conventional displacement fluctuations*. The reduced strain fluctuation space is spanned by basis functions satisfying, by construction, the kinematical compatibility conditions, this guaranteeing that, after reduction, any solution in the strain fluctuation space also satisfies the kinematical compatibility.
3. *A* specific *Reduced Optimal Quadrature* (ROQ) is used as a key technique to obtain relevant computational cost reduction from the ROM. This technique, consisting of replacing the standard Gauss integration rule over the whole finite element mesh by a different optimal quadrature involving much less sampling points, has been proposed in recent works [16,17] as an ingredient of HyPer-Reduced Order Modeling (HPROM) strategies. In these works, the reduced numerical integration technique is applied to the variational equations of the problem (i.e. internal forces, involving n-dimensional vector entities) whereas, in the herein proposed approach, a similar reduced integration technique is applied, again unconventionally, *to the primitive problem,* i.e: the functional (a scalar entity) in the micro-scale saddle-point problem *that supplies the RVE variational equations*. In the present proposal, this functional turns out to be the stored energy (free energy) at the RVE, which, being a scalar entity, is less demanding as for the integration rule is concerned. This is expected to translate into very relevant savings in both the numerical ROQ derivation procedure and the resulting computational cost.

The above techniques are combined to provide the proposed HPROM strategy for the RVE, using a standard two-stage (off-line/on-line) strategy:

(I) OFF-LINE STAGE
- A sampling program of the high-fidelity RVE finite element model is performed with the objective of gathering a set of solutions (snapshots) of micro-strain fluctuations, at the regular and singular RVE domains, as well as the corresponding snapshots of the energy functional in the RVE variational principle.

[2] Typically, as many times as sampling points are considered at the upper-scale times the number of iterations involved in the solution of the non-linear problem.

[3] As many of relevant strain dimensions in the RVE problem: 3 in 2-D cases and 6 in 3D cases.

- Through POD techniques, applied to the snapshots sets, the reduced bases (modes) for the regular and singular *micro-strain fluctuations and energies* are computed.
- The ROQ rule, stated in terms of the placement of the sampling points and the corresponding integration weights at the regular and singular domains, is determined on the basis of an optimal integration of the considered free energy modes.

(II) ON-LINE STAGE

- The information collected in stage (I) (micro-strain fluctuation bases and reduced integration rule) is used for the on-line solution, in stage (II) of a reduced problem. The micro-strain fluctuation projections onto the reduced strain bases are now the unknowns, and the new HPROM problem is set in terms of the new ROM saddle-point functional integrated with the Reduced Optimal Quadrature (ROQ).

In the remaining of this work, this strategy for HPROM of the RVE problem is detailed. After presenting, in Section 2, a brief summary of the multiscale fracture model already presented by the authors elsewhere in [8] (and additionally detailed in Appendix A), in Section 3 the BVP of the RVE is shown. Then, in Section 3.2, the ROM procedure is developed by approaching the micro-strain fluctuation fields through the obtained low-dimensional spaces. Finally, a technique to derive the Reduced Optimal Quadrature (ROQ) rule is proposed in Section 3.3.

It is emphasized that the ROQ rule described in this work is not only restricted to multiscale approaches, but it can rather easily be extended and applied to more general mechanical problems. In order to sketch a possible generalization of this technique, in Appendix B a summarized description of the ROQ procedure for a fairly general family of non-linear solid mechanics problems is presented.

An important issue in developing an efficient HPROM method is the design of the off-line testing program. It should be devised to obtain a representative set of RVE solutions, appropriately spanning the parametric space of the RVE actions (the macroscopic strains). In Section 4.1 a devised sampling program for the high-fidelity RVE finite element model is presented.

In Section 5, the a-priori and consistency errors of the HPROM model are studied. Additional assessment of the model is then done for realistic structural problems exhibiting propagating fracture at the macro-scale. The work ends with some concluding remarks in Section 6.

2. Multiscale mechanical model for propagating fracture

Let us consider the fracture problem of a body $\mathcal{B}$ composed of a heterogeneous material, which is modeled using the multiscale technique developed in [8]. Although, here it is not paid attention to this multiscale technique, but to its computational cost reduction, a summary of the main ingredients of the former has to be presented in order to understand the proposed cost reduction procedure. Therefore, a sketch of the multiscale model is presented in Fig. 1 and it is outlined in the following items.

- Two scales of analysis, macro and micro (or meso) scales, are considered within a semi-concurrent FE^2 approach (see [9]). Every point at the macroscale is associated with an RVE (also termed here micro-cell) which, through a convenient homogenization procedure, provides the effective constitutive response of the material at the macroscale. The micro-cell model accounts for the material morphology at the lower scale.
- The failure process at the macroscale is characterized by cracks modeled with the Continuum-Strong Discontinuity Approach (CSDA). This approach, with applications to monoscale fracture analysis, has been widely developed by the authors in the past, see [18–21]. Briefly, this technique assumes a strong discontinuity kinematics (i.e including displacement jumps in the crack paths), which is regularized by introducing very thin bands, denoted $\mathcal{B}_{loc}$, see Fig. 1(a), where strain localization can take place. The actual displacement discontinuities are regularized inside these bands as highly localized strains (Fig. 1(b)). Cohesive forces in $\mathcal{B}_{loc}$ are calculated with the stresses $\boldsymbol{\sigma}$, evaluated inside these bands *in terms of the localized strains* rather than in terms of the *displacement jumps* used in the more conventional *discrete approaches* governed by traction-separation laws. The fact that in the CSDA the constitutive model is a continuum (stress–strain) one, makes this approach more compatible with conventional computational homogenization procedures. Therefore, $\boldsymbol{\sigma}$ is evaluated everywhere in $\mathcal{B}$ through a standard homogenization procedure involving a micro-cell which takes into account the material micro-heterogeneity.

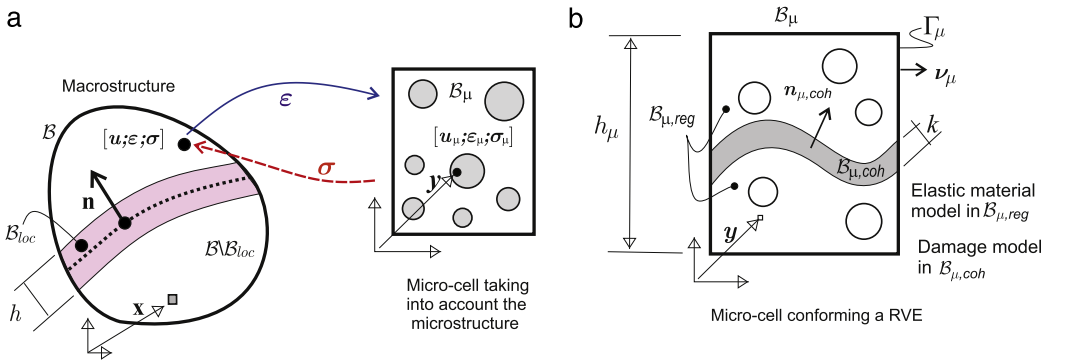


Fig. 1. Outline of the multiscale model for propagating fracture: (a) macro and micro-scales; (b) micro-cell model accounting for material failure.

- At the micro-cell (or RVE), material failure is modeled, using again the CSDA, through a large number of predefined finite thickness micro-bands[4] endowed with a very simple stress–strain continuum damage constitutive law (see [22,23]). Activation of a given failure mechanism at the micro-cell, i.e. activation of the damage process in a set of micro-bands, triggers a strain localization effect at the macroscale which finally leads to the formation of a strong discontinuity (macro-crack). The strong discontinuity mechanism at the macroscale induced by strain localization, as well as the propagating crack phenomenon represented by this mechanism, is properly captured with finite elements enriched with a specific kinematics.

Since, as commented above, the objective of this work is to provide a reduced model of the RVE problem, emphasis is only given to the RVE modeling. Additional details of the multiscale technique can be found in Appendix A and the reference work [8].

2.1. Notation adopted in this work

The body domain at the *macroscale* is denoted $\mathcal{B}$ ($\mathcal{B} \subset \mathbb{R}^{n_{dim}}$ and n_{dim} is the space dimension), while $\mathcal{B}_{loc}$ indicates the subdomain where strain localization happens, see Fig. 1(a). It is assumed that $\mathcal{B}_{loc}$ is a band of finite thickness h, with a mean surface having a unit normal vector $\boldsymbol{n}$. Points at macroscale are identified by $\boldsymbol{x}$ and displacements, strains and stresses by $\boldsymbol{u}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$, respectively.

All geometrical entities and variables at the *microscale* are identified with the subindex $(\cdot)_\mu$ (so, $\mathcal{B}_\mu$ is the micro-cell domain at the microscale). In correspondence with the macro-scale fields, displacements, strains and stresses are described at the microscale as $\boldsymbol{u}_\mu$, $\boldsymbol{\varepsilon}_\mu$ and $\boldsymbol{\sigma}_\mu$, respectively. Points at the microscale are denoted $\boldsymbol{y}$.

Tensor notation vs. Voigt's notation

In this work both notations are in accordance with the following criteria:

(a) conventional tensor notation, is used in the original, not-reduced, formulations of the micro-scale variational problem. Being the space of the symmetric second order tensors: $\mathbb{S}^{n_{dim}\times n_{dim}}$, then, typically stresses and strains, are identified as: $\boldsymbol{\varepsilon}, \boldsymbol{\sigma} \in \mathbb{S}^{n_{dim}\times n_{dim}}$. Double contraction of second order tensors is written as: $(\boldsymbol{\sigma}_\mu : \tilde{\boldsymbol{\varepsilon}}_\mu)$.

(b) Voigt's notation, utilized in the reduced model formulations and their implementations. In this notation, strains and stresses are described as column vectors: ($\boldsymbol{\varepsilon}, \boldsymbol{\sigma} \in \mathbb{R}^{n_\sigma}$, $n_\sigma = 4$ in plane states of stress or strain). Fourth order tensors are described by matrices in $\mathbb{R}^{n_\sigma \times n_\sigma}$. Double contraction of second order tensors is written as a vector product in matrix form, i.e. $(\boldsymbol{\sigma}_\mu)^T \tilde{\boldsymbol{\varepsilon}}_\mu$.

(c) The same symbol is used to identify identical entities in both notations. The proper context of every equation determines which specific notation is used. For example, by writing $\boldsymbol{\varepsilon} \in \mathbb{S}^{n_{dim}\times n_{dim}}$ it is understood that the strain, and the corresponding equation, are expressed in tensor notation, while writing $\boldsymbol{\varepsilon} \in \mathbb{R}^{n_\sigma}$, refers that the same entity is expressed in Voigt's notation.

[4] Large enough to provide a good representation of the microscopic fracture variables relevant for the macroscopic homogenized constitutive behavior.

2.2. Micro-cell model accounting for material failure: Variational boundary value problem (VBVP)

Let us consider the micro-cell $\mathcal{B}_\mu$ sketched in Fig. 1(b), of size h_μ, and boundary Γ_μ with unit normal vector denoted by $\boldsymbol{n}_\mu$. This micro-cell is large enough to satisfy the conditions of being a RVE.[5] The displacement field at the microscale, denoted $\boldsymbol{u}_\mu$, can be taken as the addition of three terms:

$$\boldsymbol{u}_\mu(\boldsymbol{y}) = \boldsymbol{u} + \boldsymbol{\varepsilon} \cdot \boldsymbol{y} + \tilde{\boldsymbol{u}}_\mu(\boldsymbol{y}). \tag{1}$$

The terms $\boldsymbol{u}$ and $\boldsymbol{\varepsilon}$ are the macro-scale displacement and strain, respectively, down-scaled to the RVE. The displacement fluctuation field, $\tilde{\boldsymbol{u}}_\mu(\boldsymbol{y})$ is the relevant term in the governing equations of the RVE problem. The micro-strain $\boldsymbol{\varepsilon}_\mu$ in the RVE is given by the addition of two terms[6]:

$$\boldsymbol{\varepsilon}_\mu(\boldsymbol{y}) = \boldsymbol{\varepsilon} + \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}) = \boldsymbol{\varepsilon} + \nabla^s \tilde{\boldsymbol{u}}_\mu(\boldsymbol{y}) \quad \forall \boldsymbol{y} \in \mathcal{B}_\mu, \tag{2}$$

where the macro-strain $\boldsymbol{\varepsilon}$ is homogeneously distributed in $\mathcal{B}_\mu$, while the compatible micro-strain fluctuation, $\tilde{\boldsymbol{\varepsilon}}_\mu$, is the symmetric gradient of the displacement fluctuation field satisfying:

$$\int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu \, d\mathcal{B} = \int_{\mathcal{B}_\mu} \nabla^s \tilde{\boldsymbol{u}}_\mu \, d\mathcal{B} = \int_{\Gamma_\mu} \tilde{\boldsymbol{u}}_\mu \otimes^s \boldsymbol{n}_\mu \, d\Gamma = \boldsymbol{0}. \tag{3}$$

The last identity in (3) is obtained by applying the Green's theorem to transform the volume integral, in $\mathcal{B}_\mu$, into the surface integral, on Γ_μ.

From the kinematical constraint (3), we define the space $\mathcal{U}^u_\mu$ of admissible displacement fluctuations[7]:

$$\mathcal{U}^u_\mu := \left\{ \tilde{\boldsymbol{u}}_\mu \quad | \quad \int_{\Gamma_\mu} \tilde{\boldsymbol{u}}_\mu \otimes^s \boldsymbol{n}_\mu \, d\Gamma = \boldsymbol{0} \right\}. \tag{4}$$

Eqs. (1)–(4), plus the constitutive and equilibrium equations that should be satisfied at the RVE (described in Appendix A but not repeated here, see Eqs. (76)–(79)), define a conventional Variational Boundary Value Problem (VBVP) at the RVE where the actions (driving force) are the macro-strains, $\boldsymbol{\varepsilon}$. A summary of the governing equations, identified as PROBLEM I, is given next.

PROBLEM I *(RVE variational displacement fluctuation-based problem)*
Given the macroscale strain, $\boldsymbol{\varepsilon}$, the space of displacement fluctuations, $\mathcal{U}^u_\mu$, and the space of admissible displacement fluctuations $\mathcal{V}^u_\mu$:

$$\mathcal{U}^u_\mu = \mathcal{V}^u_\mu := \left\{ \tilde{\boldsymbol{u}}_\mu \quad | \int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu \, d\mathcal{B} = \int_{\Gamma_\mu} \tilde{\boldsymbol{u}}_\mu \otimes^s \boldsymbol{n}_\mu \, d\Gamma = \boldsymbol{0} \right\}; \tag{5}$$

find $\tilde{\boldsymbol{u}}_\mu \in \mathcal{U}^u_\mu$ such that:

$$\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}_\mu, d_\mu) : \nabla^s \hat{\boldsymbol{u}}_\mu \; d\mathcal{B} = 0; \qquad \forall \hat{\boldsymbol{u}}_\mu \in \mathcal{V}^u_\mu; \tag{6}$$

$$\dot{d}_\mu(\boldsymbol{y}, \boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu) \tag{7}$$

Eq. (6) is the variational equilibrium equation in terms of the micro-stresses $\boldsymbol{\sigma}_\mu$. Implicitly in the same expression have been considered the constitutive equation relating $\boldsymbol{\sigma}_\mu$ in terms of the micro-strains, $\boldsymbol{\varepsilon}_\mu$, and internal damage variable, d_μ. While (7) describes the pointwise evolution law for this internal variable.

[5] Micro-cell or RVE terms will be indistinctly used in this work.

[6] Without dropping out the emphasis given to the fact that variables are typically load history-dependent, from now on references to time will be omitted unless it is necessary.

[7] The HPROM procedure presented in Section 3 is not restricted to deal with RVE models having minimum kinematical constraints defined by Eq. (3) and the vectorial space (4) associated to it. In fact, and due to the linear properties of the SVD algorithms, the HPROM can be applied, with no further modification, to the cases of homogeneous boundary conditions defined in every subdomain of the partitioned RVE ($\mathcal{B}_{\mu,\text{reg}}$ and $\mathcal{B}_{\mu,\text{coh}}$). Therefore, the proposed techniques can be used in RVE models having periodic boundary conditions.

2.3. Alternative descriptions of the micro-cell VBVP

2.3.1. Formulation in micro-strain fluctuations

As it will be shown later, for hyper-reduced order modeling purposes it may be convenient to rephrase PROBLEM I taking as the primary unknowns the micro-strain fluctuations, $\tilde{\boldsymbol{\varepsilon}}_\mu$, instead of the micro displacement fluctuations. Let us consider the space $\mathcal{E}_\mu$ of micro-strain tensorial functions which are kinematically compatible. Necessary and sufficient conditions for this to happen are the classical compatibility equations (see [24]):

$$\mathcal{E}_\mu := \{\boldsymbol{\zeta} \in \mathbb{S}^{n_{dim}\times n_{dim}} \quad | \quad e_{mjq}e_{nir}\boldsymbol{\zeta}_{ij,qr} = 0\} \tag{8}$$

where e is the permutation tensor, while Einstein summation convention applies to the equation constraining $\boldsymbol{\zeta}$. Notice that equations constraining $\boldsymbol{\zeta}$ are linear and homogeneous, which is a crucial fact to be exploited later on.

We realize that Eq. (3) allows rephrasing Eqs. (5) and (6) as the following problem:

PROBLEM I-R *(RVE rephrased variational strain fluctuation-based problem)*
Given the macroscale strains, $\boldsymbol{\varepsilon}$, and the spaces of kinematically compatible strain fluctuations, $\mathcal{U}_\mu^{\tilde{\epsilon}}$, and admissible strain fluctuations, $\mathcal{V}_\mu^{\tilde{\epsilon}}$:

$$\mathcal{U}_\mu^{\tilde{\epsilon}} = \mathcal{V}_\mu^{\tilde{\epsilon}} := \left\{ \tilde{\boldsymbol{\varepsilon}}_\mu \quad | \quad \int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu \, d\mathcal{B} = \mathbf{0} \quad \text{and} \quad \tilde{\boldsymbol{\varepsilon}}_\mu \in \mathcal{E}_\mu \right\}; \tag{9}$$

find $\tilde{\boldsymbol{\varepsilon}}_\mu \in \mathcal{U}_\mu^{\tilde{\epsilon}}$ such that:

$$\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}_\mu, d_\mu) : \widehat{\boldsymbol{\varepsilon}}_\mu \, d\mathcal{B} = 0; \qquad \forall \widehat{\boldsymbol{\varepsilon}}_\mu \in \mathcal{V}_\mu^{\tilde{\epsilon}}; \tag{10}$$

$$\dot{d}_\mu(\mathbf{y}, \boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu) \tag{11}$$

PROBLEM I-R is completely equivalent to PROBLEM-I as for the micro-stresses, $\boldsymbol{\sigma}_\mu$, and the homogenized macroscopic results, but now they are given in terms of the micro-strain fluctuations $\boldsymbol{\varepsilon}_\mu$. The original displacement fluctuations, $\tilde{\boldsymbol{u}}_\mu$, can be recovered, if necessary for the deformed RVE visualization purposes, through an additional spatial integration of the equation $\tilde{\boldsymbol{\varepsilon}}_\mu = \nabla^s \tilde{\boldsymbol{u}}_\mu$. A variational procedure for doing this is provided in Appendix C.

2.3.2. Formulation as a saddle point problem

Let us consider the material free energy, φ_μ, for the isotropic damage model in $\mathcal{B}_{\mu,\text{coh}}$, which can be expressed as (see [8]):

$$\varphi_\mu(\boldsymbol{\varepsilon}_\mu, d_\mu(\boldsymbol{\varepsilon}_\mu^\tau)) = \frac{1}{2}(1 - d_\mu(\boldsymbol{\varepsilon}_\mu^\tau))\boldsymbol{\varepsilon}_\mu \cdot \mathbb{C} \cdot \boldsymbol{\varepsilon}_\mu \tag{12}$$

where the micro-internal variable (isotropic damage), d_μ, has been identified as a functional of the micro-strain history, denoted $\boldsymbol{\varepsilon}_\mu^\tau(t)$. The supra-index τ refers to all the registered historical values until the current time t: $\tau \in [0, t]$. Typically, the internal variable d_μ is specified through evolution equations, generally expressed as:

$$\dot{d}_\mu(\mathbf{y}, \boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu). \tag{13}$$

Additionally, and following the standard framework in constitutive modeling, and in consonance with Eq. (12), the micro-stress, $\boldsymbol{\sigma}_\mu$, can be written as:

$$\boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}_\mu) = \partial\varphi_\mu(\boldsymbol{\varepsilon}_\mu, d_\mu)/\partial\boldsymbol{\varepsilon}_\mu, \tag{14}$$

where, it is implicitly assumed that the variable d_μ, in the arguments of φ_μ, satisfies the constitutive relation (13). Monolithic schemes to compute the stress $\boldsymbol{\sigma}_\mu$, in Eq. (14), jointly with Eq. (13), are conventional techniques in non-linear solid mechanics (see for example [25]).

For subsequent developments, and accounting for Eq. (2), it is convenient to introduce a *parameterized expression of the free energy*, identified as:

$$\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu) = \varphi_\mu(\underbrace{\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu}_{\boldsymbol{\varepsilon}_\mu}, d_\mu) = \frac{1}{2}(1-d_\mu)(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu)\mathbb{C}\cdot(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu) \tag{15}$$

where the sub-index $[\boldsymbol{\varepsilon}, d_\mu]$ indicates that the macro-strain, $\boldsymbol{\varepsilon}$ (the action, or external driving force for the RVE problem), and $d_\mu \in [0,1]$ are considered free *parameters* of the mapping $\tilde{\boldsymbol{\varepsilon}}_\mu \mapsto \varphi_\mu$, not necessarily satisfying Eq. (13). Therefore, in expression (15), d_μ is assumed to be independent of $\tilde{\boldsymbol{\varepsilon}}_\mu$. Under this condition, the expression

$$\boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu, d_\mu) = \frac{\partial}{\partial\tilde{\boldsymbol{\varepsilon}}_\mu}\varphi_\mu(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu, d_\mu) = \frac{\partial}{\partial\boldsymbol{\varepsilon}}\varphi_\mu(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu, d_\mu) \tag{16}$$

defines the stress only if it is combined with the evolution equation (13).

Next, in consonance with the concepts introduced in Appendix B.1, we define the parameterized functional:

$$\Pi_{[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda}) = \int_{\mathcal{B}} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu)\, d\mathcal{B} + \boldsymbol{\lambda} : \int_{\mathcal{B}} \tilde{\boldsymbol{\varepsilon}}_\mu\, d\mathcal{B}. \tag{17}$$

In Eq. (17) $\boldsymbol{\lambda}(t) \in \mathbb{S}^{n\times n}$, with uniform distribution on $\mathcal{B}_\mu$, is a symmetric second order tensor Lagrange multiplier enforcing condition (3) on the micro-strain fluctuations $\tilde{\boldsymbol{\varepsilon}}_\mu$.

With the parameterized functional $\Pi_{[\boldsymbol{\varepsilon},d_\mu]}$, we define the following:

PROBLEM II (*RVE saddle point problem)*:
Given the macroscale strain, $\boldsymbol{\varepsilon}$, find $\tilde{\boldsymbol{\varepsilon}}_\mu$ and $\boldsymbol{\lambda}$ satisfying

$$\{\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{\varepsilon}, d_\mu), \boldsymbol{\lambda}(\boldsymbol{\varepsilon}, d_\mu)\} = \arg\Big\{\min_{\tilde{\boldsymbol{\varepsilon}}_\mu\in\mathcal{E}_\mu}\max_{\boldsymbol{\lambda}\in\mathbb{S}^{n\times n}} \Pi_{[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda})\Big\} \tag{18}$$

such that:

$$\dot{d}_\mu(\boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu) \tag{19}$$

After considering that the micro-stress $\boldsymbol{\sigma}_\mu$ is given by the identity $\boldsymbol{\sigma}_\mu = \partial\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}/\partial\tilde{\boldsymbol{\varepsilon}}_\mu$ in conjunction with the evolution equation (19), from the saddle point PROBLEM II the following optimality conditions emerge:

$$\int_{\mathcal{B}_\mu} [\boldsymbol{\sigma}_\mu(\tilde{\boldsymbol{\varepsilon}}_\mu)(\boldsymbol{y}) + \boldsymbol{\lambda}] : \hat{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y})\, d\mathcal{B} = 0; \quad \forall\hat{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}) \in \mathcal{E}_\mu \tag{20}$$

$$\widehat{\boldsymbol{\lambda}} : \int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu\, d\mathcal{B} = \mathbf{0}; \quad \forall\widehat{\boldsymbol{\lambda}} \in \mathbb{S}^{n\times n}. \tag{21}$$

Comparison of Eqs. (20) and (21) with Eqs. (9) and (10) shows that the saddle-point PROBLEM II is equivalent to the rephrased PROBLEM I-R and, therefore, PROBLEM II *constitutes a primitive statement* of PROBLEM I, in Eqs. (5) and (6).

It is remarkable that the *primitive* PROBLEM II involves the numerical integration of a single *scalar* functional $\Pi_{[\boldsymbol{\varepsilon},d_\mu]}$ (see Eqs. (18) and (17)) instead of the *integration of vector-dimensional* entities, as the ones stemming from the optimality conditions in Eqs. (20) and (21). Therefore, taking Eq. (17) as the start point for the solution of the problem, instead of Eqs. (9) and (10), offers an opportunity for a simplified reduced-integration strategy, as the one explored in Section 3.3. Also notice, in Eq. (20), that the micro-strain variations, $\hat{\boldsymbol{\varepsilon}}_\mu$, are not constrained by the restriction defined in Eq. (3).

3. RVE reduced order model

Let us start by introducing a standard finite element technique based on the formulation in PROBLEM I. The so-defined discrete problem is termed the *High-Fidelity (HF) finite element model*, and solutions associated with it

are denoted HF solutions. High-dimensional fields of displacement fluctuations, $\tilde{\boldsymbol{u}}_\mu$, micro-strain fluctuations, $\tilde{\boldsymbol{\varepsilon}}_\mu$, and micro-stresses, $\boldsymbol{\sigma}_\mu$, are obtained as an outcome, during a sampling process, of successive solutions of the original HF micro-cell problem. Next, a low-dimensional approximation of the micro-strain fluctuation field is derived.

Voigt's notation, as an alternative to the tensor description, is used from now on. In particular, micro-strains and micro-stresses are denoted as column vectors: $\tilde{\boldsymbol{\varepsilon}}_\mu \in \mathbb{R}^{n_\sigma}$ and $\boldsymbol{\sigma}_\mu \in \mathbb{R}^{n_\sigma}$.

3.1. Problem formulated in reduced strains

In the HF formulation, a finite element mesh covering the RVE domain $\mathcal{B}_\mu$ is considered, while vectors $\tilde{\boldsymbol{\varepsilon}}_\mu$ and $\boldsymbol{\sigma}_\mu$ are defined at each Gauss Point of this mesh. Being the number of Gauss Points N_g, then, the dimension of the HF finite element micro-strain fluctuation and micro-stress spaces, are of the order: $\mathcal{O}(N_g)$.

We search for a reduced vectorial space of dimension n_ε, with $n_\varepsilon \ll N_g$, for modeling the micro-strain fluctuations. This low-dimension space is obtained as the linear expansion of an orthogonal basis of n_ε spatial functions: $\{\boldsymbol{\Psi}(\boldsymbol{y})\} = \{\boldsymbol{\Psi}_1(\boldsymbol{y}), \ldots, \boldsymbol{\Psi}_{n_\varepsilon}(\boldsymbol{y})\}$ (with $\boldsymbol{\Psi}_i(\boldsymbol{y}) \in \mathbb{R}^{n_\sigma}; i = 1, \ldots, n_\varepsilon$), as follows:

$$\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}, t) = \sum_{i=1}^{n_\varepsilon} \boldsymbol{\Psi}_i(\boldsymbol{y}) c_i(t) = \boldsymbol{\Psi}(\boldsymbol{y})\boldsymbol{c}(t) \tag{22}$$

where each element $\boldsymbol{\Psi}_i$, of the basis $\{\boldsymbol{\Psi}\}$, is recognized as a micro-strain fluctuation mode and the vector of time dependent coefficients $\boldsymbol{c}(t) = [c_1, \ldots, c_{n_\varepsilon}]$ ($\boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}$) represents the amplitude of these modes. In the last identity of (22), the matrix $\boldsymbol{\Psi}(\boldsymbol{y}) = [\boldsymbol{\Psi}_1, \ldots, \boldsymbol{\Psi}_{n_\varepsilon}]$, with $\boldsymbol{\Psi}(\boldsymbol{y}) \in \mathbb{R}^{\boldsymbol{n}_\sigma \times n_\varepsilon}$ collects, in column, the n_ε micro-strain modes of the basis $\{\boldsymbol{\Psi}\}$. Note that, in order to preserve a simplified notation, an identical symbol is used to identify the micro-strain fluctuation field in the HF approach and in the low-dimensional approach.

The basis $\{\boldsymbol{\Psi}\}$, composed of n_ε basis vectors (or modes), is obtained through a representative off-line sampling program of the micro-cell model, solved with the HF formulation of PROBLEM I.[8] From these solutions, the dominant reduced modes of the micro-strain fluctuations are extracted through a Proper Orthogonal Decomposition (POD) technique. Additional details of this procedure are given in the next sub-Section.

The variations of the micro-strain fluctuations are adopted with an identical approach to (22):

$$\hat{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}) = \boldsymbol{\Psi}(\boldsymbol{y})\hat{\boldsymbol{c}} \tag{23}$$

where the vector $\hat{\boldsymbol{c}}$ (with $\hat{\boldsymbol{c}} \in \mathbb{R}^{n_\varepsilon}$) collects the coefficients of the linear expansion.

Introducing (22) and (23) into the variational equations (20) and (21), derived from PROBLEM II, results in a new model written in a reduced basis. It is termed the Reduced-Order Model (ROM) and it is presented in the following box:

PROBLEM III (ROM) (*RVE saddle point problem)*:
Given the macroscale strain, $\boldsymbol{\varepsilon}$, find $\boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{n_\sigma}$ satisfying:

$$\begin{aligned}\{\boldsymbol{c}(\boldsymbol{\varepsilon}, d_\mu), \boldsymbol{\lambda}(\boldsymbol{\varepsilon}, d_\mu)\} &= \arg\left\{\min_{\boldsymbol{c}\in\mathbb{R}^{n_\varepsilon}} \max_{\boldsymbol{\lambda}\in\mathbb{R}^{n_\sigma}} \Pi_{[\boldsymbol{\varepsilon}, d_\mu]}(\boldsymbol{\Psi}\boldsymbol{c}, \boldsymbol{\lambda})\right\} \\ &= \arg\left\{\min_{\boldsymbol{c}\in\mathbb{R}^{n_\varepsilon}} \max_{\boldsymbol{\lambda}\in\mathbb{R}^{n_\sigma}} \left[\int_{\mathcal{B}} \varphi_{\mu[\boldsymbol{\varepsilon}, d_\mu]}(\boldsymbol{\varepsilon} + \boldsymbol{\Psi}\boldsymbol{c})\, d\mathcal{B} + \boldsymbol{\lambda}^T \left(\int_{\mathcal{B}} \boldsymbol{\Psi} d\mathcal{B}\right)\boldsymbol{c}\right]\right\}\end{aligned} \tag{24}$$

such that:

$$\dot{d}_\mu(\boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu) \tag{25}$$

[8] In the next Section, it is shown that every function $\boldsymbol{\Psi}_i$ in the basis spanning the reduced strain fluctuation fields, belongs by construction, to the space $\mathcal{E}_\mu$ (i.e.: $\boldsymbol{\Psi}_i \in \mathcal{E}_\mu$). This guarantees that any function $\tilde{\boldsymbol{\varepsilon}}_\mu$, obtained by linear combination of these bases, also satisfies the condition: $\tilde{\boldsymbol{\varepsilon}}_\mu \in \mathcal{E}_\mu$.

After straightforward variational manipulations, the optimality conditions for the problem above yield,

$$\frac{\partial}{\partial \boldsymbol{c}^T}\Pi_{[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi c},\boldsymbol{\lambda}) = \int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}+\boldsymbol{\Psi c})d\mathcal{B} + \left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi}^T\, d\mathcal{B}\right)\boldsymbol{\lambda} = \boldsymbol{0}; \tag{26}$$

$$\frac{\partial}{\partial \boldsymbol{\lambda}^T}\Pi_{[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi c},\boldsymbol{\lambda}) = \left(\int_{\mathcal{B}_\mu} \boldsymbol{\Psi} d\mathcal{B}\right)\boldsymbol{c} = \boldsymbol{0}. \tag{27}$$

The system of equations (26) and (27) consists of n_ε and n_σ algebraic equations, respectively. They are enough to find the vector of coefficients, $\boldsymbol{c}$, and the Lagrange multiplier vector $\boldsymbol{\lambda}$.

3.2. Construction of a POD basis for the micro-strain fluctuations

Typically, the integrals in Eqs. (26) and (27) are evaluated by means of a numerical quadrature, such as the conventional Gauss integration rule. Then, the kernel of the integrals has to be only evaluated in the quadrature points of the original finite element mesh. In accordance with this procedure, by denoting N_g the number of Gauss points, only the values $\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_i,t)$ of the reduced field at Gauss Point positions: $\boldsymbol{y}_i$, with $i=1,\ldots,N_g$, should be determined:

$$\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_i,t) = \boldsymbol{\Psi}(\boldsymbol{y}_i)\boldsymbol{c}(t); \qquad \text{for } \boldsymbol{y}_i \;\text{ with }\; i=1,\ldots,N_g. \tag{28}$$

A convenient way to express the PROBLEM III in a compact notation is to collect in one-column vector $[\tilde{\boldsymbol{\varepsilon}}_\mu]_G$ all the micro-strain fluctuation vectors $\tilde{\boldsymbol{\varepsilon}}_\mu \in \mathbb{R}^{n_\sigma}$ corresponding to the set of Gauss points of the original HF finite element mesh and the modes (basis vectors) evaluated in the Gauss point position in a matrix $\boldsymbol{\Psi}_G$, resulting:

$$[\tilde{\boldsymbol{\varepsilon}}_\mu]_G = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_1) \\ \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_2) \\ \vdots \\ \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_{N_g}) \end{bmatrix} = \boldsymbol{\Psi}_G\boldsymbol{c}; \qquad \boldsymbol{\Psi}_G = \begin{bmatrix} \boldsymbol{\Psi}(\boldsymbol{y}_1) \\ \boldsymbol{\Psi}(\boldsymbol{y}_2) \\ \vdots \\ \boldsymbol{\Psi}(\boldsymbol{y}_{N_g}) \end{bmatrix} \tag{29}$$

with:

$$[\tilde{\boldsymbol{\varepsilon}}_\mu]_G \in \mathbb{R}^{(n_\sigma N_g)}; \qquad \boldsymbol{\Psi}_G \in \mathbb{R}^{(n_\sigma N_g \times n_\varepsilon)}; \qquad \boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}. \tag{30}$$

Subindex G (standing for "Global"), which is added to the matrix of modes, means that this matrix is built by piling up, in a single column, the modes evaluated at each Gauss point of the HF finite element model, as it is shown in (29) (and unambiguously defined by the dimension of the vector). In the following, all matrices with subindex G will have a similar meaning. It is also important to notice that the SVD technique, when applied to the snapshot matrix $[\boldsymbol{X}]$, determines the basis $\{\boldsymbol{\Psi}(\boldsymbol{y})\}$ only through the matrix $\boldsymbol{\Psi}_G$. Therefore, in the present approach, this basis is not determined for arbitrary $\boldsymbol{y}$.

Due to the linear character of the procedure by which the basis $\boldsymbol{\Psi}_G$ is constructed, and the linear homogeneous character of the compatibility equations for strains, (8), every vector of $\boldsymbol{\Psi}_G$ fulfills the compatibility equations, then: $\boldsymbol{\Psi}_G \in \mathcal{E}_\mu$. This property is preserved for the reduced basis $\{\boldsymbol{\Psi}\}$ even after performing the spatial domain partition described in Section 3.2.2. Therefore, any tensor field obtained by a linear span of $\{\boldsymbol{\Psi}\}$, through an arbitrary vector of parameters $\boldsymbol{c}$, as in Eq. (28), also belongs to $\mathcal{E}_\mu$.

Finally, introducing Eq. (29) into the ROM equations (26)–(27), the following system of equations, in compact notation, is obtained:

$$\boldsymbol{\Psi}_G^T[\mathbb{W}]\Big([\boldsymbol{\sigma}_\mu(\boldsymbol{c})]_G + [\boldsymbol{\lambda}]_G\Big) = \boldsymbol{0}; \tag{31}$$

$$[W]\,\boldsymbol{\Psi}_G\boldsymbol{c} = \boldsymbol{0}; \tag{32}$$

where $[\boldsymbol{\sigma}_\mu(\boldsymbol{c})]_G \in \mathbb{R}^{n_\sigma N_g}$ is the column vector constituted by piling-up the N_g stress vectors, $\boldsymbol{\sigma}_\mu(\boldsymbol{c}) \in \mathbb{R}^{n_\sigma}$, evaluated at the Gauss points. The column vector $[\boldsymbol{\lambda}]_G$ (with $[\boldsymbol{\lambda}]_G \in \mathbb{R}^{n_\sigma N_g}$) is also the piled-up of N_g repetitive values of the same vector $\boldsymbol{\lambda} \in \mathbb{R}^{n_\sigma}$. The square diagonal matrix $[\mathbb{W}] \in \mathbb{R}^{(N_g n_\sigma \times N_g n_\sigma)}$ and the rectangular matrix $[W] \in \mathbb{R}^{n_\sigma \times N_g n_\sigma}$

collect the Gauss point weights: $\{W_1, W_2, \ldots, W_{N_g}\}$, which for plane stress or plane strain cases (with $n_\sigma = 4$) are distributed in sub-block matrices $\mathbb{W}_i \in \mathbb{R}^{4\times 4}$ ($i = 1, \ldots, N_g$), as follows:

$$[\mathbb{W}] = \begin{bmatrix} \mathbb{W}_1 & \ldots & \mathbb{O} \\ \mathbb{O} & \ddots & \mathbb{O} \\ \mathbb{O} & \ldots & \mathbb{W}_{N_g} \end{bmatrix}; \qquad [W] = \begin{bmatrix} \mathbb{W}_1 & \ldots & \mathbb{W}_{N_g} \end{bmatrix}; \tag{33}$$

$$\mathbb{W}_i = \begin{bmatrix} W_i & 0 & 0 & 0 \\ 0 & W_i & 0 & 0 \\ 0 & 0 & W_i & 0 \\ 0 & 0 & 0 & W_i \end{bmatrix}.$$

In the set of equations (31)–(32), the internal damage variables, $d_\mu(\boldsymbol{y}_i)$, defining the stresses $\boldsymbol{\sigma}_\mu(\boldsymbol{c})$ at each Gauss point, $\boldsymbol{y}_i$, implicitly satisfy the evolution equation (13).

3.2.1. Collecting sampled solutions: the sampling program

Transition from the high-dimensional finite element space corresponding to the micro-strain fluctuations, $\tilde{\boldsymbol{\varepsilon}}_\mu$, to the desired reduced-order space spanned by the basis $\{\boldsymbol{\Psi}\}$, is accomplished by resorting to the Proper Orthogonal Decomposition (POD) technique, such as explained in [26] (see also [27]). Following that technique, the first step consists of generating a collection of solutions, supplied by an off-line performed sampling program, of the High Fidelity (HF) micro-cell finite element problem governed by Eqs. (5) and (6) (PROBLEM I).

This micro-cell sampling program consists of taking samples of RVE solutions for different trial cases. In each trial case, the RVE is subjected to a given predefined history of the input driving force (or action). The sampling program, should be designed such that the collected data determines a good-enough span of all possible loading histories. The larger is the number of loading history cases represented in the data set the larger is the accuracy obtained with the reduced model. Therefore, a good sampling program is the one whose loading HF model histories appropriately cover a large part of the parametric space of solutions.

The actions, or loading system, of the micro-cell VBVP are the macro-strains $\boldsymbol{\varepsilon}$. In each trial case corresponding to a specific loading condition, the micro-strain fluctuation $\tilde{\boldsymbol{\varepsilon}}_\mu$, at every Gauss point and for a given loading-step, is picked up and collected. Section 4 gives additional details about the RVE sampling program designed for the present problem.

Let us consider the kth trial solution collected from the sampling program and which corresponds to a given time step increment and loading case. The micro-strain fluctuation of all Gauss points provided by this solution is denoted with supra-index k: $\tilde{\boldsymbol{\varepsilon}}^k_\mu(\boldsymbol{y}_i)$ (with $i = 1, \ldots, N_g$). Then, each column vector $\boldsymbol{X}_k$ (with $\boldsymbol{X}_k \in \mathbb{R}^{(N_g n_\sigma)}$) of the snapshot matrix, $[\boldsymbol{X}]$, is defined as follows:

$$\boldsymbol{X}_k = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}^k_\mu(\boldsymbol{y}_1) \\ \tilde{\boldsymbol{\varepsilon}}^k_\mu(\boldsymbol{y}_2) \\ \vdots \\ \tilde{\boldsymbol{\varepsilon}}^k_\mu(\boldsymbol{y}_{N_g}) \end{bmatrix}. \tag{34}$$

Furthermore, the snapshot matrix:

$$[\boldsymbol{X}] = [\boldsymbol{X}_1\, \boldsymbol{X}_2 \ldots \boldsymbol{X}_{p_{snp}}]; \qquad \in \mathbb{R}^{(N_g n_\sigma)\times p_{snp}} \tag{35}$$

is constituted by n_{nsp} snapshot vectors. Therefore, $[\boldsymbol{X}]$ represents a number of time-step solutions, in terms of micro-strain fluctuations, obtained with the HF model of PROBLEM I, under different loading conditions.

3.2.2. Partition of the micro-strain reduced basis

In order to get a more accurate estimation of the dominant modes of the micro-strain fluctuation fields, it is convenient to partition the snapshot matrix components in accordance with: (i) the material response observed during the load history in each sampled trajectory, and (ii) the different micro-cell domains. This procedure is sketched in Fig. 2.

The snapshots taken from a given sampled trajectory are categorized depending on whether they are taken in the micro-cell elastic regime (therefore all points of the micro-cell are in elastic state), or snapshots taken during the

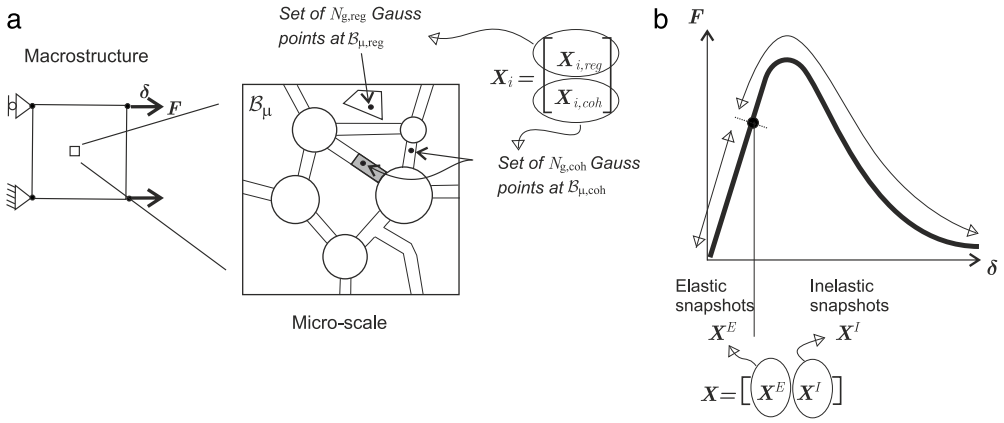


Fig. 2. Partition of the snapshot matrix $\boldsymbol{X}$. (a) Entries are partitioned into two sub-blocks: micro-strain fluctuations in $\boldsymbol{X}_{i,reg}$ are from points at $\mathcal{B}_{\mu,\text{reg}}$, while micro-strain fluctuations in $\boldsymbol{X}_{i,coh}$ are from Gauss points at $\mathcal{B}_{\mu,\text{coh}}$. (b) Snapshots taken during the elastic regime of the micro-cell (all Gauss points in $\mathcal{B}_{\mu,\text{coh}}$ are in elastic states) correspond to the sub-block $\boldsymbol{X}^E$. Snapshots taken during the inelastic regime (at least one Gauss point in $\mathcal{B}_{\mu,\text{coh}}$ is in inelastic state) correspond to the sub-block $\boldsymbol{X}^I$ (domains $\mathcal{B}_{\mu,\text{reg}}$ and $\mathcal{B}_{\mu,\text{coh}}$ are defined in Appendix A).

inelastic regime (at least one point in the macro-cell is damaging). In accordance with this criterion, the full snapshot matrix can be partitioned into two sub-blocks of columns, i.e.:

$$[\boldsymbol{X}] = [\boldsymbol{X}^E\ \boldsymbol{X}^I]. \tag{36}$$

In the sub-block of elastic snapshots, denoted $\boldsymbol{X}^E$, each column is one solution of the elastic regime, for some of the sampling trajectories. The complementary sub-block is built with inelastic snapshots, and is denoted $\boldsymbol{X}^I$.

An additional partition of $[\boldsymbol{X}]$ is done by considering the singular domain of the RVE, represented by $\mathcal{B}_{\mu,\text{coh}}$, and its disjoint domain, denoted regular domain $\mathcal{B}_{\mu,\text{reg}}$. These domains are defined in Appendix A, see Figs. 1(b) and 12. Since each Gauss point in the HF finite element model lies on one of those domains, without loss of generality the snapshot entries are organized so that the first, $N_{g,reg}$ entries correspond to Gauss points in $\mathcal{B}_{\mu,\text{reg}}$, while the remaining $N_{g,coh}$ entries correspond to Gauss points in $\mathcal{B}_{\mu,\text{coh}}$. Clearly: $N_g = N_{g,reg} + N_{g,coh}$.

Then, in accordance with this criterion, the snapshots (36) can be now partitioned into rows as follows:

$$[\boldsymbol{X}] = \begin{bmatrix} \boldsymbol{X}_{\text{reg}} \\ \boldsymbol{X}_{\text{coh}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^E_{\text{reg}}\ \boldsymbol{X}^I_{\text{reg}} \\ \boldsymbol{X}^E_{\text{coh}}\ \boldsymbol{X}^I_{\text{coh}} \end{bmatrix}; \tag{37}$$

where $\boldsymbol{X}_{\text{reg}}$ collects the micro-strain fluctuations of the first N_g^{reg} Gauss points and $\boldsymbol{X}_{\text{coh}}$ collects the micro-strain fluctuations of the remaining $N_{g,coh}$ Gauss points. In (37), the second identity emphasizes the double partition done in accordance with the history loading stage of each snapshot and the spatial domain where the micro-strains have been picked up from.

The second step of this procedure consists of finding a reduced POD basis, utilizing a SVD technique, of the elastic snapshots. The SVD technique is applied, in a separate way, for each partition corresponding to $\mathcal{B}_{\mu,\text{reg}}$ and $\mathcal{B}_{\mu,\text{coh}}$, respectively, as follows:

$$\boldsymbol{\Psi}^E_{\text{G,reg}} \leftarrow \text{SVD}[\boldsymbol{X}^E_{\text{reg}}]; \tag{38}$$

$$\boldsymbol{\Psi}^E_{\text{G,coh}} \leftarrow \text{SVD}[\boldsymbol{X}^E_{\text{coh}}]. \tag{39}$$

Therefore, by construction, the basis of the elastic reduced-order micro-strain fluctuation space:

$$\boldsymbol{\Psi}^E_G = \begin{bmatrix} \boldsymbol{\Psi}^E_{\text{G,reg}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Psi}^E_{\text{G,coh}} \end{bmatrix}; \qquad \boldsymbol{\Psi}^E_{\text{G,reg}} \in \mathbb{R}^{(N_{g,reg}\times 3)};\ \ \boldsymbol{\Psi}^E_{\text{G,coh}} \in \mathbb{R}^{(N_{g,coh}\times 3)}; \tag{40}$$

reproduces the elastic response of the micro-cell. Furthermore, each sub-block, $\boldsymbol{\Psi}^E_{\mathrm{G,reg}}$ and $\boldsymbol{\Psi}^E_{\mathrm{G,coh}}$, has dimension 3. This dimension results from the linear response of the mechanical system during the elastic regime and from the loading parameter space dimension, which is 3. By normalizing each column of the matrix, $\boldsymbol{\Psi}^E_G$, an orthonormal basis is obtained, in the sense that:

$$[\boldsymbol{\Psi}^E_G]^T_i\,[\boldsymbol{\Psi}^E_G]_j = \delta_{ij}; \qquad \delta_{ij} = 1;\ \text{if } i = j; \qquad \text{otherwise } \delta_{ij} = 0; \tag{41}$$

where $[\cdot]_i$ and $[\cdot]_j$ denote columns ith and jth, respectively, of the matrix of modes of the reduced micro-strain fluctuations.

In order to preserve the orthogonality of the full basis $\{\boldsymbol{\Psi}\}$, the method in [17] is used. Therefore, to obtain the remaining elements of $\{\boldsymbol{\Psi}\}$, not contained in the space spanned by $\{\boldsymbol{\Psi}^E\}$, the inelastic snapshots $\boldsymbol{X}^I$ are projected into the orthogonal subspace spanned by the basis $\boldsymbol{\Psi}^E$:

$$\boldsymbol{X}^{I\perp}_{\mathrm{reg}} = \boldsymbol{X}^I_{\mathrm{reg}} - \sum_i \Big([\boldsymbol{\Psi}^E_{\mathrm{G,reg}}]_i \cdot \boldsymbol{X}^I_{\mathrm{reg}}\Big)[\boldsymbol{\Psi}^E_{\mathrm{G,reg}}]_i \tag{42}$$

$$\boldsymbol{X}^{I\perp}_{\mathrm{coh}} = \boldsymbol{X}^I_{\mathrm{coh}} - \sum_i \Big([\boldsymbol{\Psi}^E_{\mathrm{G,coh}}]_i \cdot \boldsymbol{X}^I_{\mathrm{coh}}\Big)[\boldsymbol{\Psi}^E_{\mathrm{G,coh}}]_i \tag{43}$$

then, similarly to (38)–(39), the POD technique is separately applied to the orthogonal projections of the snapshots $\boldsymbol{X}^{I\perp}_{\mathrm{reg}}$ and $\boldsymbol{X}^{I\perp}_{\mathrm{coh}}$ to obtain:

$$\boldsymbol{\Psi}^I_{\mathrm{G,reg}} \leftarrow \mathrm{SVD}[\boldsymbol{X}^{I\perp}_{\mathrm{reg}}]; \tag{44}$$

$$\boldsymbol{\Psi}^I_{\mathrm{G,coh}} \leftarrow \mathrm{SVD}[\boldsymbol{X}^{I\perp}_{\mathrm{coh}}]. \tag{45}$$

In this way, and by construction, the basis:

$$\boldsymbol{\Psi}^I_G = \begin{bmatrix} \boldsymbol{\Psi}^I_{\mathrm{G,reg}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^I_{\mathrm{G;coh}} \end{bmatrix}; \qquad \boldsymbol{\Psi}^I_{\mathrm{G,reg}} \in \mathbb{R}^{(N_{g,reg}\times n^I_{\varepsilon,reg})};\ \ \boldsymbol{\Psi}^I_{\mathrm{G,coh}} \in \mathbb{R}^{(N_{g,coh}\times n^I_{\varepsilon,coh})}; \tag{46}$$

reproduces the inelastic micro-strain fluctuations response in the micro-cell. In this equation, $n^I_{\varepsilon,reg}$ and $n^I_{\varepsilon,coh}$ are the number of dominant micro-strain inelastic modes in $\mathcal{B}_{\mu,\mathrm{reg}}$ and $\mathcal{B}_{\mu,\mathrm{coh}}$, respectively.

By construction, the space spanned by $\{\boldsymbol{\Psi}^I_G\}$ is orthogonal to that spanned by $\{\boldsymbol{\Psi}^E_G\}$. Therefore, from (40) and (46) it is possible to combine them into the reduced orthonormal basis:

$$\{\boldsymbol{\Psi}_G\} = \{\boldsymbol{\Psi}^E_G;\ \boldsymbol{\Psi}^I_G\} \tag{47}$$

which spans the full snapshot matrix $\boldsymbol{X}$.

The number of basis vectors in $\{\boldsymbol{\Psi}_G\}$ is: $n_\varepsilon = 6 + n^I_{\varepsilon,reg} + n^I_{\varepsilon,coh}$.[9] The values $n^I_{\varepsilon,reg}$ and $n^I_{\varepsilon,coh}$ are obtained from the solution of the SVD procedure,[10] Eqs. (44) and (45), by taking those modes associated with the largest singular values (see [26]).

3.3. Hyper-reduced order model (HPROM)

In many cases, it is an accepted fact that, although ROM markedly reduces the number of unknowns in the problem (in the present case, the dimension n_ε of the vector $\boldsymbol{c}$ in Eq. (22)), this does not translate into an actual reduction of the computational cost and, consequently, in a problem speedup (see [26] for details). Therefore, further actions should be taken. These actions are known in the literature with the term *hyper-reduction* [12] which gives rise to the HyPer-Reduced Order Model (HPROM).

[9] In general, n_ε is smaller than the dimension of the space expanded by the columns of $\boldsymbol{X}$. Therefore, there exists an inherent error (the a-priori error) for capturing arbitrary snapshots $\boldsymbol{X}_i$ through a linear expansion of the base $\{\boldsymbol{\Psi}_G\}$. This issue is studied in Section 5.

[10] For the present case, the behavior in $\mathcal{B}_{\mu,\mathrm{reg}}$ is elastic, therefore $n^I_{\varepsilon,reg} = 0$.

The reason for the low-efficiency of the ROM is that the numerical integration of the integrals in (31) is yet associated with a high computational cost, since it is necessary to evaluate the kernels in all the N_g Gauss points. Therefore, the objective is to introduce an additional cost-reduction technique aiming at diminishing the computational burden in evaluating these integrals.

To pursue this objective, a *reduced integration technique* has been developed by resorting to a nonconventional method, termed Reduced Optimal Quadrature (ROQ), following the ideas that are presented, within a more general context, in Appendix B. It is based on two actions:

1. Derive a reduced numerical integration scheme for some terms of the primitive PROBLEM III in Eqs. (24) and (25): i.e. to evaluate the integral functional in Eq. (24), the classical Gauss integration rule is replaced by an optimal reduced quadrature. More specifically, the integral term involving the free energy, $\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}$ in Eq. (24) is evaluated as

$$\int_{\mathcal{B}_\mu}^{Gauss} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\boldsymbol{y})\boldsymbol{c})\, d\mathcal{B} \simeq \sum_{j=1}^{N_r} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\boldsymbol{z}_j)\boldsymbol{c})\ \omega_j\ d\mathcal{B} := \int_{\mathcal{B}_\mu}^{Reduced} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\boldsymbol{y})\boldsymbol{c})\, d\mathcal{B}. \tag{48}$$

   The reduced quadrature is based on selecting, through an adequate algorithm, a set of discrete sampling points, $\boldsymbol{z}_j;\ j = 1,\dots,N_r$ and the corresponding weights $\omega_j;\ j = 1,\dots,N_r$. The kernel of the integrand is evaluated in those discrete points, $\boldsymbol{z}_j$, and weighted, with factors, ω_j, to approach the exact integral as shown in Eq. (48). The success of the reduced integration numerical scheme, in front of the conventional Gaussian quadrature, lies on the fact that it is possible to reduce notably the number of involved quadrature points to $N_r \ll N_g$, with respect to the number of Gauss points, N_g, in the original HF model, but keeping under strict control the numerical error introduced by the reduced quadrature.
2. Replace the reduced integrated functional (48) into the primitive statement, Eq. (24) of PROBLEM III and find the corresponding optimality conditions (26) and (27). Considering that the derivative of Eq. (48) with respect to the parameters $\boldsymbol{c}$, can be written as:

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{c}} \int_{\mathcal{B}_\mu}^{Reduced} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\boldsymbol{y})\boldsymbol{c})\, d\mathcal{B} &= \int_{\mathcal{B}_\mu}^{Reduced} \frac{\partial}{\partial \boldsymbol{c}} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\underbrace{\boldsymbol{\Psi}(\boldsymbol{y})\boldsymbol{c}}_{=\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},\boldsymbol{c})})\, d\mathcal{B} \\ &= \int_{\mathcal{B}_\mu}^{Reduced} \underbrace{\left(\frac{\partial \tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},\boldsymbol{c})}{\partial \boldsymbol{c}}\right)^T}_{\boldsymbol{\Psi}(\boldsymbol{y})^T} \underbrace{\frac{\partial \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},\boldsymbol{c}))}{\partial \tilde{\boldsymbol{\varepsilon}}_\mu}}_{\boldsymbol{\sigma}_\mu(\boldsymbol{y},\boldsymbol{c})}\, d\mathcal{B} \\ &= \int_{\mathcal{B}_\mu}^{Reduced} \boldsymbol{\Psi}(\boldsymbol{y})^T \boldsymbol{\sigma}_\mu(\boldsymbol{y},\boldsymbol{c})\, d\mathcal{B} \simeq \sum_{j=1}^{N_r} \boldsymbol{\Psi}(\boldsymbol{z}_j)^T\ \boldsymbol{\sigma}_\mu(\boldsymbol{z}_j,\boldsymbol{c})\ \omega_j.\end{aligned} \tag{49}$$

   Eqs. (49) can be utilized to evaluate the derivatives in the optimality condition (26) yielding the following set of equations (from now on termed HPROM formulation):

PROBLEM IV (HPROM) (*RVE saddle point problem*):
Given the macroscale strain, $\boldsymbol{\varepsilon}$, find $\boldsymbol{c} \in \mathbb{R}^{n_\varepsilon}$ and $\boldsymbol{\lambda} \in \mathbb{R}^{n_\sigma}$ satisfying:

$$\int_{\mathcal{B}_\mu}^{Reduced} \boldsymbol{\Psi}(\boldsymbol{y})^T \boldsymbol{\sigma}_\mu(\boldsymbol{y},\boldsymbol{c})\, d\mathcal{B} + \left(\int_{\mathcal{B}_\mu}^{Gauss} \boldsymbol{\Psi}(\boldsymbol{y})^T\, d\mathcal{B}\right)\boldsymbol{\lambda} = \boldsymbol{0}; \tag{50}$$

$$\left(\int_{\mathcal{B}_\mu}^{Gauss} \boldsymbol{\Psi}(\boldsymbol{y})\, d\mathcal{B}\right)\boldsymbol{c} = \boldsymbol{0}; \tag{51}$$

A similar procedure could be also used for the other integral terms in Eqs. (50) ($\int_{\mathcal{B}}^{Gauss} \boldsymbol{\Psi}^T\ d\mathcal{B}$) and (51) ($\int_{\mathcal{B}}^{Gauss} \boldsymbol{\Psi}\ d\mathcal{B}$), but, these being constant terms (not depending on the unknowns of the problem) they can be integrated

once for all (presumably in the off-line stage), using the standard Gauss quadrature, the result being stored and used, when necessary, in the on-line stage.

3.3.1. Determination of the Reduced Optimal Quadrature (ROQ)

The minimum number of quadrature points providing an admissible integration error in the free energy integral (48) can be determined by considering an *optimal* linear expansion of $\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}$ in terms of energy modes Φ_i $(i = 1, \dots, n_\varphi)$ and amplitudes f_i, i.e. a similar expression to that adopted for constructing the reduced micro-strain fluctuation in Eq. (22), as follows:

$$\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{y})\boldsymbol{c}) = \sum_{i=1}^{n_\varphi} \Phi_i(\mathbf{y})\, f_i(\boldsymbol{c}, \boldsymbol{\varepsilon}, d_\mu). \tag{52}$$

Replacing this linear expansion into the integral (48) yields

$$\int_{\mathcal{B}_\mu} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{y})\boldsymbol{c})\; d\mathcal{B} = \sum_{i=1}^{n_\varphi} \Big(\int_{\mathcal{B}_\mu} \Phi_i(\mathbf{y}) d\mathcal{B} \Big) f_i(\boldsymbol{c}, \boldsymbol{\varepsilon}, d_\mu). \tag{53}$$

From Eq. (52) it is clear that the error associated with the numerical integration of the right-hand side, *relies only on the error associated with the integration of every one of the n_φ modes, $\Phi_i(\mathbf{y})$*. Therefore, the strategy, for finding an optimal numerical integration translates into seeking a reduced optimal quadrature rule (48) *for every mode Φ_i*:

$$\int_{\mathcal{B}_\mu} \Phi_i(\mathbf{y})\; d\mathcal{B} \approx \int_{\mathcal{B}_\mu}^{Reduced} \Phi_i(\mathbf{y})\; d\mathcal{B} = \sum_{j=1}^{N_r} \Phi_i(\mathbf{z}_j)\; \omega_j; \qquad \text{for: } i = 1, \dots, n_\varphi. \tag{54}$$

Following the strategy proposed in [17], the positions, $\mathbf{z}_j$, weights, ω_j, and number, N_r, of quadrature points can be determined by solving an adjoint minimization problem, where the objective is to minimize the error introduced by evaluating the n_φ reduced integrals in Eq. (54), that is

$$\big(\underbrace{\mathbf{z}_j, \omega_j}_{j=1,\dots,N_r} \big) = \arg\Big\{ \min_{\mathbf{z}_j,\omega_j} \sum_{i=1}^{n_\varphi} \Big\{ \Big(\Big| \underbrace{\int_{\mathcal{B}_\mu}^{Gauss} \Phi_i(\mathbf{y})\; d\mathcal{B}}_{\sum_{j=1}^{N_g} \Phi_i(\mathbf{y}_j)\, \mathrm{W}_j} - \underbrace{\int_{\mathcal{B}_\mu}^{Reduced} \Phi_i(\mathbf{z})\; d\mathcal{B}}_{\sum_{j=1}^{N_r} \Phi_i(\mathbf{z}_j)\, \omega_j} \Big| + \Big| |\mathcal{B}_\mu| - \sum_{j=1}^{N_r} \omega_j \Big| \Big) \Big\} \Big\}. \tag{55}$$

The last term of the cost function in (55) forces the satisfaction of the additional constraint:

$$\int_{\mathcal{B}_\mu}^{Reduced} d\mathcal{B} = \sum_{j=1}^{N_r} \omega_j = |\mathcal{B}|. \tag{56}$$

By construction, the set of positions, $(\{\mathbf{z}_1, \dots, \mathbf{z}_{N_r}\})$ of the reduced integration scheme sampling points are selected among the positions, $(\{\mathbf{y}_1, \dots, \mathbf{y}_{Ng}\})$ of the standard Gauss integration points. A greedy algorithm for solving (55), preserving the positiveness of all the weight factors ω_j, has been proposed in [17]. This algorithm is here adopted to determine the position of points $\mathbf{z}_i$ and their corresponding weights ω_i.

Replacing the reduced quadrature rule (54) in (53), the integral of the free energy can finally be approached as:

$$\int_{\mathcal{B}_\mu} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{y})\boldsymbol{c})\; d\mathcal{B} \approx \int_{\mathcal{B}_\mu}^{reduced} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{y})\boldsymbol{c})\; d\mathcal{B} = \sum_{j=1}^{N_r} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{z}_j)\boldsymbol{c})\, \omega_j. \tag{57}$$

Notice that Eq. (57) defines an integration rule with specific sampling points $\mathbf{z}_j$ and weights ω_j $(j = 1, \dots, N_r)$, which is distinctly different to the original Gauss integration quadrature, i.e.:

$$\int_{\mathcal{B}_\mu} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{y})\boldsymbol{c})\; d\mathcal{B} \backsimeq \sum_{j=1}^{N_r} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{z}_j)\boldsymbol{c})\, \omega_j := \int_{\mathcal{B}_\mu}^{Reduced} \varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{\Psi}(\mathbf{z}_j)\boldsymbol{c})\; d\mathcal{B}. \tag{58}$$

Also the standard rules for integral's derivatives, with respect to parameters apply, i.e.:

$$\frac{\partial}{\partial \boldsymbol{\gamma}} \int_{\mathcal{B}_\mu} \varphi_\mu(\mathbf{y}, \boldsymbol{\gamma})\, d\mathcal{B} \backsimeq \frac{\partial}{\partial \boldsymbol{\gamma}} \int_{\mathcal{B}_\mu}^{Reduced} \varphi_\mu(\mathbf{y}, \boldsymbol{\gamma})\, d\mathcal{B} = \frac{\partial}{\partial \boldsymbol{\gamma}} \sum_{j=1}^{N_r} \varphi_\mu(\mathbf{z}_j, \boldsymbol{\gamma})\, \omega_j$$
$$= \sum_{j=1}^{N_r} \frac{\partial}{\partial \boldsymbol{\gamma}} \varphi_\mu(\mathbf{z}_j, \boldsymbol{\gamma})\, \omega_j := \int_{\mathcal{B}_\mu}^{Reduced} \frac{\partial}{\partial \boldsymbol{\gamma}} \varphi_\mu(\mathbf{y}, \boldsymbol{\gamma})\, d\mathcal{B} \tag{59}$$

the symbol $\boldsymbol{\gamma} \equiv \{\boldsymbol{c}, \boldsymbol{\varepsilon}, d_\mu\}$ standing for any of the parameters in Eq. (58) as it has been done in Eq. (49).

The utilized greedy algorithm provides the optimum quadrature. Some additional interesting properties are (see [17]):

- The set of N_r quadrature points ($\{\mathbf{z}_1, \ldots, \mathbf{z}_{N_r}\}$) is a sub-set of the Gauss point set associated with the HF model (by construction).
- An exact integration rule can be achieved for $N_r = n_\varphi + 1$ quadrature points, where n_φ is the number of the free energy modes in Eq. (52) (see [17] for the proof).
- The error introduced by using the (ROQ) scheme in Eq. (57), is governed by the neglected terms in the expansion (52). The quadrature error diminishes by increasing n_φ.[11]

It is remarked that, the determination of the reduced integration quadrature in Eq. (55) *depends only on the free-energy modes*, Φ_i, in Eq. (52) and *it is independent of the modal coefficients* f_i.

3.4. POD basis for the free energy

The development of the reduced quadrature rule in Eq. (54) is based on the concept of spanning the free energy through an *optimal* linear combination of the basis $\{\boldsymbol{\Phi}\}$, and integrating exactly every element of the basis with the reduced quadrature rule.

Regarding the computation of this basis in the off-line process, similar to the one described in Section 3.2 and based on the construction of a snapshot matrix and the subsequent computation of the basis from a SVD technique, two options appear:

1. Construct the free energy snapshots by collecting solutions of the ROM model in PROBLEM III (Eqs. (24) and (25)). This *sequential* strategy implies that: (a) the ROM is constructed, (b) a number of micro-strain fluctuation modes n_ε is considered to find the basis, and (c) this selected basis, with n_ε modes, is used to built the energy snapshots (which, consequently, depends on the selected basis). This is an algorithmically *consistent* strategy displaying, however, a practical flaw because step (c) has to be re-computed every time that n_ε is changed (for instance to increase the HPROM accuracy or speedup).
2. Construct the free energy snapshots by collecting solutions of the HF model in PROBLEM II (Eqs. (18) and (19)). This *simultaneous* strategy computes the free energy snapshots concurrently with the fluctuating strain snapshots (both on the basis of the HF problem) the resulting free energy basis being independent of the chosen micro-strain fluctuation modes. The number of modes of the basis can be changed arbitrarily, independent from each other, to modify the resulting HPROM accuracy or speed-up, with no additional calculations. This is the strategy chosen in this work. Of course, consistency (convergence to the HF solution) is achieved when the number of modes of *both bases* is increased.

Therefore, for different sampling trajectories, a series of q_{snp} snapshots of the free energy, φ_μ, are evaluated and collected for each Gauss point. Then, the free energy snapshot matrix is built as:

$$[\boldsymbol{X}^\varphi] = [\boldsymbol{X}^\varphi_1, \ldots, \boldsymbol{X}^\varphi_{q_{snp}}]; \qquad \text{with: } [\boldsymbol{X}^\varphi_i] = \begin{bmatrix} \varphi_\mu(\mathbf{y}_1) \\ \varphi_\mu(\mathbf{y}_2) \\ \vdots \\ \varphi_\mu(\mathbf{y}_{N_g}) \end{bmatrix}_i \in \mathbb{R}^{N_g} \tag{60}$$

[11] Sensitivity analyses of the integration errors, with respect to the number of energy bases n_φ, are presented in Section 5.

where φ_μ is evaluated, in each Gauss point $\boldsymbol{y}_p$ ($p = 1, \ldots, N_g$), with the reduced micro-strain fluctuations: $\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_p) = \boldsymbol{\Psi}(\boldsymbol{y}_p)\boldsymbol{c}$, and the internal variable d_μ satisfying the constitutive relation for that micro-strain fluctuation:

$$\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\boldsymbol{y}_p) = \frac{1}{2}(1-d_\mu)\Big[(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_p))\cdot\mathbb{C}\cdot(\boldsymbol{\varepsilon}+\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}_p))\Big]. \tag{61}$$

In accordance with the position of the Gauss point, $\boldsymbol{y}_p$, in the original finite element mesh (HF model) and the ordering given to the snapshot matrix $[\boldsymbol{X}^\varphi]$, following a similar procedure to that adopted in expression (37), this matrix can be partitioned in components associated with the domains: $\mathcal{B}_{\mu,\text{coh}}$ and $\mathcal{B}_{\mu,\text{reg}}$:

$$[\boldsymbol{X}^\varphi] = \begin{bmatrix} \boldsymbol{X}^\varphi_{\text{reg}} \\ \boldsymbol{X}^\varphi_{\text{coh}} \end{bmatrix} \tag{62}$$

and the SVD technique is then separately applied to both partitions of $[\boldsymbol{X}^\varphi]$ to obtain two separated (orthogonal) bases:

$$\boldsymbol{\Phi}_{\text{G,reg}} \leftarrow \text{SVD}[\boldsymbol{X}^\varphi_{\text{reg}}]; \tag{63}$$

$$\boldsymbol{\Phi}_{\text{G,coh}} \leftarrow \text{SVD}[\boldsymbol{X}^\varphi_{\text{coh}}], \tag{64}$$

which are subsequently used, through the algorithm in Eq. (55), to obtain the positions and weights of the ROQ scheme.

3.5. Evaluation of the homogenized stresses and effective constitutive tensor

By deriving Eq. (57), with respect to the macro-strain argument, $\boldsymbol{\varepsilon}$, and accounting for Eqs. (59) and (16) we obtain:

$$\begin{aligned} \frac{\partial}{\partial\boldsymbol{\varepsilon}}\int_{\mathcal{B}_\mu}\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},\boldsymbol{c}))\,d\mathcal{B} &= \int_{\mathcal{B}_\mu}\underbrace{\frac{\partial}{\partial\boldsymbol{\varepsilon}}\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y},\boldsymbol{c}))}_{\boldsymbol{\sigma}_\mu(\boldsymbol{y},\boldsymbol{c})}\,d\mathcal{B} \\ &\simeq \sum_{j=1}^{N_r}\underbrace{\frac{\partial}{\partial\boldsymbol{\varepsilon}}\varphi_{\mu[\boldsymbol{\varepsilon},d_\mu]}(\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{z}_j,\boldsymbol{c}))}_{\boldsymbol{\sigma}_\mu(\boldsymbol{z}_j,\boldsymbol{c})}\,\omega_j = \sum_{j=1}^{N_r}\boldsymbol{\sigma}_\mu(\boldsymbol{z}_j,\boldsymbol{c})\,\omega_j, \end{aligned} \tag{65}$$

where, additionally, the evolution equation (19) is assumed to be satisfied in each quadrature point $\boldsymbol{z}_j$:

$$\dot{d}_\mu(\boldsymbol{z}_j,\boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu(\boldsymbol{z}_j)), \qquad \forall\ j = 1,\ldots,N_r. \tag{66}$$

Then, from Eqs. (65), (66) and (80) (defined in Appendix B), the following rule can be utilized to approach the homogenized stress:

$$\boldsymbol{\sigma} = \frac{1}{|\mathcal{B}_\mu|}\int_{\mathcal{B}_\mu}\boldsymbol{\sigma}_\mu(\boldsymbol{y},\boldsymbol{c})\,d\mathcal{B} \approx \frac{1}{|\mathcal{B}_\mu|}\sum_{j=1}^{N_r}\boldsymbol{\sigma}_\mu(\boldsymbol{z}_j,\boldsymbol{c})\,\omega_j. \tag{67}$$

Notice that evaluation of the homogenized values in Eq. (67) involves only the reduced number of quadrature points, N_r, where the micro-stresses are sampled. A similar development can be utilized for evaluating the homogenized constitutive tensor given in Eq. (75) (defined in Appendix B), yielding:

$$\boldsymbol{C} \simeq \frac{1}{|\mathcal{B}_\mu|}\sum_{j=1}^{N_r}\boldsymbol{C}_\mu(\boldsymbol{z}_j,\boldsymbol{c})\Big(\mathbb{I}+\mathcal{A}_\mu(\boldsymbol{z}_j)\Big)\,\omega_j. \tag{68}$$

4. RVE off-line sampling program design

In accordance with the concepts introduced in Section 3.2, the snapshot matrix, $[\boldsymbol{X}]$, of $\tilde{\boldsymbol{\varepsilon}}_\mu$, and the posterior POD basis evaluation, are obtained from a sampling program of the RVE high-fidelity (HF) finite element model by

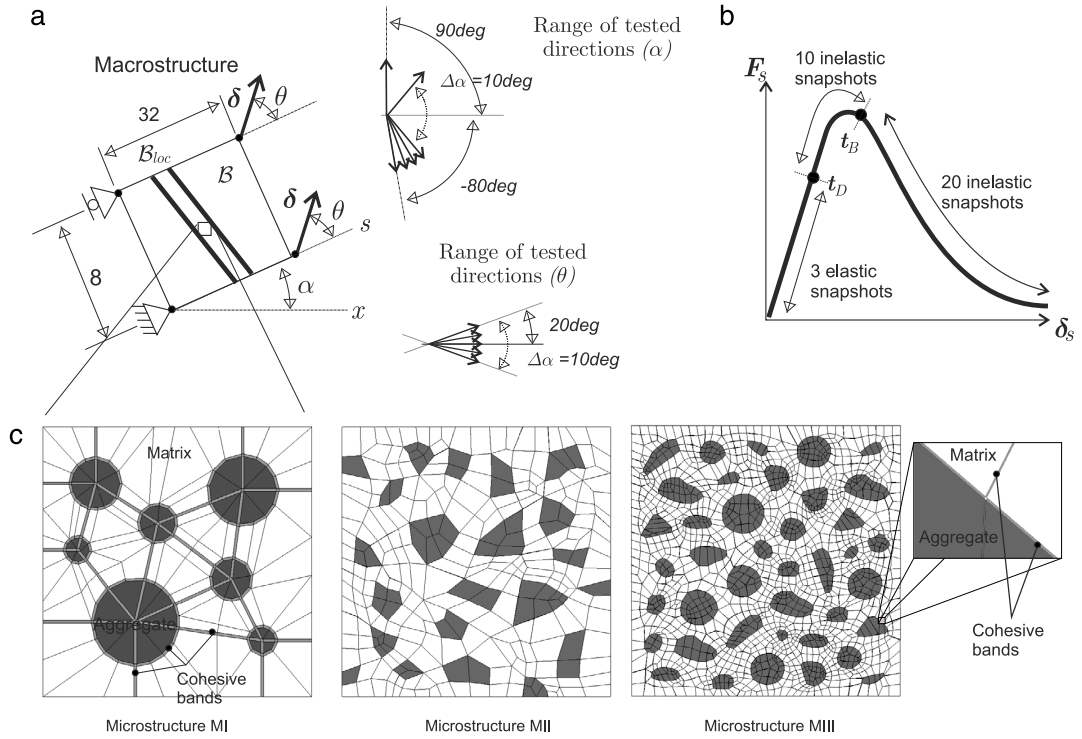


Fig. 3. Parametrization of the sampling trajectories. (a) Loading histories at the macro-scale defined by angles α and θ. Finite element model of the micro-cells with a detail showing the cohesive bands for capturing the possible macroscopic failure mechanisms. (b) Selection of elastic and inelastic snapshots during a sampled trajectory represented by the macrostructural response: loading component in the axial strip direction $\boldsymbol{F}_s$ versus axial displacement component $\boldsymbol{\delta}_s$. (c) Microstructures MI, MII and MIII representing concrete, three phases are modeled: matrix, aggregates and interfaces matrix–matrix and matrix–aggregates. The interfaces are simulated with cohesive bands.

gathering solutions from different trial cases. Each trial case corresponds to a micro-cell problem which is subjected to a specific loading condition, and from each loading history, several solutions of $\tilde{\boldsymbol{\varepsilon}}_\mu$ and energies φ_μ (snapshots) are taken. The success for constructing suitable POD basis, for spanning $\tilde{\boldsymbol{\varepsilon}}_\mu$ and φ_μ, lies on gathering the most appropriate (discrete) solutions representing the space of solutions.

In the micro-cell VBVP, the input parameters, or driving forces, defining the loading history are the components of the macro-strains $\boldsymbol{\varepsilon}$. In plane problems, where the macro-strain component orthogonal to the plane is assumed zero ($\boldsymbol{\varepsilon}_{zz} = 0$),[12] these parameters live in a space of dimension 3 ($\{\boldsymbol{\varepsilon}_{xx}, \boldsymbol{\varepsilon}_{yy}, \boldsymbol{\varepsilon}_{xy}\}$). This space can be typically parameterized using the three components of the strain tensor in a canonical basis.

The exploration of the micro-cell (HF) finite element model is performed using the following strategy:

1. Multiscale tests described in Fig. 3(a) are solved.
2. A macroscale strip is considered as a simple structure to, indirectly, induce in the RVE the complex macro-strain path leading to material instability. The strip is oriented along the direction s, forming an angle, α, with the horizontal axis x.
3. This strip is stretched by imposing, on one side of the strip, uniform displacements $\boldsymbol{\delta}$ forming an angle θ with respect to the local axis s.
4. Displacements $\boldsymbol{\delta}$ are then monotonically increased until getting the full structural degradation. Different test cases are performed by sweeping the angle α in the interval: $\alpha \in [-80, 90]deg$ (every $\Delta\alpha = 10.deg$) and the angle θ in the interval: $\theta \in [-20, 20]deg$ (every $\Delta\theta = 10.deg$).

[12] In the conventional coordinate system $\{x, y, z\}$, with $\{x, y\}$ defining the plane of analysis.

5. In this way, 90 structural trajectories (every one with 400 loading steps for increasing values of $\boldsymbol{\delta}$, until failure) are performed. A selection of the corresponding micro-strain fluctuations $\tilde{\boldsymbol{\varepsilon}}_\mu$, and free energy, φ_μ, snapshots are stored to configure the corresponding snapshot matrices (see Fig. 3(b)) as explained in Section 4.1

The strip is modeled with only one EFEM finite element. At some loading level, the formation of one macro-crack band $\mathcal{B}_{loc}$, is captured by the EFEM element. For different tests, and depending on $\boldsymbol{\delta}$, the cracks would display different directions.

The micro-cell related to the central singular Gauss point of the EFEM element is surveyed. And the snapshot matrix $[\boldsymbol{X}]$ is built accordingly with the responses obtained in this surveyed cell during the sampling process. The square micro-cell is designed with sides parallel to the horizontal and vertical axis.

Note that for every α, a uniaxial stretching is simulated when $\theta = 0$.

With the devised sampling program, a wide spectrum of representative solutions of the problem, displaying macro-cracks, induced by a large number of different micro-cell failure modes, can be captured.

4.1. Methodology for sampling snapshots in given trajectories

In the examples presented below, the following methodology has been used for sampling snapshots along the sampling trajectories. Let us consider a given sampling trajectory, and the corresponding macroscopic structural response in terms of the force–displacement, e.g. the one sketched in Fig. 3(b). A total of 33 snapshots are collected (for each of the micro strain-fluctuation, $\tilde{\boldsymbol{\varepsilon}}_\mu$, and free energy, φ_μ, fields) along three different stages of the trajectory:

1. The *elastic regime* (when no dissipation has taken place yet),[13] which ends when some zero-dissipation threshold is overcome at time t_D, so that the elastic stage is characterized by the time interval $[0, t_D]$. Three (approximately equidistant) snapshots are taken in this stage, contributing to the matrices $[\boldsymbol{X}^E]$ and $[\boldsymbol{X}^{\varphi,E}]$ in Eqs. (36) and (60), respectively.
2. The *inelastic-stable regime*, characterized by the time interval $[t_D, t_B]$, where t_B stands for the *bifurcation time*,[14] which triggers the onset of a macro-crack at the macro-scale. Ten, equidistant snapshots are taken at his stage, contributing to the snapshot matrices $[\boldsymbol{X}^I]$ and $[\boldsymbol{X}^{\varphi,I}]$, in Eqs. (36) and (60), respectively.
3. The *inelastic-unestable regim*e, when the material is propagating a crack at the macro-scale $t > t_B$. Twenty snapshots are taken at this stage, contributing to the snapshot matrices $[\boldsymbol{X}^I]$ and $[\boldsymbol{X}^{\varphi,I}]$, respectively.

In turn, every of those snapshots is allocated to the regular, $(\bullet)_{\text{reg}}$, and cohesive $(\bullet)_{\text{coh}}$ blocks in Eqs. (37) and (60), respectively, depending on the spatial position of the corresponding sampling points.

4.2. Sampled RVE's

Three squared micro-cell models, made of a matrix and aggregates, are devised and sampled see (Fig. 3(a)) to simulate the meso-structure of a cementitious like material (concrete). They are of increasing complexity, with finite element details shown in Table 1. Trying to mimic the concrete material response, the cells are modeled with three components: *aggregates*, which are assumed elastic, *bulk matrix*, which is assumed elastic, and *interfaces* (matrix–matrix and matrix–aggregates), which are simulated with banded elements and modeled with the isotropic damage constitutive law. The properties of the materials in the micro-cell are defined in Table 2.

Typical deformed meshes of Microstructure III are depicted in Fig. 4. These deformed meshes have been obtained by the superposition of the micro-displacement fluctuations to the original configuration. The micro-displacement fluctuation of each picture is in correspondence with one of the micro-strain modes of the basis $\{\boldsymbol{\Psi}^E\}$ or $\{\boldsymbol{\Psi}^I\}$ applying the POD procedure to the snapshot matrix $[\boldsymbol{X}]$ which has been constructed using the methodology above described. Given one mode $\boldsymbol{\Psi}_i$, the procedure for evaluating the micro-displacements fluctuations at the RVE has been described in Appendix C.

[13] A procedure to obtain the macro-scale dissipation from the micro-scale one is provided in [8].

[14] The bifurcation time t_B is characterized by the loss of ellipticity of the localization tensor, obtained in terms of the homogenized tangent constitutive tensor, given in Eq. (75) (see [28] for numerical algorithms for detecting this condition).

Table 1
Finite element models for Microstructures MI, MII and MIII.

Microstructure	Number of FE	Number of D.o.f.'s	Number of cohesive bands	Total number of Gauss points (N_g)
M I	311	668	145	1244
M II	445	1036	161	1780
M III	5409	14256	2189	21636

Table 2
Material properties of the three sampled micro-cells. MI, MII and MIII refer to the micro-cells denoted: Microstructure I, II and II respectively in Fig. 3. Properties are: E_μ (Young's modulus), ν_μ (Poisson ratio), $\sigma_{\mu u}$ (ultimate tensile stress) and $G_{\mu,f}$ (fracture energy).

	Model	E_μ [MPa]	ν_μ	$\sigma_{\mu u}$[MPa]	$G_{\mu f}$[N/m]
Elastic matrix	MI	1.85e4	0.18	–	–
	MII	1.85e4	0.18	–	–
	MIII	1.85e4	0.18	–	–
Elastic aggregate	MI	3.70e4	0.18	–	–
	MII	3.70e4	0.18	–	–
	MIII	3.70e4	0.18	–	–
Cohesive bands of matrix–matrix interface	MI	1.85e4	0.18	2.60	140
	MII	1.85e4	0.18	2.60	140
	MIII	1.85e4	0.18	2.60	140
Cohesive bands of matrix–aggregate interface	MI	1.85e4	0.18	1.30	70
	MII	1.85e4	0.18	–	–
	MIII	1.85e4	0.18	–	–

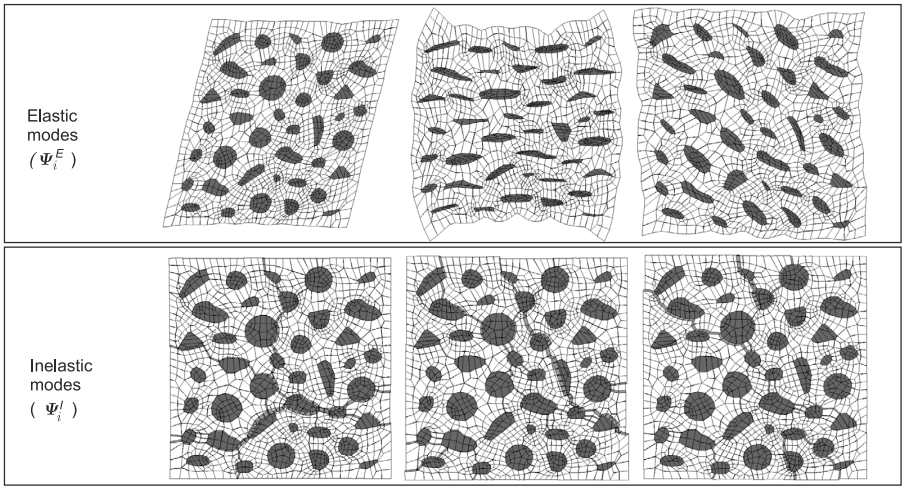


Fig. 4. Typical deformed meshes for micro-strain modes in microstructure MIII. The micro-displacements, required for post-processing, are obtained with the recovery procedure in Appendix C. The amplification factors are different for elastic and inelastic modes. Active cohesive bands, in inelastic modes, display unstructured crack opening.

5. Numerical assessment

The accuracy of the ROM model mainly depends on the number of modes, n_ε, taken in the basis $\{\boldsymbol{\Psi}\}$ spanning the micro-strain fluctuation field. Additionally, the accuracy of the HPROM model also depends on the number of quadrature points, N_r, adopted in the ROQ scheme. Accuracy of the combined HPROM strategy depends, then, on

both parameters. In this section the issue is addressed to evaluate sensitivity of the committed errors with respect to those parameters, in different scenarios. Three different kinds of tests are presented:

(i) *Consistency tests*. A number of trajectories, already solved with the HF model during the micro-cell sampling process in the off-line stage, are re-evaluated using the ROM and HPROM models. Errors of both approaches with respect to the HF solutions are analyzed. This kind of assessment provides an estimation of the quality and richness of basis to reproduce actual snapshots, $\boldsymbol{X}_i$, and the accuracy of the ROQ scheme. It is expected that the error with respect to the HF solutions *(consistency-error)* tends to zero as the number of considered modes for, both, the micro strain-fluctuation, n_ε and N_r increase.

(ii) *Accuracy tests*. Similar to the aforementioned consistency tests, but for the RVE subjected to a *random loading trajectory* (different from any previously sampled trajectory in the snapshot matrices). These tests also evaluate the richness of the reduced bases but, in this case, HF solutions not tested during the sampling process are not expected to be captured exactly, because of a remaining *sampling error*.

(iii) A *multiscale structural test.* With the objective of evaluating convergence of the solutions when the proposed overall HPROM strategy is applied to a complex structural crack propagation problem, the classical benchmark of crack propagation in a concrete L-shaped specimen is reproduced. Solution convergence, in terms of the increasing number of quadrature points defining the ROQ scheme, is studied.

5.1. Consistency tests

Three types of errors are involved in consistency tests. First, by considering that the POD procedure, to select the matrix of modes $\boldsymbol{\Psi}_G$, neglects modes associated with singular values smaller than a given limit, then, $\boldsymbol{\Psi}_G$ cannot expand completely the snapshot set defined by $[\boldsymbol{X}]$. This error is identified as the *a-priori error* of the ROM model and can be determined in the off-line stage, without constructing the ROM model. A second source of error is associated with the snapshot sampling technique, in the sense that only a few snapshots, of each sampled trajectory, are taken to build $[\boldsymbol{X}]$. So, the ROM model reproducing sampled trajectories has associated an error, identified as the *a-posteriori error*, which is the combination of the a-priori and the sampling errors. Finally, the HPROM model has also associated an additional source of error which is governed by the reduced quadrature rule.

In this section first the a-priori and a-posteriori consistency errors of the ROM model, through solutions obtained with Microstructures MI, II and III, are studied. Then, the convergence properties of the HPROM, in consistency tests, are analyzed.

5.1.1. A-priori (off-line) consistency error of the ROM model

Once obtained the POD basis, $\{\boldsymbol{\Psi}\}$, for micro-strain fluctuations, the a-priori (percentage) error in approaching every snapshot $\boldsymbol{X}_i$, with $i = 1, \ldots, p_{snp}$, can be estimated through:

$$\text{error}_i\% = \frac{\|\boldsymbol{X}_i - \overbrace{\boldsymbol{\Psi}_G \boldsymbol{c}^i}^{\boldsymbol{X}_i^\Psi}\|}{\|\boldsymbol{X}_i\|} 100; \qquad \boldsymbol{c}^i = \boldsymbol{\Psi}_G^T \boldsymbol{X}_i \tag{69}$$

where $\boldsymbol{c}^i$ are coefficients of the vector $\boldsymbol{X}_i^\Psi$ (the orthogonal projection of $\boldsymbol{X}_i$ onto the space spanned by $\{\boldsymbol{\Psi}\}$). Clearly, if the space spanned by $\{\boldsymbol{\Psi}\}$ is rich enough to include $\boldsymbol{X}_i$, then, the error (69) is zero. Then, the maximum error:

$$\text{Error} = \max_i \Big(\text{error}_i\%\Big); \; i = 1, \ldots, p_{snp} \tag{70}$$

identifies the richness of $\{\boldsymbol{\Psi}\}$ for capturing all the snapshots in the matrix $[\boldsymbol{X}]$.

In accordance with the snapshot decomposition in Eq. (37), the POD bases $\{\boldsymbol{\Psi}\}$ and $\{\boldsymbol{\Phi}\}$, for micro-strain fluctuations and energies, respectively, are partitioned into two domains: $\{\boldsymbol{\Psi}_{reg}\}$, $\{\boldsymbol{\Psi}_{coh}\}$ and $\{\boldsymbol{\Phi}_{reg}\}$, $\{\boldsymbol{\Phi}_{coh}\}$. Then, using expression (69) and (70) for every partition of $\{\boldsymbol{\Psi}\}$ and $\{\boldsymbol{\Phi}\}$, the errors: $\text{Error}^{\Psi}_{reg}$ and $\text{Error}^{\Psi}_{coh}$, for the regular and cohesive band domains, respectively, of the micro-strain basis $\{\boldsymbol{\Psi}\}$, and $\text{Error}^{\Phi}_{reg}$ and $\text{Error}^{\Phi}_{coh}$, for the regular and cohesive band domains, respectively, of the energy basis $\{\boldsymbol{\Phi}\}$ are computed. Fig. 5 plots these errors. There it can be checked that the HPROM is fully consistent (zero errors are achieved for large enough values of n_ϵ and n_φ).

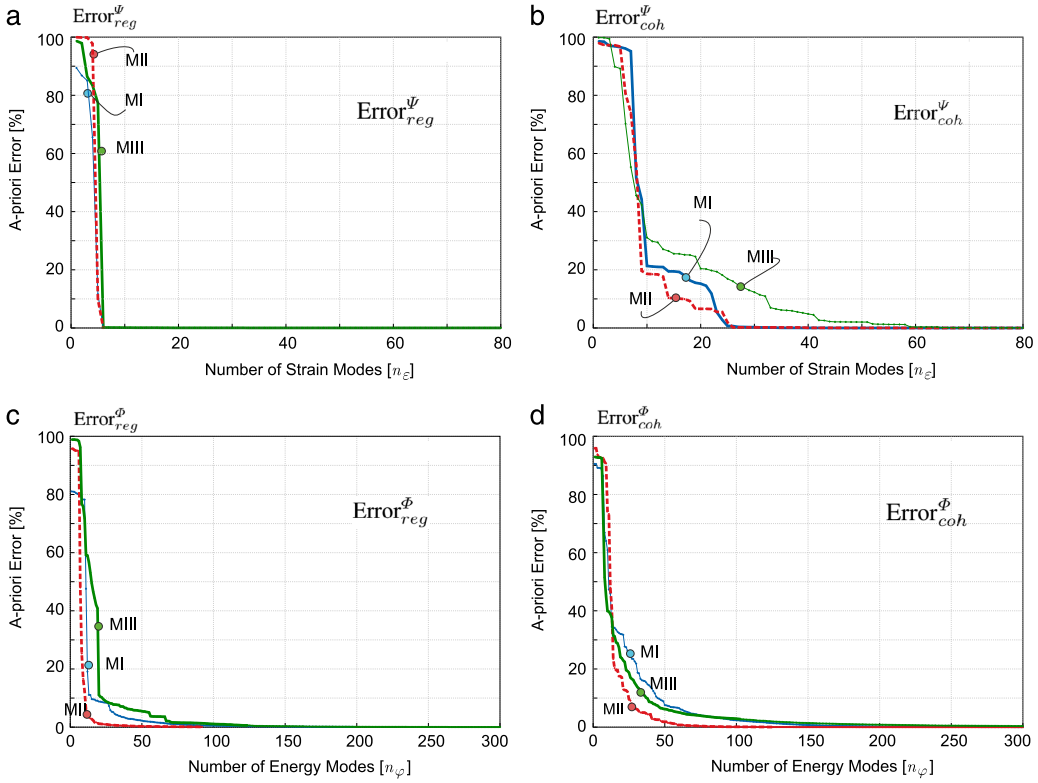


Fig. 5. Consistency test. A-priori consistency error for microstructures MI, MII and MIII ($\{\boldsymbol{\Psi}\}$ is the basis for the micro-strain fluctuation space and $\{\boldsymbol{\Phi}\}$ is the basis for the energy space): (a) $\text{Error}^{\Psi}_{reg}$, (b) $\text{Error}^{\Psi}_{coh}$ (c) $\text{Error}^{\Phi}_{reg}$ and (d) $\text{Error}^{\Phi}_{coh}$.

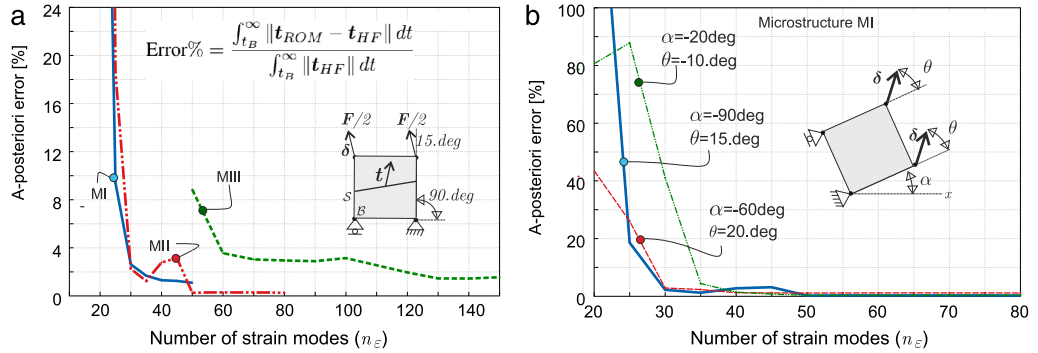


Fig. 6. Consistency test. Analysis of results obtained with the ROM model. Error of the homogenized traction-vector norm vs. number of modes ($\boldsymbol{t}_{ROM}$ and $\boldsymbol{t}_{HF}$ are the traction-vectors obtained with the ROM and HF models, respectively). (a) Results obtained with microstructures MI, MII and MIII for the sampled trajectory characterized by $\alpha = 90.deg$ and $\theta = 15.deg$. (b) Results obtained with microstructure MII for three different sampled trajectories.

5.1.2. A-posteriori consistency error of the ROM model

Fig. 6 plots the error curves for evaluation of the homogenized traction vector norm, at the macro-scale, versus the number of modes (n_ε) defining the basis $\{\boldsymbol{\Psi}\}$ of the reduced micro-strain fluctuation space. The relative error is

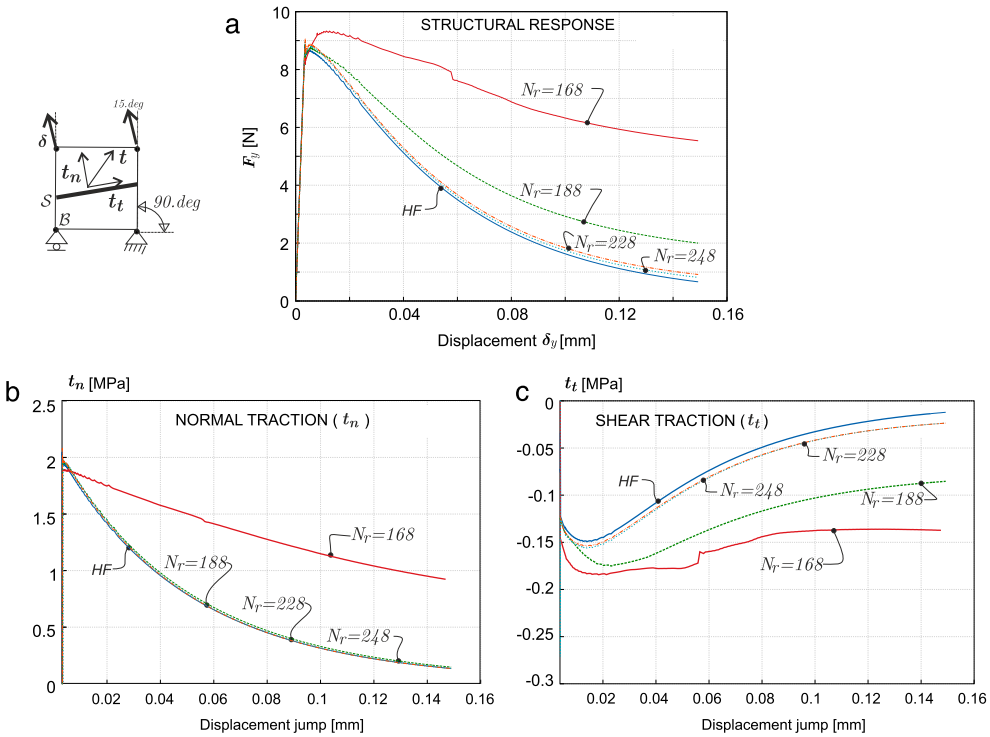


Fig. 7. Accuracy test for microstructure MIII with 21 636 Gauss points in the original HF model. Analysis of results obtained with the HPROM model in terms of the number of quadrature points (N_r) of the ROQ scheme (with $n_\varepsilon = 106$). (a) Vertical components of the reaction force ($\boldsymbol{F}_y$) vs. imposed displacement ($\boldsymbol{\delta}_y$) at macro-scale (HF curve refers to the High-Fidelity finite element model); (b) normal component of homogenized traction vector vs. displacement jump at the macro-scale (in $\mathcal{S}$); (c) shear component of the homogenized traction vector vs. displacement jump at the macro-scale.

given by:

$$\text{Error}\% = \frac{\int_{t_B}^{\infty} \|\boldsymbol{t}_{ROM} - \boldsymbol{t}_{HF}\|\, dt}{\int_{t_B}^{\infty} \|\boldsymbol{t}_{HF}\|\, dt} \tag{71}$$

where $\boldsymbol{t}_{ROM}$ and $\boldsymbol{t}_{HF}$ are the traction-vectors determined with the ROM and HF models, and t_B stands for the bifurcation time. Fig. 6(a) compares the a-posteriori errors obtained with the microstructures MI, MII and MII by simulating one sampled trajectory. While, Fig. 6(b) compares the errors obtained with microstructure MII by simulating three different sampled trajectories.

Results in Fig. 6 corroborate the consistency of the ROM approach, i.e. the convergence of the reduced model (ROM) solutions toward the HF solution by increasing the number n_ε of micro-strain fluctuation modes. It is interesting to note that, in all different cases, it can be identified a minimum value of n_ε, above which the error keeps small and constant. Also, according to these results, a residual error of the order of 1%–2% remains, even for large values of n_ε. This residual error is associated with the *sampling error*, which could be diminished by increasing the number of snapshots taken from every trajectory.

5.2. HPROM solutions for non-sampled trajectories

Fig. 7(a) plots the macrostructural responses provided by the HPROM model in microstructure MIII, with an increasing number of quadrature points N_r. The simulated trajectory is characterized by the angles $\alpha = 90.deg$ and $\theta = 15.deg$.

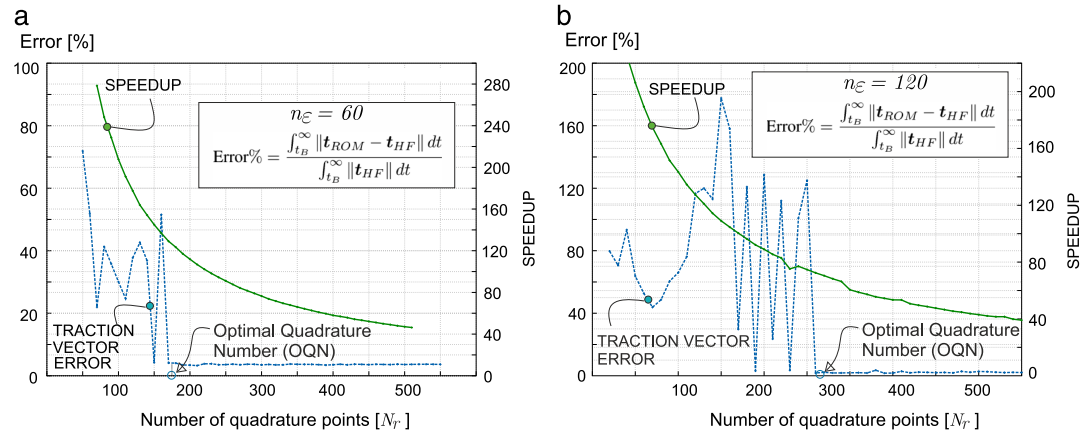


Fig. 8. Accuracy test for the microstructure MIII with 21 636 Gauss points in the original HF model. Analysis of results obtained with the HPROM model in terms of the number of quadrature points (N_r) of the reduced integration scheme with $n_\varepsilon = 106$. (a) Vertical components of the reaction force ($\boldsymbol{F}_y$) vs. imposed displacement ($\boldsymbol{\delta}_y$) at macro-scale; (b) normal component of homogenized traction vector vs. $\boldsymbol{\delta}_y$; (c) shear component of the homogenized traction vector vs. $\boldsymbol{\delta}_y$. The *Optimal Quadrature Number* (OQN) is shown in both cases.

The macrostructural response is given in terms of the vertical component of the reaction force $\boldsymbol{F}_y$ versus the vertical component of the imposed displacement $\boldsymbol{\delta}_y$. Convergence toward the HF solution with increasing N_r is observed. Similarly, Fig. 7(b) plots the normal component of the macro-scale homogenized traction vector $\boldsymbol{t}_n$ for increasing N_r, and Fig. 7(c) plots the shear component of the homogenized traction-vector for increasing N_r. The traction-vector, $\boldsymbol{t}$, governs the post-critical (post-bifurcation) macro-scale response.

Results in Fig. 7 are studied by selecting a fixed number of micro-strain fluctuation modes, $n_\varepsilon = 106$. Therefore, these results depict the sensitivity of the HPROM solutions in terms of the number of quadrature points N_r.

The plots in Fig. 8 display the error of the homogenized macro traction vector evaluated with the HPROM approach, obtained with $n_\varepsilon = 60$ and $n_\varepsilon = 120$ micro-strain modes, by varying the number of quadrature points N_r. In the same plots, the computational speed-up (the inverse of the ratio of computational times), with respect to the HF case is also depicted. It can be observed the error behavior is very noisy for low values of N_r.

However, by selecting an enough high value of N_r, which depends on the number of modes n_ε, the HPROM approach error holds small and almost constant, indicating that there exist a minimum number of quadrature points, *the Optimal Quadrature Number* (OQN), that must be taken to get an enough accurate solution. This threshold is also a maximum, since significant increase of accuracy is not obtained with higher values.

Fig. 9 displays a set of pictures that help to understand the reason for such response. They show the deformed RVE (obtained through the displacement recovery method in Appendix C) displaying the failure mode obtained for $n_\varepsilon = 100$ and $n_\varepsilon = 140$ and increasing number of quadrature points N_r. For comparison, it is also shown the HF failure mode for the simulated trajectory.

There, it can be observed the effects of increasing values of the HPROM parameters, n_ε and N_r, on capturing the correct failure mode (crack path) at the RVE. When this crack path is exactly fitted i.e. for values equal to or larger than the corresponding OQN, the job is already done and larger values of N_r do not translate into additional accuracy.

5.3. Design of the HPROM strategy

Fig. 10 shows the summary of a number of results obtained by running the HPROM strategy in a large number of cases for microstructure MIII.

The plots can be used as an "abacus" for a-priori design by the user of the HPROM strategy in a multiscale problem (for a given microstructure at the RVE). By selecting the admissible error, in the top figure, (say error $= 3.5\%$), the number of strain modes, ($n_\varepsilon = 80$) is obtained. Entering in the lower plot, with this result, ($n_\varepsilon = 80$), one obtains the suitable number of integration points ($OQN \simeq 200$) and the speedup that can be expected ($speedup \simeq 110$).

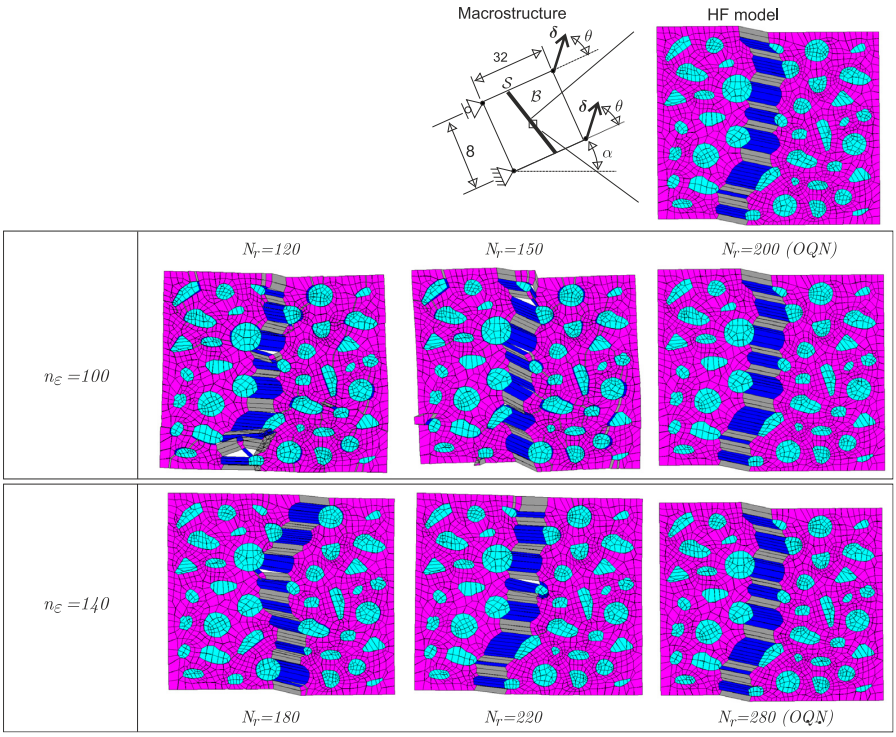


Fig. 9. Non-sampled trajectory defined by $\alpha = -2.5deg$ and $\theta = -15.deg$. Capture of failure modes with HPROM, by increasing the number of quadrature points N_r and micro-strain modes n_ε, in comparison with the failure mode captured with the High-Fidelity (HF) finite element model.

The availability of such an abacus (constructed off-line) for a specific RVE microstructure, allows the user's selection of the most appropriate HPROM strategy, by balancing the admissible error vs. the desired speedup, for the on-line multiscale computation stage.

5.4. Multiscale crack propagation problem: L-shaped panel

The panel in Fig. 11 is a benchmark commonly used for testing propagating fracture models. This concrete specimen is considered here to test the qualitative results and convergence properties of the proposed HPROM approach, when utilized in real FE^2 multiscale crack propagation problems.

The geometry of the simulated specimen is depicted in Fig. 11(a). As shown in Fig. 11(b), the domain of the L-shaped panel is split into two domains: (1) the multiscale domain (with 721 finite elements) corresponding to the region where the crack may propagate, which is modeled with HPROM of the Microstructure MIII, depicted in Fig. 3, and (2) the remaining part of the panel, which is modeled with an elastic monoscale approach (using 1709 finite elements), where the elasticity tensor is obtained through an *elastic homogenization* of the micro-structure elastic properties. Even for this (rather coarse) multiscale problem the high fidelity (HF) computational solution is extremely costly to handle, till the point that, with the available computational resources,[15] it has not been possible to display the complete action–response curve (in Fig. 11(d)).

However, the remaining structural responses in Fig. 11(d), obtained through a number of HPROM strategies, involve very reasonable computational costs, and they were obtained in advance *with no previous knowledge of the HF*results. The accuracies are very good, and a response indistinguishable from the HF can be obtained 60 times

[15] A cluster of 500 cores, of recent generation, is used. The multiscale finite element code is written in a Matlab© environment.

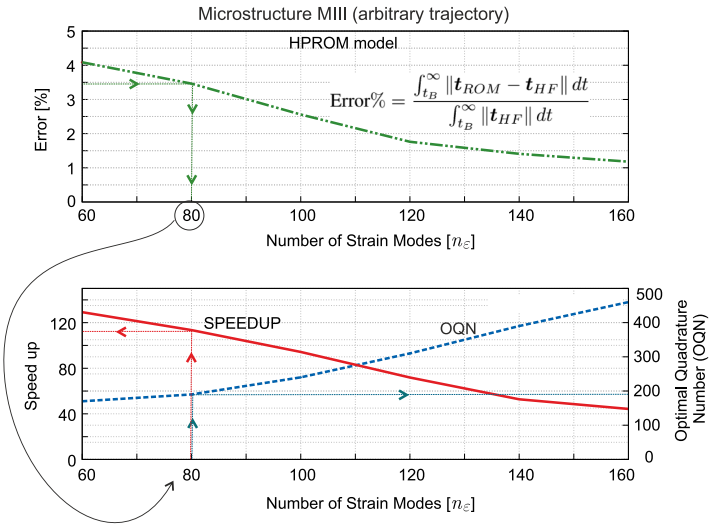


Fig. 10. HPROM design diagrams. *Top*: HPROM error in terms of number of strain modes. *Bottom:* OQN and obtained speedup in terms of the number of strain modes. By selecting the admissible error (say 3.5%) in the upper diagram, one obtains the requested number of strain modes, $n_\varepsilon = 80$. Entering with this result in the lower diagram one obtains the suitable number of integration points ($TQN = 200$) and the resulting speedup ($speedup = 110$).

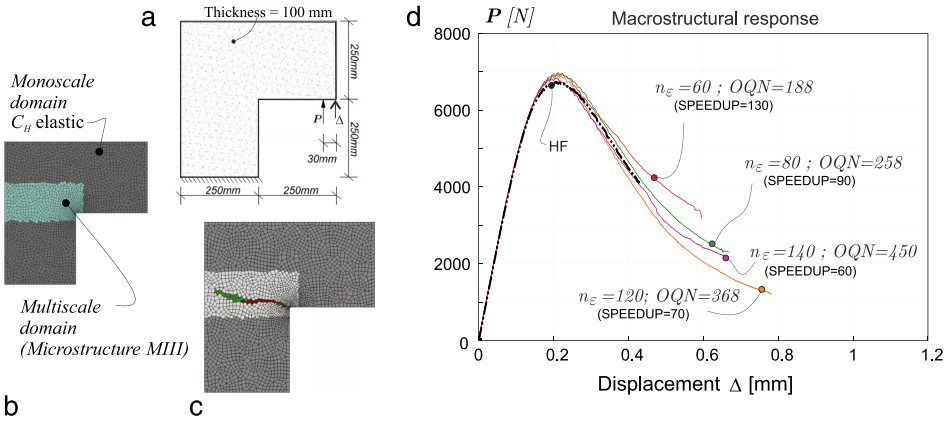


Fig. 11. L-shaped panel. (a) Specimen geometry; (b) finite element mesh; (c) macro-scale obtained crack path (red and green); d) structural responses in terms of force P vs. vertical displacement Δ, for different RVE HPROM strategies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

faster ($speedup = 60$). A less accurate response, but with a fairly good agreement with the HF can be obtained with $speedup = 130$. This illustrates the new paradigm and computational possibilities open by HPROM strategies in computational multiscale modeling.

6. Concluding remarks

Along this paper a new computational strategy for developing hyper-reduced order modeling (HPROM) of multiscale fracture problems has been presented.

The proposed strategy consists of a series of new techniques, here developed and used to explore up to what extent multiscale FE^2 computational multiscale modeling of fracture can be made affordable in terms of computational

cost. It essentially consists of using a standard two-stage (off-line/on-line) strategy to produce a HyPer Reduced Order Model (HPROM) with some specific techniques to exploit the computational-cost reduction.

Several aspects of the proposed methodology can be highlighted as new contributions in the literature:

- The RVE *domain separation technique*, to account for distinct constitutive models used at the RVE and take the maximum advantage of this distinction.
- A *strain-based formulation of the variational RVE problem* allowing a simpler application of the previous technique, *without the need of introducing compatibility constraints*.
- A specific *sampling program*, for the construction of the sets of snapshots in the off-line stage of the HPROM procedure, in accordance with the rest of elements of the proposed strategy.
- The *Reduced Optimal Quadrature* (ROQ) technique, an evolution of some recent HPROM methods [16,17], which resorts to the primitive formulation of the RVE problem as a saddle-point problem. This allows identifying the involved functional as the appropriate (scalar) entity for designing an optimal reduced-cost integration rule, *instead of the internal forces (vector-entities)* obtained from the corresponding variational principle.

The resulting methodology has been extensively tested, both at the RVE level and in multiscale modeling of propagating fracture benchmarks, this being totally unconventional for fracturing problems. The obtained results, in terms of error vs. computational speedup, are certainly encouraging. Speedups of two orders of magnitude, much higher than the ones currently reported in the literature for the same type of problems, have been obtained.

At this point it can be argued that only idealized, two-dimensional, problems have been considered. The real interest of many multiscale modeling problems residing on actual three-dimensional problems, the following question arises: to what extent this "encouraging" results can be extended to three-dimensional problems, where the involved RVE complexity and the associated computational cost can be two or three orders of magnitudes larger? The answer can be stated in two parts: (1) the proposed HPROM strategy methodology can be conceptually extended from 2D to 3D cases without fundamental changes, and difficulties are not envisaged for this purpose, and (2) the obtained results in terms of speedup can be understood as fully scalable from 2D to 3D problems. This means that one could think of achievable values of 10^4–10^5 for the speedups in 3D problems... and this fact would turn the 3D multiscale case into affordable (in conjunction with, currently available, HPC procedures). This would provide additional scientific and industrial credit to the multiscale modeling concept. Although, this has to be proved in forthcoming works.

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement No. 320815 (ERC Advanced Grant Project "Advanced tools for computational design of engineering materials" COMP-DES-MAT).

Funding received from the Spanish Ministry of Economy and Competitiveness through the National Research Plan 2014: MAT2014-60919-R is gratefully acknowledged.

Appendix A. Multiscale model for propagating fracture

A brief overview of the multiscale model for propagating fracture fully developed in [8] is here presented. The main interest in this contribution lies on the RVE model cost reduction and not on the multiscale problem itself. However, and in order to introduce an appropriate setting for this purpose, the considered multiscale fracture model is introduced here. This point is shown in Appendix A.2 of the Appendix, while the RVE finite element implementation, which is taken as a starting point for constructing the HPROM model, is presented in Appendix A.3.

A.1. Macroscopic model

A.1.1. Kinematics

Let us consider a regularized strong discontinuity kinematics at the macroscale. The strain field stemming from a regularized jump in $\mathcal{B}_{loc}$, the localization band with a finite width h (the regularization parameter), is introduced as follows:

$$\boldsymbol{\varepsilon}(\boldsymbol{x},t) = \begin{cases} \bar{\boldsymbol{\varepsilon}}(\boldsymbol{x},t); & \forall \boldsymbol{x} \in \mathcal{B}\backslash\mathcal{B}_{loc}\, \forall t \in [0,T]\,; \\ \bar{\boldsymbol{\varepsilon}}(\boldsymbol{x},t) + \dfrac{\chi}{h}\big(\boldsymbol{m}(t) \otimes^s \boldsymbol{n}(\boldsymbol{x})\big); & \forall \boldsymbol{x} \in \mathcal{B}_{loc}\, \forall t \in [0,T]\,; \end{cases} \tag{72}$$

where $\bar{\boldsymbol{\varepsilon}}$ is a smooth contribution to the strain field, and the term $\frac{\chi}{h}\big(\boldsymbol{m}(t)\otimes^s \boldsymbol{n}(\boldsymbol{x})\big)$ is the strain jump component with strength given by the factor χ/h. The unit vector $\boldsymbol{m}(t)$, the polarization direction, indicates the direction of the displacement jump across the band $\mathcal{B}_{loc}$, and $\boldsymbol{n}(\boldsymbol{x})$ is the unit normal vector to the band. Supra index $(\cdot)^s$ indicates the symmetric part of the second order tensor and $[0, T]$ is the time (or pseudo-time) interval of interest.

A.1.2. Equilibrium and constitutive equations

Specific details of the Boundary Value Problem at the macroscale domain, $\mathcal{B}$, are not necessary for the intended development of the micro-cell reduced model. The only relevant fact is that the macroscopic stresses, $\boldsymbol{\sigma}(\boldsymbol{x}, t)$, are point-wise obtained, in terms of the macroscopic strains, $\boldsymbol{\varepsilon}(\boldsymbol{x}, t)$, as

$$\boldsymbol{\sigma}(\boldsymbol{x}, t) = \Sigma\big(\boldsymbol{\varepsilon}(\boldsymbol{x}, t)\big) \tag{73}$$

where $\Sigma(\bullet)$, stands for the homogenized constitutive model obtained from the computational homogenization in the micro-cell.

After assuming the existence of a Representative Volume Element (RVE) and the variational Hill–Mandel Principle (see [29]), which introduces the connection between both scales, the constitutive law Σ is defined through the following sequence of operations: (i) first, the micro-stress field $\boldsymbol{\sigma}_\mu$, in $\mathcal{B}_\mu$, is determined by solving a standard variational boundary value problem (VBVP) defined in the RVE. The actions, or driving forces, of this VBVP are the macro-strains $\boldsymbol{\varepsilon}$, which are homogeneously distributed in $\mathcal{B}_\mu$. The boundary conditions are prescribed according to the rules defined in the next section, and (ii) the macro-stresses are computed as the volumetric average of $\boldsymbol{\sigma}_\mu$ in $\mathcal{B}_\mu$, as follows:

$$\boldsymbol{\sigma}(\boldsymbol{x}, t) = \Sigma\big(\boldsymbol{\varepsilon}(\boldsymbol{x}, t)\big) = \frac{1}{|\mathcal{B}_\mu|}\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\boldsymbol{x}, \boldsymbol{y}, t)\, d\mathcal{B}_\mu. \tag{74}$$

From this expression, the effective constitutive tensor: $\boldsymbol{C} = \partial\boldsymbol{\sigma}/\partial\boldsymbol{\varepsilon}$, can be written as:

$$\boldsymbol{C}(\boldsymbol{x}, t) = \frac{1}{|\mathcal{B}_\mu|}\int_{\mathcal{B}_\mu} \boldsymbol{C}_\mu(\boldsymbol{y})\Big(\mathbb{I} + \mathcal{A}_\mu(\boldsymbol{x}, \boldsymbol{y}, t)\Big)\, d\mathcal{B}_\mu \tag{75}$$

where $\mathbb{I}$ is the fourth order identity tensor and $\mathcal{A}_\mu$ is the conventional localization tensor.

A complete description of the insertion of the resulting homogenized macroscopic constitutive model into a propagating fracture scheme at the macro-scale can be found in [8].

A.2. Micro-cell model accounting for material failure

Let us consider a micro-cell $\mathcal{B}_\mu$ model accounting for material failure. The micro-strain $\boldsymbol{\varepsilon}_\mu$ in this RVE is given by expression (2), where the relevant term is the micro-strain fluctuation $\tilde{\boldsymbol{\varepsilon}}_\mu$, satisfying the natural conditions (3).

Let us assume that a pre-defined, network of cohesive bands, $\mathcal{B}_{\mu,\text{coh}}$, having finite thickness k (with $k \ll h_\mu$) and characterized by a high aspect ratio, is introduced in the cell, such as shown in Fig. 1(b). It is assumed that fracture at the micro-cell can only appear and propagate along this network, which is dense enough to capture the relevant micro-cell failure mechanisms. A conventional phenomenological non-linear dissipative model describes the constitutive response of these bands. For this purpose the continuum damage model in [8] has been chosen, which, though being extremely simple, provides the basic ingredients (strain softening controlled by fracture energy) involved in the onset and propagation of the micro-cracks.

The cohesive bands constitutive response is governed by the damage model, while the remaining part of the micro-cell, $\mathcal{B}_{\mu,\text{reg}}$, is assumed to behave as linear elastic.

Therefore, the following rules apply, see Fig. 1(b):

- Elastic material model is considered in $\mathcal{B}_{\mu,\text{reg}}$, so:

$$\boldsymbol{\sigma}_\mu(\boldsymbol{y}) = \boldsymbol{C}^e : \boldsymbol{\varepsilon}_\mu; \qquad \forall\, \boldsymbol{y} \in \mathcal{B}_{\mu,\text{reg}} \tag{76}$$

$\boldsymbol{C}^e$ being the fourth order Hooke's isotropic elastic tensor.

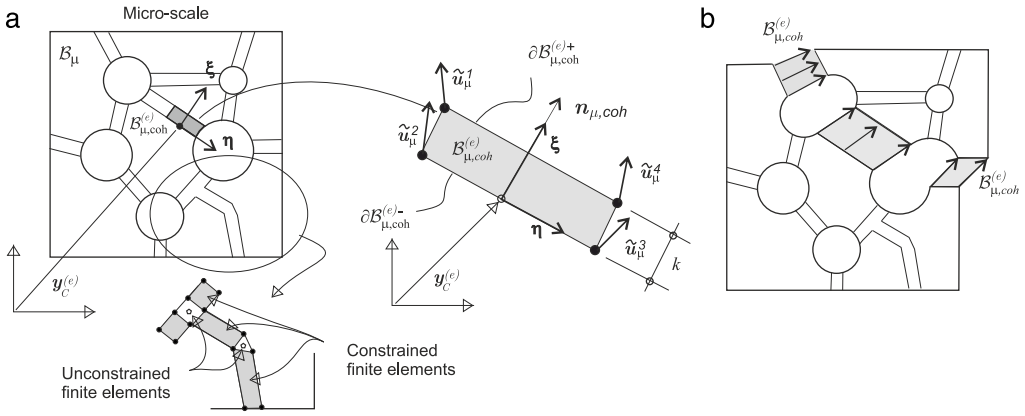


Fig. 12. Finite element discretization at the micro-scale. (a) Finite element, $\mathcal{B}^{(e)}_{\mu,\text{coh}}$, in the cohesive band $\mathcal{B}_{\mu,\text{coh}}$. Constrained and unconstrained finite elements in $\mathcal{B}_{\mu,\text{coh}}$; (b) deformed RVE after activation of a micro-cell failure mode.

- The continuum damage model is considered in the bands $\mathcal{B}_{\mu,\text{coh}}$, so:

$$\boldsymbol{\sigma}_\mu(\boldsymbol{y}, \boldsymbol{\varepsilon}_\mu, d_\mu) = (1 - d_\mu)\boldsymbol{C}^e : \boldsymbol{\varepsilon}_\mu; \qquad \forall\, \boldsymbol{y} \in \mathcal{B}_{\mu,\text{coh}} \tag{77}$$

where $d_\mu \in [0, 1]$ is the scalar, micro-scale damage variable which plays the role of internal variable. Its evolution equation is given by:

$$\dot{d}_\mu(\boldsymbol{y}, \boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu) \tag{78}$$

where g is a function, of the instantaneous mechanical state, ruling the damage evolution and the material hardening/softening. In Eq. (78) the upper dot is used to denote time (or pseudo-time) derivative. A uniform strain is assumed within the cohesive bands, $\mathcal{B}_{\mu,\text{coh}}$, and therefore, the bandwidth k can be used as a regularization parameter on the model softening modulus, in terms of the material fracture energy, G_f, thus providing a correct energy dissipation in the band. Damage evolution is considered only for tensile strain states. A brief summary of the damage model can be found in [8].

To connect both scales of analysis, the Hill–Mandel variational principle is adopted, which assumes the equality of the macro-scale virtual power density and the volumetric average of the micro-scale virtual power density. Two consequences yield from this principle (see [29]):

(i) The micro-stresses should be, in a variational sense, self-equilibrated. Let us consider the space $\mathcal{V}^u_\mu$ of kinematically admissible micro-displacement fluctuations. Functions in $\mathcal{V}^u_\mu$ have to satisfy the kinematical constraint (3). So, $\mathcal{V}^u_\mu := \mathcal{U}^u_\mu$. Then, the self-equilibrium equation reads:

$$\int_{\mathcal{B}_\mu} \nabla^s \delta\tilde{\boldsymbol{u}}_\mu : \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \nabla^s \tilde{\boldsymbol{u}}_\mu, d_\mu)\, d\mathcal{B}_\mu = 0; \qquad \forall\, \delta\tilde{\boldsymbol{u}}_\mu \in \mathcal{V}^u_\mu \tag{79}$$

where $\tilde{\boldsymbol{u}}_\mu \in \mathcal{U}^u_\mu$.

(ii) The macro-stress is related to the micro-stresses through:

$$\boldsymbol{\sigma} = \frac{1}{|\mathcal{B}_\mu|} \int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \nabla^s \tilde{\boldsymbol{u}}_\mu, d_\mu)\, d\mathcal{B}_\mu \tag{80}$$

which is the stress homogenization rule for stress up-scaling.

A.3. Finite element model of the RVE

A standard finite element model is adopted for the numerical simulation of the RVE failure response. One single quadrilateral finite element is used along the thickness of the cohesive bands in $\mathcal{B}_{\mu,\text{coh}}$, as shown in Fig. 12. As

mentioned above, these elements are endowed with a damage model displaying a regularized softening, the parameter k playing the role of regularization length.

In [8] it has been shown that, during the post-critical loading regime of the micro-cell endowed with the cohesive bands network, the kinematical constraints (3) are not sufficient to preclude loss of uniqueness of the numerical solution. It is also shown there that this instability can be easily avoided by the imposition of an internal restriction on the displacement fluctuations at the cohesive bands.

Let us consider that $\tilde{\boldsymbol{u}}_\mu(\xi,\eta)$ is the displacement fluctuation at the point with coordinates (ξ,η), in $\mathcal{B}_{\mu,\text{coh}}$ (increasing values of the coordinate, ξ, are in the sense of $\boldsymbol{n}_{\mu,coh}$), see Fig. 12(a). The displacement fluctuation jump across $\mathcal{B}_{\mu,\text{coh}}$, denoted as $[\![\tilde{\boldsymbol{u}}_\mu]\!]$, is defined by:

$$[\![\tilde{\boldsymbol{u}}_\mu]\!](\eta) := \tilde{\boldsymbol{u}}_\mu(k,\eta) - \tilde{\boldsymbol{u}}_\mu(0,\eta) \tag{81}$$

where $\tilde{\boldsymbol{u}}_\mu(k,\eta)$ and $\tilde{\boldsymbol{u}}_\mu(0,\eta)$ are the displacement fluctuation on the boundary $\partial\mathcal{B}^{(e),+}_{\mu,\text{coh}}$ and $\partial\mathcal{B}^{(e),-}_{\mu,\text{coh}}$, respectively. Then, the field $\tilde{\boldsymbol{u}}_\mu(\xi,\eta)$ is constrained by the condition:

$$[\![\tilde{\boldsymbol{u}}_\mu]\!](\eta) = \text{constant}. \tag{82}$$

The internal constraint (82) is implemented in each finite element in $\mathcal{B}_{\mu,\text{coh}}$, through the nodal constraints :

$$\tilde{\boldsymbol{u}}^1_\mu - \tilde{\boldsymbol{u}}^2_\mu = \tilde{\boldsymbol{u}}^4_\mu - \tilde{\boldsymbol{u}}^3_\mu \tag{83}$$

where $\tilde{\boldsymbol{u}}^1_\mu$, $\tilde{\boldsymbol{u}}^2_\mu$, $\tilde{\boldsymbol{u}}^3_\mu$ and $\tilde{\boldsymbol{u}}^4_\mu$ are defined in Fig. 12(a).

Restrictions in Eq. (83) can become troubleshooting in elements of $\mathcal{B}_{\mu,\text{coh}}$ resulting from cohesive band intersection, because they might over-constrain the adjacent bands. Therefore, they are only imposed on those cohesive bands not producing those this over-constraint[16] (see Fig. 12(a)). For the $\mathcal{B}_{\mu,\text{coh}}$ morphologies used in this work, this setting has shown an excellent behavior and no instability has appeared.

Appendix B. Hyper-reduced saddle-point formulation for inelastic mechanical problems

B.1. Saddle-point formulation

Let us consider a body $\mathcal{B}$, constituted by a dissipative local inviscid material. For the present mechanical description, the relevant state variables are displacement $\boldsymbol{u}(\boldsymbol{x},t)$, strains $\boldsymbol{\varepsilon}(\boldsymbol{u})$, stresses $\boldsymbol{\sigma}$, and a set of internal variables $\boldsymbol{\alpha}(\boldsymbol{x},t)$. The specific aspect considered here is that $\boldsymbol{\alpha}$ typically depends on the strain history, and therefore, it is described in terms of evolution equations having the form:

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{x},t) = g(\boldsymbol{\varepsilon},\boldsymbol{\alpha}) \tag{84}$$

which, after time integration returns the internal variables, $\boldsymbol{\alpha}$, in terms of the strain history $\boldsymbol{\varepsilon}^\tau$:

$$\boldsymbol{\alpha}(\boldsymbol{x},t) = \hat{\boldsymbol{\alpha}}(\boldsymbol{x},\boldsymbol{\varepsilon}^\tau); \qquad \tau\in[0,t]. \tag{85}$$

The free energy for any inelastic model can be generally defined by: $\varphi(\boldsymbol{\varepsilon},\boldsymbol{\alpha})$. Then, the stress $\boldsymbol{\sigma}$ is:

$$\boldsymbol{\sigma} = \frac{\partial\varphi(\boldsymbol{\varepsilon},\alpha)}{\partial\boldsymbol{\varepsilon}} \tag{86}$$

where it is assumed that α and $\boldsymbol{\varepsilon}$ are related through Eq. (85).

Following the same concept described in Section 2.3.1, we can introduce the parameterized free energy $\varphi_{[\alpha]}$. In this function, the internal variable α is assumed as an independent parameter, in the sense that once the strain history $\boldsymbol{\varepsilon}^\tau$ is defined, α does not necessarily satisfy the evolution equation (85). Then, we can write that $\boldsymbol{\sigma}$ is given by the set of equations:

$$\boldsymbol{\sigma} = \frac{\partial\varphi_{[\alpha]}(\boldsymbol{\varepsilon})}{\partial\boldsymbol{\varepsilon}}; \tag{87}$$

$$\boldsymbol{\alpha}(\boldsymbol{x},t) = \hat{\boldsymbol{\alpha}}(\boldsymbol{x},\boldsymbol{\varepsilon}^\tau); \qquad \tau\in[0,t]. \tag{88}$$

[16] An alternative option, which is used in the examples presented in Section 5, is removal of these elements from the finite element mesh.

Next, we define a parameterized functional $\Pi_{[\boldsymbol{t}^*,\boldsymbol{u}^*,\alpha]}$, where $\boldsymbol{t}^*(\boldsymbol{x},t)$ (the imposed traction at the boundary Γ_t) and $\boldsymbol{u}^*(\boldsymbol{x},t)$ (the imposed displacements at the boundary Γ_u) are the problem driving parameters (actions), and $\boldsymbol{\alpha}(\boldsymbol{x},t)$ is, again, considered as an additional free parameter:

$$\Pi_{[\boldsymbol{t}^*,\boldsymbol{u}^*,\alpha]}(\boldsymbol{u},\boldsymbol{\lambda}_u) = \int_{\mathcal{B}} \varphi_{[\alpha]}(\boldsymbol{\varepsilon}(\boldsymbol{u}))\, d\mathcal{B} - \int_{\Gamma_t} \boldsymbol{t}^* \cdot \boldsymbol{u}\, d\Gamma - \int_{\Gamma_u} \boldsymbol{\lambda}_u \cdot (\boldsymbol{u}-\boldsymbol{u}^*)\, d\Gamma \tag{89}$$

where $\boldsymbol{\lambda}_u(\boldsymbol{x},t)$ (with $\boldsymbol{\lambda}_u \in \mathbb{R}^{n_{dim}}$) is a Lagrange multiplier defined in Γ_u enforcing the Dirichlet's condition ($\boldsymbol{u}=\boldsymbol{u}^*$ at Γ_u).

Now, let us consider the saddle-point problem:

$$\Big(\boldsymbol{u}(\boldsymbol{t}^*,\boldsymbol{u}^*,\alpha),\boldsymbol{\lambda}_u(\boldsymbol{t}^*,\boldsymbol{u}^*,\alpha)\Big) = \arg\Big\{\min_{\boldsymbol{u}}\max_{\boldsymbol{\lambda}_u} \Pi_{[\boldsymbol{t}^*,\boldsymbol{u}^*,\boldsymbol{\alpha}]}(\boldsymbol{u},\boldsymbol{\lambda}_u)\Big\} \tag{90}$$

whose optimality conditions yield the following variational equations:

$$\int_{\mathcal{B}} \frac{\partial \varphi_{[\alpha]}}{\partial \boldsymbol{\varepsilon}} : \nabla^s \delta\boldsymbol{u}\, d\mathcal{B} - \int_{\Gamma_t} \boldsymbol{t}^* \cdot \delta\boldsymbol{u}\, d\Gamma - \int_{\Gamma_u} \boldsymbol{\lambda}_u \cdot \delta\boldsymbol{u}\, d\Gamma = 0; \qquad \forall \delta\boldsymbol{u} \in \mathcal{V}_u; \tag{91}$$

$$-\int_{\Gamma_u} \delta\boldsymbol{\lambda}_u \cdot (\boldsymbol{u}-\boldsymbol{u}^*)\, d\Gamma = 0; \qquad \forall \delta\boldsymbol{\lambda}_u \in \mathcal{V}_\lambda; \tag{92}$$

while, the corresponding Euler–Lagrange equations are obtained:

$$\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{0}; \qquad \boldsymbol{\sigma}(\boldsymbol{\varepsilon},\boldsymbol{\alpha}) = \partial_{\boldsymbol{\varepsilon}}\varphi_{[\alpha]}(\boldsymbol{\varepsilon}); \qquad \forall \boldsymbol{x} \in \mathcal{B} \tag{93}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{v} = \boldsymbol{t}^* \qquad \forall \boldsymbol{x} \in \Gamma_t \tag{94}$$

$$\boldsymbol{\lambda}_u = \boldsymbol{\sigma}\cdot\boldsymbol{v} \qquad \forall \boldsymbol{x} \in \Gamma_u \tag{95}$$

$$\boldsymbol{u} = \boldsymbol{u}^* \qquad \forall \boldsymbol{x} \in \Gamma_u. \tag{96}$$

Eqs. (93)–(96) in combination with (85) constitute the conventional BVP in $\mathcal{B}$. The saddle-point problem (90) together with (84) (or (85)) defines the full set of equations governing the inelastic mechanical problem.

Therefore, one can state the entire formulation in terms of the saddle point problem of the parameterized energy functional $\Pi_{[\boldsymbol{t}^*,\boldsymbol{u}^*,\boldsymbol{\alpha}]}$ (Eq. (90)) complemented by the evolution equation (84). In this sense, *the saddle point problem* (Eq. (90)) *can be considered as the primitive statement of the standard variational formulation in* Eqs. (91) and (92). This is the paradigm of the proposed HPROM strategy, that considers as the start point the functional in Eq. (90).

B.2. Hyper-reduction of the problem

B.2.1. Low-dimensional problem. Reduced order model (ROM)

Initially, we assume that we have available a conventional finite element approach for the displacement, $\boldsymbol{u}$, and Lagrange multipliers, $\boldsymbol{\lambda}_u$, fields solving the above described BVP (the HF finite element model). Then, using a procedure similar to that described in [26], (performing an off-line sampling program of the HF model followed by a POD technique) we can extract a reduced order model to approach both fields.[17]

In accordance with this procedure, the reduced displacement field can be written as a linear combination of vectors defining the basis $\{\boldsymbol{U}\} := \{\boldsymbol{U}_1,\ldots,\boldsymbol{U}_{n_q}\}$, as follows:

$$\boldsymbol{u}(\boldsymbol{x},t) = \sum_{i=1}^{n_q} \boldsymbol{U}_i(\boldsymbol{x}) q_i(t) = \boldsymbol{U}(\boldsymbol{x})\boldsymbol{q}(t); \tag{97}$$

where

$$\boldsymbol{U} = [\boldsymbol{U}_1,\ldots,\boldsymbol{U}_{n_q}] \in \mathbb{R}^{n_{dim}\times n_q}; \quad \boldsymbol{U}_i = \begin{bmatrix} U_{i,x} \\ U_{i,y} \end{bmatrix}; \quad \boldsymbol{q} = [q_1,\ldots,q_{n_q}] \in \mathbb{R}^{n_q}. \tag{98}$$

[17] Voigt's notation is used from now on.

Element $\boldsymbol{U}_i$, of the basis $\{\boldsymbol{U}\}$, has the same number of components as the displacement field, and, contrarily to the finite element shape functions, they are global functions, in the sense that its support is $\mathcal{B}$. The matrix $\boldsymbol{U}$ collects all the displacement modes defined by the basis functions, while the vector $\boldsymbol{q}$ collects all the coefficients scaling the displacement modes.

The Lagrange multiplier field can also be expanded using a low-dimension spatial basis $\{\boldsymbol{\Upsilon}\} := \{\boldsymbol{\Upsilon}_1, \ldots, \boldsymbol{\Upsilon}_{n_\gamma}\}$:

$$\boldsymbol{\lambda}_u(\boldsymbol{x}, t) = \sum_{i=1}^{n_\gamma} \boldsymbol{\Upsilon}_i(\boldsymbol{x})\boldsymbol{\gamma}_i(t) = \boldsymbol{\Upsilon}(\boldsymbol{x})\boldsymbol{\gamma}(t); \tag{99}$$

where now, matrices $\boldsymbol{\Upsilon} \in \mathbb{R}^{n_{dim} \times n_\gamma}$ and vector $\boldsymbol{\gamma} \in \mathbb{R}^{n_\gamma}$ have similar meaning to $\boldsymbol{U}$ and $\boldsymbol{q}$.

In an infinitesimal strain setting, the strains are given by:

$$\boldsymbol{\varepsilon}(\boldsymbol{x}, t) = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{q}(t) \tag{100}$$

where $\boldsymbol{B}(\boldsymbol{x})$ is the strain–displacement matrix, which for 2-D ($n_{dim} = 2$) problems is:

$$\boldsymbol{B}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial}{\partial x} & 0 \\ 0 & \dfrac{\partial}{\partial y} \\ 0 & 0 \\ \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} \end{bmatrix} \boldsymbol{U}. \tag{101}$$

Replacing the approximations (97) and (100) into the functional (89), results:

$$\Pi_{[\boldsymbol{t}^*, \boldsymbol{u}^*, \alpha]}(\boldsymbol{q}, \boldsymbol{\gamma}) = \underbrace{\int_{\mathcal{B}} \varphi_{[\alpha]}(\boldsymbol{B}\boldsymbol{q})\, d\mathcal{B}}_{I_1} - \underbrace{\int_{\Gamma_t} (\boldsymbol{t}^*)^T \boldsymbol{U}\boldsymbol{q}\, d\Gamma}_{I_2} - \underbrace{\int_{\Gamma_u} \boldsymbol{\gamma}^T \boldsymbol{\Upsilon}^T (\boldsymbol{U}\boldsymbol{q} - \boldsymbol{u}^*)\, d\Gamma}_{I_3}. \tag{102}$$

The min–max problem of Eq. (90) reads:

$$\Big(\boldsymbol{q}(\boldsymbol{t}^*, \boldsymbol{u}^*, \alpha), \boldsymbol{\gamma}(\boldsymbol{t}^*, \boldsymbol{u}^*, \alpha)\Big) = \arg\Big\{\min_{\boldsymbol{q}} \max_{\boldsymbol{\gamma}} \Pi_{[\boldsymbol{t}^*, \boldsymbol{u}^*, \boldsymbol{\alpha}]}(\boldsymbol{q}, \boldsymbol{\gamma})\Big\} \tag{103}$$

$$\partial_{\boldsymbol{q}} \Pi_{[\boldsymbol{t}^*, \boldsymbol{u}^*, \alpha]}(\boldsymbol{q}, \boldsymbol{\gamma}) = \int_{\mathcal{B}} \boldsymbol{B}^T \frac{\partial \varphi_{[\alpha]}}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{B}\boldsymbol{q})\, d\mathcal{B} - \int_{\Gamma_t} \boldsymbol{U}^T \boldsymbol{t}^*\, d\Gamma - \int_{\Gamma_u} \boldsymbol{U}^T \boldsymbol{\Upsilon}\boldsymbol{\gamma}\, d\Gamma \tag{104}$$

$$\partial_{\boldsymbol{\lambda}_u} \Pi_{[\boldsymbol{t}^*, \boldsymbol{u}^*, \alpha]}(\boldsymbol{q}, \boldsymbol{\gamma}) = -\int_{\Gamma_u} \boldsymbol{\Upsilon}^T (\boldsymbol{U}\boldsymbol{q} - \boldsymbol{u}^*) d\mathcal{B} = \boldsymbol{0}; \tag{105}$$

which, in combination with the evolution equation (85):

$$\boldsymbol{\alpha}(\boldsymbol{x}, t) = \hat{\boldsymbol{\alpha}}(\boldsymbol{x}, (\boldsymbol{B}\boldsymbol{q})^\tau); \qquad \tau \in [0, t] \tag{106}$$

supplies the solution of the reduced problem.

Let us now assume that the actions $\boldsymbol{t}^*$ and $\boldsymbol{u}^*$ can be also projected on appropriated low-dimension spaces with bases, $\boldsymbol{U}^*$ and $\boldsymbol{T}^*$, respectively, and expressed as follows:

$$\boldsymbol{u}^*(\boldsymbol{x}, t) = \boldsymbol{U}^*(\boldsymbol{x})\boldsymbol{r}^*(t); \qquad \boldsymbol{t}^*(\boldsymbol{x}, t) = \boldsymbol{T}^*(\boldsymbol{x})\boldsymbol{s}^*(t). \tag{107}$$

Eqs. (104)–(107) constitute a set of non-linear algebraic equation in $(\boldsymbol{q}, \boldsymbol{\gamma})$ solved in terms of the *action parameters* $(\boldsymbol{r}^*, \boldsymbol{s}^*)$, as:

$$\boldsymbol{q}(t) = \hat{\boldsymbol{q}}(\boldsymbol{r}^*(t), \boldsymbol{s}^*(t)); \qquad \boldsymbol{\gamma}(t) = \hat{\boldsymbol{\gamma}}(\boldsymbol{r}^*(t), \boldsymbol{s}^*(t)) \tag{108}$$

They constitute the Reduced Order Model, the number of unknowns being reduced to $n_q + n_\gamma$ (ROM problem).

B.2.2. Reduced Optimal Quadrature (ROQ)

As a matter of example, let us assume that we are interested in reducing *the numerical integration cost* of the term I_1 in Eq. (102). Replacing the original conventional quadrature, based on the discrete Gauss points $\boldsymbol{x}_i : i = 1, \dots, N_g$, in $\mathcal{B}$, and Gauss weights $\mathrm{W}_i : i = 1, \dots, N_g$,

$$\int_{\mathcal{B}}^{Gauss} \varphi_{[\alpha]}(\boldsymbol{B}(\boldsymbol{x})\boldsymbol{q})\, d\mathcal{B} := \sum_{i=1}^{N_g} \varphi_{[\alpha]}(\boldsymbol{B}(\boldsymbol{x}_i)\boldsymbol{q})\, \mathrm{W}_i \tag{109}$$

by a Reduced Optimal Quadrature, based on a new set of discrete points $\boldsymbol{z}_j; j = 1, \dots, N_r$ and weights $\omega_j; j = 1, \dots, N_r$, with $N_r \ll N_g$, i.e.:

$$\int_{\mathcal{B}}^{Reduced} \varphi_{[\alpha]}(\boldsymbol{B}(\boldsymbol{z})\boldsymbol{q})\, d\mathcal{B} := \sum_{j=1}^{N_r} \varphi_{[\alpha]}(\boldsymbol{B}(\boldsymbol{z}_j)\boldsymbol{q})\, \omega_j. \tag{110}$$

The goal is to determine the weights and positions of these quadrature points so as to minimize the differences provided by both integrations on the same kernel:

$$(\underbrace{\boldsymbol{z}_j, \omega_j}_{j=1,\dots,N_r}) = \arg\left\{ \min_{\boldsymbol{z}_j, \omega_j} \left(\left| \int_{\mathcal{B}}^{Gauss} \varphi_{[\alpha]}(\boldsymbol{B}(\boldsymbol{z})\boldsymbol{q})\, d\mathcal{B} - \int_{\mathcal{B}}^{Reduced} \varphi_{[\alpha]}(\boldsymbol{B}(\boldsymbol{z})\boldsymbol{q})\, d\mathcal{B} \right| \right) \right\}. \tag{111}$$

In order to achieve this goal, let us consider that kernel, $\varphi_{[\alpha]}$, evaluated *at the solutions of the ROM problem* (Eqs. (108)):

$$\begin{aligned} I_1 &= \int_{\mathcal{B}} \varphi_{[\alpha]}(\underbrace{\boldsymbol{B}(\boldsymbol{x})\hat{\boldsymbol{q}}(\boldsymbol{r}^*, \boldsymbol{s}^*)}_{\boldsymbol{\varepsilon}(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*)})\, d\mathcal{B} = \int_{\mathcal{B}} \underbrace{\varphi(\boldsymbol{\varepsilon}(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*), \boldsymbol{\alpha}(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*))}_{\Psi(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*)}\, d\mathcal{B}; \\ &= \int_{\mathcal{B}} \Psi(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*)\, d\mathcal{B}. \end{aligned} \tag{112}$$

And let us assume that an appropriate separated expansion of the kernel of the integral in (112) is available:

$$\Psi(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*) = \sum_{i=1}^{n_\psi} \boldsymbol{F}_i(\boldsymbol{x}) \boldsymbol{f}_i(\boldsymbol{r}^*, \boldsymbol{s}^*). \tag{113}$$

Then, I_1 can be replaced by:

$$I_1 = \int_{\mathcal{B}} \Psi(\boldsymbol{x}, \boldsymbol{r}^*, \boldsymbol{s}^*)\, d\mathcal{B} = \sum_{i=1}^{n_\psi} \Big(\int_{\mathcal{B}} \boldsymbol{F}_i(\boldsymbol{x})\, d\mathcal{B} \Big) \boldsymbol{f}_i(\boldsymbol{r}^*, \boldsymbol{s}^*). \tag{114}$$

A sufficient condition for Eq. (114) to be fulfilled is that every element of the basis, $\boldsymbol{F}_i(\boldsymbol{z})$, is exactly integrated by the proposed reduced integration scheme:

$$\int_{\mathcal{B}} \boldsymbol{F}_k(\boldsymbol{x})\, d\mathcal{B} = \int_{\mathcal{B}}^{Reduced} \boldsymbol{F}_k(\boldsymbol{x})\, d\mathcal{B} = \sum_{j=1}^{N_r} \boldsymbol{F}_k(\boldsymbol{z}_j)\, \omega_j; \qquad \text{for: } k = 1, \dots, N_r. \tag{115}$$

In addition, let us require that the volume is preserved by the reduced integration rule, e.g.:

$$\int_{\mathcal{B}}^{Reduced} d\mathcal{B} = |\mathcal{B}| = \sum_{j=1}^{N_r} \omega_j. \tag{116}$$

Eqs. (115) and (116) can be reduced to

$$\underbrace{(\ z_j,\omega_j\ )}_{j=1,\dots,n_\psi} = \arg\Big\{\min_{z_j,\omega_j}\sum_{k=1}^{N_r}\Big\{\Big(\Big|\underbrace{\int_{\mathcal{B}}^{Gauss}\boldsymbol{F}_k(\boldsymbol{x})\,d\mathcal{B}}_{\sum_{i=1}^{N_g}\boldsymbol{F}_k(\boldsymbol{x}_i)\,\mathrm{W}_i}-\underbrace{\int_{\mathcal{B}}^{Reduced}\boldsymbol{F}_k(\boldsymbol{z})\,d\mathcal{B}}_{\sum_{j=1}^{N_r}\boldsymbol{F}_k(\boldsymbol{z}_j)\,\omega_j}\Big|+\Big||\mathcal{B}|-\sum_{j=1}^{N_r}\omega_j\Big|\Big)\Big\}\Big\}. \tag{117}$$

Problem in Eq. (117) may be solved by means of a specific greedy algorithm (see [17] for a detailed description). It is proven that, in general, an exact reduced integration scheme, keeping the weights $\omega_j > 0$, can be found for: $N_r = n_{\psi+1}$. So, replacing this reduced integration scheme into the variational problem in (102)–(106) yields

$$\partial_{\boldsymbol{q}}\Pi(\boldsymbol{q},\boldsymbol{\gamma}) = \underbrace{\int_{\mathcal{B}}^{Reduced}\boldsymbol{B}^T\partial_{\boldsymbol{\varepsilon}}\varphi_{[\alpha]}(\boldsymbol{B}\boldsymbol{q})\,d\mathcal{B}}_{\partial_{\boldsymbol{q}}I_1} - \int_{\Gamma_t}^{Gauss}\boldsymbol{U}^T\boldsymbol{t}^*\,d\Gamma - \int_{\Gamma_u}^{Gauss}\boldsymbol{U}^T\boldsymbol{\Upsilon}\boldsymbol{\gamma}\,d\Gamma \tag{118}$$

$$\partial_{\boldsymbol{\lambda}_u}\Pi(\boldsymbol{q},\boldsymbol{\gamma}) = -\int_{\Gamma_u}^{Gauss}\boldsymbol{\Upsilon}^T(\boldsymbol{U}\boldsymbol{q}-\boldsymbol{u}^*)d\mathcal{B} = \boldsymbol{0}; \tag{119}$$

$$\boldsymbol{\alpha}(\boldsymbol{x},t) = \hat{\boldsymbol{\alpha}}(\boldsymbol{x},\boldsymbol{B}\boldsymbol{q}^\tau); \qquad \tau\in[0,t]. \tag{120}$$

A substantial cost reduction can be expected from the use of the reduced quadrature in the first integral term (I_1) in (118). If necessary, a similar treatment can be given to the remaining integrals (I_2, I_3) in the functional (102). Combination of the ROM, in Appendix B.2.1, and ROQ, in Appendix B.2.2, constitutes the proposed HPROM approach.

Appendix C. Displacement fluctuation field recovery

Solutions of the HPROM problem in Eqs. (50) and (51), provide the coefficient vectors $\boldsymbol{q}$ which, jointly with the basis $\{\boldsymbol{\Psi}\}$, define the micro-strain fluctuation field: $\tilde{\boldsymbol{\varepsilon}}_\mu = \boldsymbol{\Psi}\boldsymbol{c}$.

Although, within the context of the FE^2 multiscale approach, computation of $\tilde{\boldsymbol{\varepsilon}}_\mu$ and $\boldsymbol{\sigma}_\mu$ is sufficient to determine the macroscale homogenized constitutive law, in same cases[18] it could be necessary to evaluate the micro-displacement fluctuation field $\tilde{\boldsymbol{u}}_\mu$. In this case, $\tilde{\boldsymbol{u}}_\mu$ can be obtained by the projection of $\tilde{\boldsymbol{\varepsilon}}_\mu$ into the strain space approached by the HF finite element model, as follows.

GIVEN $\boldsymbol{\Psi}\boldsymbol{c}$, FIND: $\tilde{\boldsymbol{u}}_\mu \in \mathcal{V}^{HF}$ such that:

$$\int_{\mathcal{B}}(\boldsymbol{\Psi}\boldsymbol{c}-\nabla^s\tilde{\boldsymbol{u}}_\mu):\nabla^s\hat{\boldsymbol{u}}_\mu d\mathcal{B} = 0; \quad \forall\hat{\boldsymbol{u}}_\mu\in\mathcal{V}^{HF} \tag{121}$$

where $\mathcal{V}^{HF}$ is the displacement space obtained with the HF conventional finite element model and removing the rigid body motion modes. Note that functions in $\mathcal{V}^{HF}$ are not subjected to the kinematical constraint defining the space $\mathcal{U}^u_\mu$ in expression (4).

Introducing the finite element approaches of $\tilde{\boldsymbol{u}}_\mu$ and $\hat{\boldsymbol{u}}_\mu$, and considering that $[\tilde{\boldsymbol{u}}_\mu]$ is the global vector of nodal parameters, interpolating $\tilde{\boldsymbol{u}}_\mu$ in the original finite element mesh, Eq. (121) can be re-written as follows:

$$\underbrace{\Big(\int_{\mathcal{B}}\boldsymbol{B}_u^T\boldsymbol{\Psi}d\mathcal{B}\Big)}_{\mathbb{F}}\boldsymbol{c} - \underbrace{\Big(\int_{\mathcal{B}}\boldsymbol{B}_u^T\boldsymbol{B}_u d\mathcal{B}\Big)}_{\mathbb{K}}[\tilde{\boldsymbol{u}}_\mu] = 0; \tag{122}$$

where $\boldsymbol{B}_u$ is the strain–displacement matrix of the conventional HF finite element method. Finally:

$$[\tilde{\boldsymbol{u}}_\mu] = \mathbb{K}^{-1}\mathbb{F}\boldsymbol{c}; \tag{123}$$

Computation of $[\tilde{\boldsymbol{u}}_\mu]$ with (123) destroys the speed-up obtained with the HPROM approach. Therefore, this field should be only determined as a post-processing step, at the end of the computation.

[18] For instance, for geometric representation of the RVE deformation.

References

[1] J.T. Oden, T. Belytschko, J. Fish, T.J.R. Hughes, C. Johnson, L.A. Keyes, L. Petzold, L. Srolovitz, S. Yip, Simulation-based Engineering Science. Report of the nsf sbes Panel to the nsf Engineering Advisory Committee, National Science Foundation (USA), 2006.

[2] T. Belytschko, S. Loehnert, J.H. Song, Multiscale aggregating discontinuities: A method for circumventing loss of material stability, Internat. J. Numer. Methods Engrg. 73 (2008) 869–894.

[3] M.G.D. Geers, V.G. Kouznetsova, W.A.M. Brekelmans, Multi-scale computational homogenization: Trends and challenges, J. Comput. Appl. Math. 234 (2010) 2175–2182.

[4] J.F. Unger, S. Eckardt, Multiscale modeling of concrete, Arch. Comput. Methods Eng. 18 (3) (2011) 341–393.

[5] V.P. Nguyen, O. Lloberas-Valls, L.J. Sluys, M. Stroeven, Homogenization-based multiscale crack modelling: from micro diffusive damage to macro cracks, Comput. Methods. Appl. Mech. Engrg. 200 (2010) 1220–1236.

[6] P.J. Sánchez, P.J. Blanco, A.E. Huespe, R.A. Feijóo, Failure-oriented multi-scale variational formulation: micro-structures with nucleation and evolution of softening bands, Comput. Methods Appl. Mech. Engrg. 257 (2013) 221–247.

[7] S. Toro, P.J. Sánchez, P.J. Blanco, E.A. de Souza Neto, A.E. Huespe, R.A. Feijóo, Multiscale formulation for material failure accounting for cohesive cracks at the macro and micro scales, Int. J. Plast. 76 (2016) 75–110.

[8] J. Oliver, M. Caicedo, E. Roubin, A.E. Huespe, J.A. Hernández, Continuum approach to computational multiscale modeling of propagating fracture, Comput. Methods Appl. Mech. Engrg. 294 (2015) 384–427.

[9] F. Feyel, J.L. Chaboche, FE^2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials, Comput. Methods Appl. Mech. Engrg. 183 (2000) 309–330.

[10] J.L. Michel, H. Moulinec, P. Suquet, A computational scheme for linear and non-linear composites with arbitrary phase contrast, Internat. J. Numer. Methods Engrg. 52 (1–2) (2001) 139–160.

[11] J.C Yvonnet, Q.-C. He, The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains, J. Comput. Phys. 223 (1) (2007) 341–368.

[12] D. Ryckelynck, Hyper-reduction of mechanical models involving internal variables, Internat. J. Numer. Methods Engrg. 77 (1) (2009) 75–89.

[13] C. Oskay, J. Fish, Eigendeformation-based reduced order homogenization for failure analysis of heterogeneous materials, Comput. Methods Appl. Mech. Engrg. 196 (7) (2007) 1216–1243.

[14] S. Zhang, C. Oskay, Reduced order variational multiscale enrichment method for elasto-viscoplastic problems, Comput. Methods Appl. Mech. Engrg. 300 (2016) 199–224.

[15] P. Kerfriden, O. Goury, T. Rabczuk, S. Bordas, A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics, Comput. Methods Appl. Mech. Engrg. 256 (2013) 169–188.

[16] C. Farhat, T. Chapman, P. Avery, Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models, Internat. J. Numer. Methods Engrg. (2015).

[17] J.A. Hernández, M. Caicedo, A. Ferrer, Dimensional hyper-reduction of nonlinear finite element models via empirical cubature, Comput. Methods Appl. Mech. Engrg. (2016). http://dx.doi.org/10.1016/j.cma.2016.10.022 (in press).

[18] J. Simo, J. Oliver, A new approach to the analysis and simulation of strain softening in solids, in: Z.B. Bazant, Z. Bittnar, M. Jirásek, J. Mazars (Eds.), Fracture and Damage in Quasi-brittle Structures, E & FN Spon, 1994, pp. 25–39.

[19] J. Oliver, Modelling strong discontinuities in solids mechanics via strain softening constitutive equations. Part 1: Fundamentals, Internat. J. Numer. Methods Engrg. 39(21) (1996) 3575–3600.

[20] J. Oliver, On the discrete constitutive models induced by strong discontinuity kinematics and continuum constitutive equations, Int. J. Solids Struct. 37 (2000) 7207–7229.

[21] J. Oliver, A.E. Huespe, M.D.G. Pulido, E. Chaves, From continuum mechanics to fracture mechanics: the strong discontinuity approach, Eng. Fract. Mech. 69 (2002) 113–136.

[22] O.L. Manzoli, A.L. Gamino, E.A. Rodrigues, G.K.S. Claro, Modeling of interfaces in two-dimensional problems using solid finite elements with high aspect ratio, Comput. Struct. 94 (2012) 70–82.

[23] E.A. Rodrigues, O.L. Manzoli, L.A.G. Bitencourt Jr., T. Bittencourt, 2D mesoscale model for concrete based on the use of interface element with a high aspect ratio, Int. J. Solids Struct. 94 (2016) 112–124.

[24] J. Oliver, C.A. de Saracíbar Bosch, Mecánica de medios continuos para ingenieros, vol. 92, Univ. Politèc. de Catalunya, 2002.

[25] J. Simo, T. Hughes, Computational inelasticity, Springer-Verlag, 1998.

[26] J.A. Hernández, J. Oliver, A.E. Huespe, M.A. Caicedo, J.C. Cante, High-performance model reduction techniques in computational multiscale homogenization, Comput. Methods Appl. Mech. Engrg. 276 (2014) 149–189.

[27] Anindya Chatterjee, An introduction to the proper orthogonal decomposition, Current Sci. 78 (7) (2000) 808–817.

[28] J. Oliver, A.E. Huespe, J.C. Cante, G. Díaz, On the numerical resolution of the discontinuous material bifurcation problem, Internat. J. Numer. Methods. Engrg. 83 (2010) 786–804.

[29] P.J. Blanco, P.J. Sánchez, E.A. de Souza Neto, R.A. Feijóo, Variational foundations and generalized unified theory of RVE-based multiscale models, Arch. Comput. Methods Eng. (2014) 1–63.

B.4 Paper #4

Title: Dimensional hyper-reduction of nonlinear finite element models via empirical cubature.
Authors:

- **J. A. Hernández**: Assistant Professor of Structural Engineering and Strength of Materials at the School of Industrial and Aeronautic Engineering of Terrassa, of the Technical University of Catalonia. Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **A. Ferrer**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).

Journal of Computer Methods in Applied Mechanics and Engineering
Editors: Thomas J.R. Hughes, J. Tinsley Oden, Manolis Papadrakakis
ISSN: 0045-7825
Elsevier Editors

Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 313 (2017) 687–722
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Dimensional hyper-reduction of nonlinear finite element models via empirical cubature

J.A. Hernández[a,b,*], M.A. Caicedo[a], A. Ferrer[a,b]

[a] *Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), Technical University of Catalonia, Edificio C1, Campus Norte, Jordi Girona 1-3, Barcelona 08034, Spain*
[b] *Escola Superior d'Enginyeries Industrial, Aeroespacial i Audiovisual de Terrassa, C/ Colom, 11, Terrassa 08222, Spain*

Received 24 March 2016; received in revised form 18 July 2016; accepted 11 October 2016
Available online 20 October 2016

Abstract

We present a general framework for the dimensional reduction, in terms of number of degrees of freedom as well as number of integration points ("hyper-reduction"), of nonlinear parameterized finite element (FE) models. The reduction process is divided into two sequential stages. The first stage consists in a common Galerkin projection onto a reduced-order space, as well as in the condensation of boundary conditions and external forces. For the second stage (reduction in number of integration points), we present a novel cubature scheme that efficiently determines optimal points and associated positive weights so that the error in integrating reduced internal forces is minimized. The distinguishing features of the proposed method are: (1) The minimization problem is posed in terms of orthogonal basis vector (obtained via a *partitioned* Singular Value Decomposition) rather that in terms of snapshots of the integrand. (2) The volume of the domain is exactly integrated. (3) The selection algorithm need not solve in all iterations a nonnegative least-squares problem to force the positiveness of the weights. Furthermore, we show that the proposed method converges to the absolute minimum (zero integration error) when the number of selected points is equal to the number of internal force modes included in the objective function. We illustrate this model reduction methodology by two nonlinear, structural examples (quasi-static bending and resonant vibration of elastoplastic composite plates). In both examples, the number of integration points is reduced three order of magnitudes (with respect to FE analyses) without significantly sacrificing accuracy.

Keywords: Reduced-order model; Hyper-reduction; Optimized cubature; Finite elements; Singular Value Decomposition

1. Introduction

Generally speaking, model order reduction refers to any endeavor aimed at constructing a simpler model from a more complex one. The simpler model is usually referred to as the *reduced-order model* (ROM), while the more

* Corresponding author at: Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), Technical University of Catalonia, Edificio C1, Campus Norte, Jordi Girona 1-3, Barcelona 08034, Spain.
E-mail address: jhortega@cimne.upc.edu (J.A. Hernández).

http://dx.doi.org/10.1016/j.cma.2016.10.022

complex one is termed the full-order or high-fidelity model. This full-order model may be, for instance, – as is the case here – a finite element (FE) model.

The focus of the present paper is on the so-called *projection-based, reduced-order models*. The existence of such low-dimensional representations for a given parametrized finite element problem relies on the premise that the state variable can be accurately approximated by a linear combination of a few global basis vectors. The most common approach is to determine these basis vectors by applying some type of dimensionality reduction strategy (such as the Proper Orthogonal Decomposition, POD) over a so-called *training sample*. This sample is obtained by previously solving – in an *offline* stage – the full-order model for judiciously chosen values of the input parameters.

1.1. Approximation of nonlinear terms

In the general case of governing equations featuring terms that bear a nonaffine relationship with both the state variable and input parameters, the construction of an inexpensive low-dimensional model entails two sequential stages [1], namely: (1) projection onto the reduced basis, and (2) approximation of the nonlinear term. Once a basis matrix for the state variable is available, the projection stage is a standard operation consisting in introducing the approximation of the state variables in the governing equation, and then in posing the resulting equation in the space spanned by the basis vectors. This operation naturally leads to a significant reduction in the number of unknowns, and hence diminishes considerably the equation solving effort. However, in a general nonlinear case, the computational cost of evaluating the residual still depends on the size of the underlying finite element mesh—hence the need for a second reduction stage.

In contrast to the first reduction stage, which is more or less standard, the second stage of dimensionality reduction – Ryckelynck [2] coined the term *hyper-reduction* to refer to it – is far more challenging and still remains an issue of discussion in the model reduction community. In the following, we examine the various approaches encountered in the related literature to deal with this additional dimensionality reduction stage.

1.2. Classification of "hyper-reduction" methods

Let $\boldsymbol{F}^h \in \mathbb{R}^N$ denote[1] the full-order term bearing a general, nonaffine relationship with both the input variable and the state variable (in the context of this paper, $\boldsymbol{F}^h \in \mathbb{R}^N$ will be the vector of FE *nodal* internal forces). The corresponding projection onto the reduced order space will be represented by $\boldsymbol{F} \in \mathbb{R}^n$ ($n \ll N$), the connection between these two variables being the matrix of basis vectors $\boldsymbol{\Phi} \in \mathbb{R}^{N\times n}$ ($\boldsymbol{F} = \boldsymbol{\Phi}^T \boldsymbol{F}^h$). Existing approaches for dealing with the approximation of $\boldsymbol{F}$ can be broadly classified as *nodal vector approaches and integral approaches*.

1.2.1. Nodal vector approximation approaches ("gappy" data)

In this type of approaches, the approximation is carried out by replacing the finite element vector $\boldsymbol{F}^h$ by a low-dimensional interpolant $\boldsymbol{F}^h \approx \boldsymbol{R}_{\boldsymbol{F}} \boldsymbol{F}^h_{\mathbf{z}}$, $\boldsymbol{R}_{\boldsymbol{F}} \in \mathbb{R}^{N\times m}$ being the interpolation matrix, and $\boldsymbol{F}^h_{\mathbf{z}}$ the entries of $\boldsymbol{F}^h$ corresponding to the degrees of freedom ($\mathbf{z} \subset \{1, 2, \ldots, N\}$) at which the interpolation takes place. The interpolation matrix is obtained following the common procedure of computing a basis matrix for $\boldsymbol{F}^h$, and then determining a set of indices so that the error is minimized over a set of representative snapshots of $\boldsymbol{F}^h$. This set of interpolation indices can be determined *offline* using procedures such as the Empirical Interpolation Method (EIM) [3,4], the Best Points Interpolation Method (BPIM) [5], the Discrete BPIM [6] or the Missing Point Estimation Method [7]. The idea behind this vector approximation approach has its roots in the landmark work of Everson and Sirovich [8] for reconstruction of "gappy" data, and was historically the first proposal for dealing with nonlinear terms in model order reduction; it has been adopted by, among others, [1,9–14]. Alternatively, [2] proposes to bypass the construction of the low-dimensional interpolant and simply solve the balance equations at appropriately selected degrees of freedom (collocation).

[1] A word in notation is in order here. The superindex h is employed throughout the paper to denote finite element nodal quantities; bare symbols, on the other hand, are associated to reduced-order variables, that is, variables projected onto the reduced-order space. Likewise, FE and reduced-order dimensions are represented by upper-case and lower-case symbols, respectively. For instance, N and n denote the number of unknowns in the FE and reduced-order problems, respectively, whereas M and m represents the total number of integration points in the FE and reduced-order problem, respectively.

1.2.2. Integral approximation approaches

In a finite element context, $\boldsymbol{F}$ can be regarded, not only as a projection of a large vector into a reduced-order space ($\boldsymbol{F} = \boldsymbol{\Phi}^T \boldsymbol{F}^h$), but also as the result of integrating over the concerned domain $\Omega \subset \mathbb{R}^{\mathrm{d}}$ ($d = 2$ or 3) the corresponding reduced-order variable $\boldsymbol{f} = \boldsymbol{\Phi}^T \boldsymbol{f}^h$ ($\boldsymbol{f}^h : \Omega \to \mathbb{R}^N$), i.e.:

$$\boldsymbol{F} = \boldsymbol{\Phi}^T \int_\Omega \boldsymbol{f}^h \, d\Omega = \int_\Omega \boldsymbol{f} \, d\Omega. \tag{1}$$

Accordingly, the problem at hand can be also viewed as that of approximation of an *integral*, rather than approximation of a vector. In turn, this problem can be addressed by: (1) seeking a low-dimensional approximation of the integrand, or (2) approximating the integral itself as a weighted sum of the integrand evaluated at optimal sampling points.

1. *Interpolation of the integrand*. This type of approaches follows, in essence, the same procedure described for vector approximation approaches; the difference lies in that, rather than constructing an interpolant for the integral, in this case, it is the *integrand* that is subjected to approximation via interpolation, that is, if we make $\boldsymbol{f}(\boldsymbol{x}) \approx \sum_{g \in \mathbf{z}} \boldsymbol{R}_g(\boldsymbol{x}) \boldsymbol{f}(\boldsymbol{x}_g)$, where $\boldsymbol{R}_g$ ($g \in \mathbf{z}$) stands for the interpolation functions, then we can write

$$\boldsymbol{F} = \int_\Omega \boldsymbol{f} \, d\Omega \approx \sum_{g \in \mathbf{z}} \overbrace{\left(\int_\Omega \boldsymbol{R}_g \, d\Omega \right)}^{\boldsymbol{Q}_g} \boldsymbol{f}(\boldsymbol{x}_g) = \sum_{g \in \mathbf{z}} \boldsymbol{Q}_g \boldsymbol{f}(\boldsymbol{x}_g). \tag{2}$$

   Hence, the integral can be approximated as a sum of the product of *matrix* weights $\boldsymbol{Q}_g$ (which can be calculated offline) and the integrand evaluated at the interpolating points $\boldsymbol{x}_g \in \mathbf{z}$ —this set of points is to be chosen among the Gauss points of the underlying finite element mesh. This is the approach followed by [15–18].
2. *Cubature methods*. The last approaches to be discussed can be regarded as global *cubature methods*. To the best of the authors' knowledge, the first scheme of this type was proposed by An. et al. (2009) [19] in the context of computer graphics applications, and was recently introduced in computational mechanics circles by Farhat and co-workers [20,21]. Following the classical recipe of Gaussian quadrature of polynomial functions, An et al. advocate to approximate the integral as a finite sum of *positive scalar* weights $\{\omega_g\}_{g=1}^m$ times the integrand evaluated at appropriately chosen sampling points:

$$\boldsymbol{F} \approx \sum_{g=1}^{m} \omega_g \boldsymbol{f}(\bar{\boldsymbol{x}}_g). \tag{3}$$

   The strategy proposed by An et al. [19], and employed in successive refinements and extensions of the method [22–27], consists in determining, among the integration points of the FE mesh, a reduced set of points and associated *positive* weights so that the integration error is minimized over a set of representative samples of the integrand. The motivation behind constraining the weights to be *positive* scalars is that, in doing so, the contribution to the Jacobian matrix due to the nonlinear term inherits the spectral properties of its full-order counterpart. To put it simply, in a structural problem, if the FE stiffness matrix is symmetric and positive definite, so will be its reduced-order counterpart. It should be highlighted that this desirable attribute is not enjoyed by the other two approaches discussed above (at least in the context of standard Galerkin projection, see Ref. [28]). Indeed, interpolatory schemes ruin the symmetry and, depending on the location of the sampling points, may also destroy the positive definiteness of finite element stiffness matrices [17,14]. As a consequence, such schemes tend to be less robust than the finite element models they intend to approximate.

1.3. Goal of the paper and original contributions

All of the above described approaches are still in their infancy, and many theoretical and practical aspects still remain unaddressed. The goal of the present paper is to contribute to enhance and clarify some of these aspects in the particular case of *cubature-based ROMs*. Specifically, we propose a novel cubature approach – we call it for reasons that will become clear latter the *Empirical Cubature Method* – that addresses the following issues:

1. The cost associated to the greedy selection algorithm put forward by An et al. in their seminal paper [19] can prove unduly expensive in relatively large problems featuring high number of cubature points. To alleviate this (offline) computational cost, we propose to subject the integrand to a *dimensionality reduction process* (here, we use the Singular Value Decomposition, SVD). In doing so, the underlying minimization problem can be cast in terms of the first p dominant modes obtained from such a decomposition, rather that in terms of the set of P snapshots (in general, $p \ll P$, so the reduction in computational effort may prove substantial). Furthermore, we show that in problems in which the geometry is independent of the input parameters, the internal force modes can be alternatively obtained via an SVD of stress snapshots.
   (a) The above reduction in offline computational cost is partially made at the expense of an additional SVD, which is, in its own right, a very memory demanding operation. To overcome possible memory bottlenecks, we have concocted a *partitioned SVD* that precludes the necessity of processing the whole snapshot matrix (and that is amenable to parallelization).
2. There are certain parametric ROMs in which the nonlinear term $\boldsymbol{F}$ is by definition zero for any training configuration. This is the case, for instance, of the self-equilibrium problems appearing in multiscale hierarchical analysis, a research arena in which model reduction is particularly appealing, as evidences the intense research activity in the recent decade [17,29–36]. In this kind of problems, the minimization problem associated to the determination of optimal points and weights becomes ill-posed—it admits the trivial solution $\omega_g = 0$ $(g = 1, 2, \ldots, m)$ for any set of integration points. Based on an idea developed by the authors in Ref. [17] to deal with a similar problem in integrand-approximation approaches, we propose to decompose the snapshots into their component along the range of the integral operator, and its component along the nullspace. It is shown that this strategy, aside from eliminating the ill-posedness, leads to numerical cubature rules in which *the volume, V, of the physical domain is exactly integrated*, that is, $\sum_{g=1}^{m} \omega_g = V$.
3. Last but not least, we have devised a greedy selection algorithm able to find the absolute minimum (exact integration) of the optimization problem arising from the above mentioned decomposition. In particular, we have found that *exactly* integrating p integrand modes (plus the volume) requires $m = p + 1$ points. Furthermore, as opposed to existing approaches, the selection algorithm presented here need not solve, in all iterations, the expensive nonnegative least-squares problem to force the positiveness of the weights.

For the sake of concreteness, we have chosen as vehicle for exposing all the above ideas a classical nonlinear structural model in small strains, and consequently, the discussion that follows is couched in the terminology of structural finite element models: the state variable will be the vector of unknown nodal displacement and the nonlinear term $\boldsymbol{F}^h$ the vector of nodal internal forces. Nevertheless, the methodology is general and may be applied in many other contexts. For instance, in a nonlinear heat conduction problem, the state variable would be the vector of temperatures, and the nonlinear term $\boldsymbol{F}^h$ the heat-flux vector.

To make the paper reasonably self-contained, we throughly describe, not only the proposed cubature scheme (Section 4), and the resulting hyper-reduced order model (Section 5), but also the parameterized finite element model (Section 2), as well as the reduced-order model arising from the first reduction stage (Section 3). Guidelines of how to deal with Dirichlet boundary conditions and external forces are also provided in Section 3. Likewise, recipes for computing displacement reduced basis specially suited for the numerical examples presented in Section 6 are explained in Appendix C—these recipes are also original contributions of this work. The previously mentioned partitioned SVD, on the other hand, is not in the main body of the paper but is relegated to Appendix B.

2. Parametrized finite element model

The parametrized nonlinear equation whose complexity we wish to reduce is the standard finite element, semi-discrete motion equation in its Lagrangian form:

$$\boldsymbol{M}^h \ddot{\boldsymbol{d}}^h(\boldsymbol{\mu}) + \boldsymbol{F}^h(\boldsymbol{d}^h, \boldsymbol{d}_0^h; \boldsymbol{\mu}) = \boldsymbol{F}_{ext}^h(\boldsymbol{\mu}) - \boldsymbol{M}_0^h \ddot{\boldsymbol{d}}_0^h(\boldsymbol{\mu}). \tag{4}$$

Here, $\boldsymbol{d}^h \in \mathbb{R}^N$ and $\boldsymbol{d}_0^h \in \mathbb{R}^{N_0}$ denote the vectors of unknown and prescribed nodal displacements, respectively. The dimensions of these two vectors (i.e., the number of unrestricted and restricted degrees of freedom N and N_0, respectively) are assumed to remain constant along the analysis (constant Dirichlet boundaries). Superposed dots indicate material time derivative, i.e. $\ddot{\boldsymbol{d}}^h = \partial^2 \boldsymbol{d}^h / \partial t^2$. The set of input parameters is symbolically represented by $\boldsymbol{\mu}$,

Given $\boldsymbol{\mu} \in \mathcal{D}$, $\boldsymbol{g}^h(\boldsymbol{\mu}) : [0, T] \to \mathbb{R}^{N_0}$, $\boldsymbol{u}_0^h, \boldsymbol{v}_0^h \in \mathbb{R}^N$, $\boldsymbol{F}_{ext}^h(\boldsymbol{\mu}) : [0, T] \to \mathbb{R}^N$, find $\boldsymbol{d}^h : [0, T] \to \mathbb{R}^N$ such that

$$\boldsymbol{M}^h \ddot{\boldsymbol{d}}^h + \boldsymbol{F}^h = \boldsymbol{F}_{ext}^h - \boldsymbol{M}_0^h \ddot{\boldsymbol{d}}_0^h$$

where

$$\boldsymbol{F}^h = \int_\Omega \boldsymbol{f}^h \, d\Omega \approx \sum_{g=1}^{M} W_g \boldsymbol{f}^h(\boldsymbol{x}_g; \cdot)$$

subject to the Dirichlet boundary and initial conditions:

$$\boldsymbol{d}_0^h = \boldsymbol{g}^h, \qquad \boldsymbol{d}^h(0) = \boldsymbol{u}_0^h, \qquad \dot{\boldsymbol{d}}^h(0) = \boldsymbol{v}_0^h,$$

and the constitutive equations

$$\boldsymbol{\mathcal{H}}\Big(\boldsymbol{\sigma}, \boldsymbol{d}^h, \boldsymbol{d}_0^h, \boldsymbol{\xi}; \boldsymbol{\mu}\Big)\Big|_{\boldsymbol{x}_g} = \boldsymbol{0}, \quad g = 1, 2, \ldots, M$$

Box I. Statement of the finite element problem.

and the corresponding space by $\mathcal{D}$; this set of inputs may encompass variations of the prescribed boundary conditions, body forces, material parameters, etc... (occasionally, for notational brevity, the variable time t will be also included in $\boldsymbol{\mu}$). $\boldsymbol{M}^h \in \mathbb{R}^{N \times N}$ and $\boldsymbol{M}_0^h \in \mathbb{R}^{N \times N_0}$ are the mass matrices that give the inertial forces caused by acceleration of unrestricted and restricted DOFs, respectively; they both are assumed to be independent of $\boldsymbol{\mu}$.

On the other hand, $\boldsymbol{F}^h \in \mathbb{R}^N$ and $\boldsymbol{F}_{ext}^h \in \mathbb{R}^N$ denote the vectors of nodal internal and external forces, respectively. For simplicity of exposition, $\boldsymbol{F}_{ext}^h$ is assumed to be independent of the state variable. Thus, the previously mentioned "offending", nonlinear term requiring optimized cubature is the *vector of nodal internal forces* $\boldsymbol{F}^h$. The FE integration rule employed to evaluate this vector (e.g., element-wise Gauss quadrature) will be characterized by the pairs $\{\boldsymbol{x}_g, W_g\}_{g=1}^M, \boldsymbol{x}_g \in \Omega$ being the position of the gth integration point, W_g the corresponding weight (this weight includes the Jacobian of the finite element containing the point), and M the total number of integration points; accordingly, we can compactly write

$$\boldsymbol{F}^h = \int_\Omega \boldsymbol{f}^h \, d\Omega \approx \sum_{g=1}^{M} W_g \boldsymbol{f}^h(\boldsymbol{x}_g) \tag{5}$$

where $\boldsymbol{f}^h(\boldsymbol{x}_g) : \mathcal{D} \to \mathbb{R}^N$ is the (sparse) internal force vector at the gth integration point. In a small strain setting, $\boldsymbol{f}^h(\boldsymbol{x}_g) = \boldsymbol{B}^{h\,T}(\boldsymbol{x}_g)\boldsymbol{\sigma}(\boldsymbol{x}_g)$, $\boldsymbol{B}^h(\boldsymbol{x}_g) \in \mathbb{R}^{N \times s}$ being the classical (global) strain–displacement finite element matrix at point $\boldsymbol{x}_g$, and $\boldsymbol{\sigma} \in \mathbb{R}^s$ the stress vector ($s = 4, 6$, for plane stress/strain problems, and 3D problems, respectively). The nonlinearity between $\boldsymbol{f}^h$ and the state variable $\boldsymbol{d}^h$ may be of geometric nature (large strains) and/or material nature. Lastly, the constitutive relationship between the stress vector and both $\boldsymbol{\mu}$ and the history of deformation at each integration point will be symbolically represented here by

$$\boldsymbol{\mathcal{H}}\Big(\boldsymbol{\sigma}, \boldsymbol{d}^h, \boldsymbol{d}_0^h, \boldsymbol{\xi}; \boldsymbol{\mu}\Big)\Big|_{\boldsymbol{x}_g} = \boldsymbol{0}, \quad g = 1, 2, \ldots, M \tag{6}$$

where $\boldsymbol{\xi}$ stands for the vector of internal variables. For future references, the statement of the finite element problem is summarized in Box I.

3. First reduction stage

3.1. Unknown nodal displacements

The whole idea of model reduction relies on the premise that, for any input parameter $\boldsymbol{\mu} \in \mathcal{D}$, the displacement solution can be approximated by a set of n linearly independent basis vectors $\boldsymbol{\Phi}_i \in \mathbb{R}^N$ ($i = 1, 2, \ldots, n$), with $n \ll N$,

that is

$$\boldsymbol{d}^h(\boldsymbol{\mu}) \approx \boldsymbol{\Phi}\boldsymbol{d}(\boldsymbol{\mu}) \tag{7}$$

where $\boldsymbol{\Phi} = \begin{bmatrix}\boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \cdots & \boldsymbol{\Phi}_n\end{bmatrix}$ is the displacement basis matrix, and $\boldsymbol{d} \in \mathbb{R}^n$ the vector of (unknowns) reduced displacements—physically, the vector containing the amplitude of each displacement mode. There are various procedures for computing the basis matrix $\boldsymbol{\Phi}$. The most common one consists in, firstly, solving the full-order problem for representative values of the input parameters $\{\boldsymbol{\mu}^j\}_{j=1}^P$ ($\boldsymbol{\mu}^j \in \mathcal{D}$); then collecting the corresponding solutions in a snapshot matrix:

$$\boldsymbol{X}_d := \begin{bmatrix}\boldsymbol{d}^h(\boldsymbol{\mu}^1) & \boldsymbol{d}^h(\boldsymbol{\mu}^2) & \cdots & \boldsymbol{d}^h(\boldsymbol{\mu}^P)\end{bmatrix}; \tag{8}$$

and, finally, extracting the dominant modes by processing this matrix using any type of dimensionality reduction. If the standard Singular Value Decomposition (SVD) is the strategy of choice, then $\boldsymbol{X}_d \approx \boldsymbol{\Phi}\boldsymbol{\Sigma}_\Phi\boldsymbol{V}_\Phi^T$, $\boldsymbol{\Sigma}_\Phi$ and $\boldsymbol{V}_\Phi \in \mathbb{R}^{P\times n}$ being the truncated matrices of singular values and right singular vectors, respectively. More sophisticated dimensionality reduction techniques can, of course, be used. For instance, in the example shown in Section 6.1, we use the elastic/inelastic SVD proposed by the authors in [17], that consists, roughly, in decomposing the ensemble of solutions in those obtained in the elastic range, and those obtained when inelastic processes come into play, and then perform separated SVDs for each ensemble. Likewise, in the example of Section 6.2, the basis matrix for displacements includes both SVD modes and the dominant natural vibration modes of the structure. Further details concerning these two alternative strategies are provided in Appendix C.

3.2. Input vectors

To construct a ROM whose complexity is completely independent of the size of the underlying FE mesh, it is necessary[2] to replace by low-rank representations, not only the solution vector $\boldsymbol{d}^h$, but also the *input* vectors of prescribed displacements ($\boldsymbol{d}_0^h$) and external forces[3] ($\boldsymbol{F}_{ext}^h$). In fact, these two vectors usually admit *exact* approximations by linear combination of a few spatial basis vectors: if prescribed displacements are, say, uniform in space, just one spatial mode would suffice; if the spatial variation is linear, then two spatial modes would be needed, and so on. The coefficients in the linear combinations, on the other hand, can be obtained by *interpolation*. For instance, for the nodal vector of prescribed displacements, we can write (assuming an *exact* approximation)

$$\boldsymbol{d}_0^h(\boldsymbol{\mu}) = (\boldsymbol{\Xi}\boldsymbol{\Xi}_{\mathbf{b}}^{-1})\boldsymbol{d}_0^h\Big|_{\mathbf{b}}(\boldsymbol{\mu}) \tag{9}$$

where $\boldsymbol{\Xi} \in \mathbb{R}^{N_0\times n_0}$ is the corresponding basis matrix, $\mathbf{b} \subseteq \{1, 2, \ldots, N_0\}$ a set of n_0 admissible interpolation indices (admissible here means that the block matrix of $\boldsymbol{\Xi}$ corresponding to rows $\mathbf{b} = \{\mathrm{b}_1, \mathrm{b}_2, \ldots, \mathrm{b}_{n_0}\}$, denoted by $\boldsymbol{\Xi}_{\mathbf{b}}$, is invertible), and $\boldsymbol{d}_0^h\big|_{\mathbf{b}}$ the entries of $\boldsymbol{d}_0^h$ corresponding to indices $\mathbf{b}$. By introducing the variables

$$\boldsymbol{\Phi}_0 := \boldsymbol{\Xi}\boldsymbol{\Xi}_{\mathbf{b}}^{-1} \qquad\qquad \boldsymbol{d}_0 := \boldsymbol{d}_0^h\Big|_{\mathbf{b}}. \tag{10}$$

Eq. (9) becomes expressible in a format similar to the approximation of the unknown displacement in Eq. (7), i.e.:

$$\boldsymbol{d}_0^h(\boldsymbol{\mu}) = \boldsymbol{\Phi}_0\boldsymbol{d}_0(\boldsymbol{\mu}). \tag{11}$$

The basis matrix $\boldsymbol{\Xi}$ can be obtained by the same procedure used for the basis matrix of $\boldsymbol{d}^h$, that is, by collecting the vector of prescribed displacements for the training input parameters in a single matrix $\boldsymbol{X}_{d0} \in \mathbb{R}^{N_0\times P}$, and then applying the SVD to get the matrix of left singular vectors. Once this matrix is at one's disposal, the set of interpolation indices $\mathbf{b}$ can be easily determined by means of, for instance, the *Discrete Empirical Interpolation Method* (DEIM) [4].

[2] In many problems, computational savings achieved by compressing input vectors may be negligible in comparison with the savings obtained from approximating unknown displacement and internal forces, and therefore, this operation may be skipped.

[3] For simplicity of exposition, we have assumed that the initial conditions $\boldsymbol{u}_0^h$ and $\boldsymbol{v}_0^h$ are independent of $\boldsymbol{\mu}$. Should such conditions be dependent on $\boldsymbol{\mu}$, the same procedure explained in the following is to be applied to $\boldsymbol{u}_0^h$ and $\boldsymbol{v}_0^h$.

The interpolant of $\boldsymbol{F}^h_{ext}$ can be constructed following the same approach, i.e, by making

$$\boldsymbol{F}^h_{ext} = (\boldsymbol{\Theta}\,\boldsymbol{\Theta}_{\mathbf{c}}^{-1})\,\boldsymbol{F}^h_{ext}\Big|_{\mathbf{c}} \tag{12}$$

$\boldsymbol{\Theta} \in \mathbb{R}^{N\times n_f}$ and $\mathbf{c} \subset \{1, 2, \ldots, N\}$ being the corresponding basis matrix and set of interpolation points, respectively.

3.3. Projection onto the reduced-order space

Introducing expressions (7) and (11) into the FE balance equation (4), and multiplying by $\boldsymbol{\Phi}^T$ (Galerkin projection), we get

$$(\boldsymbol{\Phi}^T\boldsymbol{M}^h\,\boldsymbol{\Phi})\ddot{\boldsymbol{d}} + \boldsymbol{\Phi}^T\boldsymbol{F}^h = \boldsymbol{\Phi}^T\boldsymbol{F}^h_{ext} - (\boldsymbol{\Phi}^T\boldsymbol{M}^h_0\,\boldsymbol{\Phi}_0)\ddot{\boldsymbol{d}}_0. \tag{13}$$

Defining the *reduced mass matrices* $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{M}_0 \in \mathbb{R}^{n\times n_0}$ by

$$\boldsymbol{M} := \boldsymbol{\Phi}^T\boldsymbol{M}^h\,\boldsymbol{\Phi} \qquad\qquad \boldsymbol{M}_0 := \boldsymbol{\Phi}^T\boldsymbol{M}^h_0\,\boldsymbol{\Phi}_0 \tag{14}$$

and the *reduced vector of internal and external forces* ($\boldsymbol{F} \in \mathbb{R}^n$ and $\boldsymbol{F}_{ext} \in \mathbb{R}^n$, respectively) by

$$\boldsymbol{F} = \boldsymbol{\Phi}^T\boldsymbol{F}^h \qquad\qquad \boldsymbol{F}_{ext} = \boldsymbol{\Phi}^T\boldsymbol{F}^h_{ext} \tag{15}$$

balance equation (13) can be rewritten as[4]

$$\boldsymbol{M}\ddot{\boldsymbol{d}} + \boldsymbol{F} = \boldsymbol{F}_{ext} - \boldsymbol{M}_0\ddot{\boldsymbol{d}}_0. \tag{16}$$

Likewise, the initial and boundary conditions in the ROM become

$$\boldsymbol{d}_0 = \boldsymbol{g}^h_{\mathbf{b}}(\boldsymbol{\mu}), \qquad \boldsymbol{d}(0) = \boldsymbol{u}_0, \qquad \dot{\mathbf{d}}(0) = \boldsymbol{v}_0 \tag{17}$$

where $\boldsymbol{u}_0 = \boldsymbol{\Phi}^T\boldsymbol{u}^h_0$ and $\boldsymbol{v}_0 = \boldsymbol{\Phi}^T\boldsymbol{v}^h_0$.

Lastly, substitution of Eq. (12) into the second equation in (15) leads to the low-dimensional representation of the reduced external force vector:

$$\boldsymbol{F}_{ext} = \boldsymbol{R}_{ext}\,\boldsymbol{F}^h_{ext}\Big|_{\mathbf{c}}, \qquad \text{where} \qquad \boldsymbol{R}_{ext} = \boldsymbol{\Phi}^T(\boldsymbol{\Theta}\,\boldsymbol{\Theta}_{\mathbf{c}}^{-1}). \tag{18}$$

3.4. Internal forces

The only term in Eq. (16) that still remains dependent on the complexity of the FE mesh (through the number of FE integration points M) is the reduced vector of internal forces $\boldsymbol{F} \in \mathbb{R}^n$; indeed, multiplying Eq. (5) by $\boldsymbol{\Phi}^T$ leads to

$$\boldsymbol{F} = \int_{\Omega} \boldsymbol{\Phi}^T\boldsymbol{f}^h\; d\Omega \approx \sum_{g=1}^{M} W_g\boldsymbol{f}(\boldsymbol{x}_g;\cdot), \tag{19}$$

where $\boldsymbol{f} := \boldsymbol{\Phi}^T\boldsymbol{f}^h$. To culminate the model-order reduction process, thus, we have to develop a more efficient integration rule for $\boldsymbol{F}$—one that acknowledges the fact that displacements, and therefore, strains, stresses and internal forces reside now in low-dimensional spaces. Approximation consistency considerations dictate that this last step is to be carried out in a second reduction stage, one in which the "reference" model is no longer the finite element model, but the ROM described heretofore. For the reader's convenience and easy reference, the *offline* operations necessary to construct such a reduced-order problem, as well as the statement of the model itself, are summarized in Boxes II and III, respectively.

[4] It is at this point that the usefulness of our notational scheme becomes evident: to get the reduced-order balance equation (16), one just has to drop the superscript "h" in all the terms of the full-order equation (4) —bearing in mind the distinct definitions of the involved operators.

1. Solve the finite element problem (see Box I) for representative input parameters $\{\boldsymbol{\mu}^i\}_{i=1}^P$. Store the resulting vectors of unrestricted nodal displacements ($\boldsymbol{d}^h$), prescribed nodal displacements ($\boldsymbol{d}_0^h$) and external nodal forces ($\boldsymbol{F}_{ext}^h$) in the snapshot matrices $\boldsymbol{X}_d$, $\boldsymbol{X}_{d0}$, and $\boldsymbol{X}_{fext}$, respectively.
2. Apply the SVD (or any other dimensionality technique alike) to matrices $\boldsymbol{X}_d$, $\boldsymbol{X}_{d0}$, and $\boldsymbol{X}_{fext}$ to determine their respective basis matrices $\boldsymbol{\Phi} \in \mathbb{R}^{N\times n}$, $\boldsymbol{\Xi} \in \mathbb{R}^{N_0\times n_0}$ and $\boldsymbol{\Theta} \in \mathbb{R}^{N\times n_f}$.
3. Determine the interpolation indices $\mathbf{b} \subseteq \{1, 2, \ldots, N_0\}$ and $\mathbf{c} \subset \{1, 2, \ldots, N\}$ corresponding to basis matrices $\boldsymbol{\Xi}$ and $\boldsymbol{\Theta}$, respectively (using, for instance, the DEIM).
4. Compute the reduced-order matrices $\boldsymbol{\Phi}_0 = \boldsymbol{\Xi}\boldsymbol{\Xi}_{\mathbf{b}}^{-1}$, $\boldsymbol{M} = \boldsymbol{\Phi}^T\boldsymbol{M}^h\boldsymbol{\Phi}$, $\boldsymbol{M}_0 = \boldsymbol{\Phi}^T\boldsymbol{M}_0^h\boldsymbol{\Phi}_0$, $\boldsymbol{R}_{ext} = \boldsymbol{\Phi}^T\boldsymbol{\Theta}\boldsymbol{\Theta}_{\mathbf{c}}^{-1}$, $\boldsymbol{u}_0 = \boldsymbol{\Phi}^T\boldsymbol{u}_0^h$ and $\boldsymbol{v}_0 = \boldsymbol{\Phi}^T\boldsymbol{v}_0^h$.

Box II. Offline operations (first reduction stage).

Given $\boldsymbol{\mu} \in \mathcal{D}$, $\boldsymbol{g}_{\mathbf{b}}^h(\boldsymbol{\mu}) : [0, T] \to \mathbb{R}^{n_0}$, $\boldsymbol{u}_0, \boldsymbol{v}_0 \in \mathbb{R}^n$, $\boldsymbol{F}_{ext}^h\big|_{\mathbf{c}}(\boldsymbol{\mu}) : [0, T] \to \mathbb{R}^{n_f}$, find $\boldsymbol{d} : [0, T] \to \mathbb{R}^n$ such that

$$\boldsymbol{M}\ddot{\boldsymbol{d}} + \boldsymbol{F} = \boldsymbol{F}_{ext} - \boldsymbol{M}_0\ddot{\boldsymbol{d}}_0$$

where

$$\boldsymbol{F} = \int_\Omega \boldsymbol{f}\, d\Omega \approx \sum_{g=1}^{M} W_g \boldsymbol{f}(\boldsymbol{x}_g; \cdot), \qquad (\boldsymbol{f} = \boldsymbol{\Phi}^T\boldsymbol{f}^h)$$

and $\boldsymbol{F}_{ext} = \boldsymbol{R}_{ext}\boldsymbol{F}_{ext}^h\big|_{\mathbf{c}}$, subject to the Dirichlet boundary and initial conditions:

$$\boldsymbol{d}_0 = \boldsymbol{g}_{\mathbf{b}}^h, \qquad \boldsymbol{d}(0) = \boldsymbol{u}_0, \qquad \dot{\mathbf{d}}(0) = \boldsymbol{v}_0,$$

and to the constitutive equations

$$\boldsymbol{\mathcal{H}}(\boldsymbol{\sigma}, \boldsymbol{d}, \boldsymbol{d}_0, \boldsymbol{\xi}; \boldsymbol{\mu})|_{\boldsymbol{x}_g} = \mathbf{0}, \quad g = 1, 2, \ldots, M$$

Box III. Statement of the reduced-order problem (without approximation of internal forces).

4. Empirical cubature method (offline stage)

In this section, we describe the formalisms concerning the problem of approximating the integral of the reduced vector of internal forces using the proposed *Empirical Cubature Method* (ECM). To originate our considerations from a general standpoint, we first deal with the case in which a *continuous* representation of the integrand is assumed to be available. Then, we move to the more practical scenario of only having at one's disposal a discrete representation of the integrand—that is, when the integrand is only known at the integration points of the FE mesh.

4.1. Standard "optimized" cubature scheme

As pointed out in Section 1.3, our proposal for efficiently integrating the vector of reduced internal forces draws on the *optimized cubature scheme* proposed by An et al. in Ref. [19]. We begin our discussion by briefly summarizing the main ingredients of this scheme. Suppose that the reduced-order problem described in Box III is solved for the set of input parameters $\{\boldsymbol{\mu}^i\}_{i=1}^P$ (the same set employed in the first reduction stage). Let $f_I^j(\boldsymbol{x}) = f_I(\boldsymbol{x}, \boldsymbol{\mu}^j)$ denote the Ith component ($I = 1, 2, \ldots, n$) of the integrand at point $\boldsymbol{x} \in \Omega$ corresponding to the solution for input parameter $\boldsymbol{\mu}^j$ ($j = 1, 2, \ldots, P$). The idea put forward in Ref. [19] consists in approximating the integral of any $f_I^j : \Omega \to \mathbb{R}$ as the sum of positive, scalar weights multiplied by the function evaluated at appropriately chosen points, i.e.:

$$F_I^j = \int_\Omega f_I^j\, d\Omega \approx \sum_{g=1}^{m} \omega_g f_I^j(\bar{\boldsymbol{x}}_g). \tag{20}$$

The positions of the integration points $\mathcal{Z} = \{\bar{\boldsymbol{x}}_g\}_{g=1}^m$ and their associated positive weights $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_m]^T$ ($\omega_g > 0$) are determined so that the integration error for all components and training samples is minimized:

$$(\boldsymbol{\omega}, \mathcal{Z}) = \arg \min_{\boldsymbol{w}\in\mathbb{R}^m_+, \tilde{\boldsymbol{Z}}_g\in\Omega} \sqrt{\sum_{I=1}^{n}\sum_{j=1}^{P}(e_I^j)^2} \tag{21}$$

where

$$e_I^j := \sum_{g=1}^{m} \omega_g f_I^j(\bar{\boldsymbol{x}}_g) - \int_\Omega f_I^j \, d\Omega. \tag{22}$$

Minimization problem (21) can be cast in matrix format as follows

$$(\boldsymbol{\omega}, \mathcal{Z}) = \arg \min_{\boldsymbol{w}\in\mathbb{R}^m_+, \tilde{\boldsymbol{Z}}_g\in\Omega} \|\boldsymbol{J}_{\tilde{\boldsymbol{Z}}}\boldsymbol{w} - \boldsymbol{b}\| \tag{23}$$

where $\|\cdot\|$ stands for the standard Euclidean norm, and

$$\boldsymbol{J}_{\mathcal{Z}} := \begin{bmatrix} \boldsymbol{f}^1(\bar{\boldsymbol{x}}_1) & \boldsymbol{f}^1(\bar{\boldsymbol{x}}_2) & \cdots & \boldsymbol{f}^1(\bar{\boldsymbol{x}}_m) \\ \boldsymbol{f}^2(\bar{\boldsymbol{x}}_1) & \boldsymbol{f}^2(\bar{\boldsymbol{x}}_2) & \cdots & \boldsymbol{f}^2(\bar{\boldsymbol{x}}_m) \\ \cdots & \cdots & \vdots & \cdots \\ \boldsymbol{f}^P(\bar{\boldsymbol{x}}_1) & \boldsymbol{f}^P(\bar{\boldsymbol{x}}_2) & \cdots & \boldsymbol{f}^P(\bar{\boldsymbol{x}}_m) \end{bmatrix}, \qquad \boldsymbol{b} := \begin{bmatrix} \int_\Omega \boldsymbol{f}^1 \, d\Omega \\ \int_\Omega \boldsymbol{f}^2 \, d\Omega \\ \vdots \\ \int_\Omega \boldsymbol{f}^P \, d\Omega \end{bmatrix}, \qquad \boldsymbol{f}^j = \begin{bmatrix} f_1^j \\ f_2^j \\ \vdots \\ f_n^j \end{bmatrix}. \tag{24}$$

Notice that the gth column of $\boldsymbol{J}_{\mathcal{Z}} \in \mathbb{R}^{nP\times m}$ contains the value of the integrand at the gth integration point for all training samples; the vector $\boldsymbol{b} \in \mathbb{R}^{nP}$, on the other hand, is formed by arranging the *exact* integral of each function in a single column.

4.2. Dimensional reduction of the integrand

Problem (23) represents a combinatorial, nonnegative minimization problem whose complexity depends, among other factors, on the total number of samples of the integrand (P). Especially in dynamic and/or history dependent problems, this number may prove inordinately high, for, in such cases, the time variable is considered as an input parameter, and, hence, one has to store the solution at all time steps, or a representative set of time steps. To alleviate the intrinsic computational burden of solving this optimization problem, we propose to *subject the integrand to a dimensionality reduction process* – in a manner similar to that explained for the displacement vectors in the first reduction stage –, and then pose the minimization problem (23) in terms of the (orthogonal) basis functions arising from this process, rather than in terms of the integrand itself.

Suppose that using some dimensionality reduction technique, we determine a set of $p \ll P$ basis functions for f_I^j so that

$$f_I^j(\boldsymbol{x}) \approx \sum_{i=1}^{p} \Lambda_i(\boldsymbol{x}) c_{iI}^j, \qquad (j = 1, 2, \ldots, P;\ \ I = 1, 2, \ldots, n) \tag{25}$$

where $\Lambda_i : \Omega \to \mathbb{R}$ stands for the ith basis function, and $c_{iI}^j \in \mathbb{R}$ is the corresponding coefficient in the approximation. Introducing the above into the expression for the integration error (22), we get

$$e_I^j = \left(\sum_{g=1}^{m} \omega_g \Lambda_i(\bar{\boldsymbol{x}}_g) - \int_\Omega \Lambda_i \, d\Omega \right) c_{iI}^j. \tag{26}$$

Since the coefficients are independent of the position, it follows that minimizing the integration error e_I^j over all training samples and components is equivalent to *minimizing the error incurred in approximating the integral of the*

basis functions Λ_i $(i = 1, 2, \ldots, p)$. This consideration leads to a minimization problem similar to problem (23), but in which the matrices $\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}}$ and $\boldsymbol{b}$ are *defined in terms of the orthogonal basis functions* Λ_i $(i = 1, 2, \ldots, p)$ – rather that in terms of the samples of the integrand f_I^j –, i.e.:

$$\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}} := \begin{bmatrix} \Lambda_1(\bar{\boldsymbol{x}}_1) & \Lambda_1(\bar{\boldsymbol{x}}_2) & \cdots & \Lambda_1(\bar{\boldsymbol{x}}_m) \\ \Lambda_2(\bar{\boldsymbol{x}}_1) & \Lambda_2(\bar{\boldsymbol{x}}_2) & \cdots & \Lambda_2(\bar{\boldsymbol{x}}_m) \\ \cdots & \cdots & \vdots & \cdots \\ \Lambda_p(\bar{\boldsymbol{x}}_1) & \Lambda_p(\bar{\boldsymbol{x}}_2) & \cdots & \Lambda_p(\bar{\boldsymbol{x}}_m) \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} \int_\Omega \Lambda_1 \, d\Omega \\ \int_\Omega \Lambda_2 \, d\Omega \\ \vdots \\ \int_\Omega \Lambda_p \, d\Omega \end{bmatrix}. \tag{27}$$

Thus, in subjecting the integrand to a dimensionality reduction process, the number of rows of both $\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}}$ and $\boldsymbol{b}$ decreases from $P\,n$ to p (where $p \ll P$), and, consequently, so does the computational cost associated to solving the minimization problem.

4.3. Basis functions for the integrand: continuous case

4.3.1. Shortcoming of standard approaches

We now face the problem of developing a general strategy for determining the basis functions $\Lambda_i : \Omega \to \mathbb{R}$ for the integrand. In principle, one may use the standard strategy of directly applying the POD or techniques alike to the ensemble of snapshot of the integrand function $\{f_I^j(\boldsymbol{x})\}_{j=1}^P$. However, as pointed out in the introductory section, this strategy renders the minimization problem (23) ill-posed in cases in which the integral of such samples is zero for all values of the input parameters,[5] i.e, when

$$\int_\Omega f_I^j \, d\Omega = \boldsymbol{0} \tag{28}$$

for all $j = 1, 2, \ldots, P$ and $I = 1, 2, \ldots, n$. Indeed, since the basis functions are linear combination of snapshots, their integrals will be also zero, and, consequently, $\boldsymbol{b} = \boldsymbol{0}$. Under such circumstances, minimization problem (23) becomes

$$(\boldsymbol{\omega}, \boldsymbol{\mathcal{Z}}) = \arg \min_{\boldsymbol{w} \geq 0, \tilde{\boldsymbol{\mathcal{Z}}}_g \in \Omega} \|\boldsymbol{J}_{\tilde{\boldsymbol{\mathcal{Z}}}} \boldsymbol{w}\|, \tag{29}$$

which possesses the trivial solution $\boldsymbol{\omega} = 0$ (for any set of integration points). Of course, this ill-posedness may be simply alleviated by replacing the non-negative constraint by each ω_g being greater than a certain (small) tolerance. However, the quality of this solution becomes sensible to the value of the chosen tolerance and, furthermore, it leads to integration schemes requiring more integration points than strictly necessary.

4.3.2. Expanded basis approach (EBA)

The approach adopted here to eliminate the above-mentioned ill-posedness is to use an "expanded basis" (in a sense that will become clear in the ensuing discussion) for the nonlinear term. This idea was introduced by the authors in Ref. [17] for dealing with problems of similar nature appearing in the context of schemes based on interpolation of the integrand. The gist of this general idea goes as follows. Suppose that the equations arising from the Galerkin projection (first reduction stage) are compactly written as

$$\boldsymbol{A}(\boldsymbol{s}) = \boldsymbol{D}, \tag{30}$$

where $\boldsymbol{s}$ is the nonlinear term we wish to approximate, $\boldsymbol{A}$ is a linear operator independent of both the input parameter and the state variable, and $\boldsymbol{D} \in \mathbb{R}^n$ the sum of the remaining terms of the governing equation. Let us assume for the

[5] This situation is encountered in quasi-static problems with no external forces (only prescribed displacements).

sake of argument that $\boldsymbol{s} \in \mathbb{R}^M$, and that, therefore, $\boldsymbol{A}$ admits a matrix representation, denoted also by $\boldsymbol{A}$. Then suppose that $\boldsymbol{s}$ in Eq. (30) is replaced by the approximation $\boldsymbol{s} \approx \sum_{i=1}^{p} \boldsymbol{\Psi}_i c_i = \boldsymbol{\Psi c}$:

$$\boldsymbol{A}\left(\sum_{i=1}^{p} \boldsymbol{\Psi}_i c_i\right) = \boldsymbol{D} \quad \Rightarrow \quad (\boldsymbol{A}\,\boldsymbol{\Psi})\boldsymbol{c} = \boldsymbol{D} \tag{31}$$

$\{\boldsymbol{\Psi}_i\}_{i=1}^{p}$ being a set of basis vectors (and $\boldsymbol{\Psi} \in \mathbb{R}^{M\times p}$ the corresponding basis matrix). Expression (31) represents a system of n scalar equations that are *linear* in the coefficients $c_i = c_i(\boldsymbol{\mu}, \boldsymbol{d})$, although nonlinear, in general, in the state variable $\boldsymbol{d} \in \mathbb{R}^n$. It may be inferred from this fact, thus, that a *necessary* condition for Eq. (31) to have a unique solution in the state variable $\boldsymbol{d}$ is that these n equations must be all linearly independent.[6] This implies that the rank of the product of $\boldsymbol{A}$ times the basis matrix $\boldsymbol{\Psi}$ must be equal to the number of unknowns n, i.e. $\text{rank}(\boldsymbol{A}\,\boldsymbol{\Psi}) = n$.

This necessary condition poses certain restrictions on the number and form of employed basis functions $\{\boldsymbol{\Psi}_i\}_{i=1}^{p}$. One of such restrictions is obvious: the number of modes for $\boldsymbol{s}$ must be equal or greater than the number of unknowns ($p \geq n$), on the grounds that[7]

$$\text{rank}(\boldsymbol{A}\,\boldsymbol{\Psi}) \leq \min(\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{\Psi})) = \min(n, p). \tag{32}$$

Another (less obvious) condition, articulated and formulated by the authors in Ref. [17], is that the *span of the basis vectors must cover entirely the range of the operator* $\boldsymbol{A}$ (denoted by $\mathcal{R}(\boldsymbol{A})$), or, to put it another way, the intersection of the range of the operator and the span of the chosen basis functions must be the range of the operator itself:

$$\mathcal{R}(\boldsymbol{A}) \cap \text{span}\{\boldsymbol{\Psi}_i\}_{i=1}^{p} = \mathcal{R}(\boldsymbol{A}). \tag{33}$$

The proof that this condition guarantees that $\text{rank}(\boldsymbol{A}\,\boldsymbol{\Psi}) = n$ follows from the (columnwise) orthogonal decomposition of $\boldsymbol{\Psi}$ into its components along the kernel (denoted by $\mathcal{N}(\boldsymbol{A})$) and range of $\boldsymbol{A}$:

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}^R + \boldsymbol{\Psi}^N, \text{ with} \qquad \boldsymbol{\Psi}_i^R \in \mathcal{R}(\boldsymbol{A}), \;\; \boldsymbol{\Psi}_i^N \in \mathcal{N}(\boldsymbol{A}). \tag{34}$$

Indeed, since $\boldsymbol{A}\,\boldsymbol{\Psi}^N = \boldsymbol{0}$, we have that $\text{rank}(\boldsymbol{A}\,\boldsymbol{\Psi}) = \text{rank}(\boldsymbol{A}\,\boldsymbol{\Psi}^R)$. Furthermore, if Eq. (33) holds, then $\text{rank}(\boldsymbol{\Psi}^R) = n$. By elementary properties of the rank of product of matrices (see for instance Ref. [37]), if follows finally that $\text{rank}(\boldsymbol{A}\,\boldsymbol{\Psi}^R) = \text{rank}(\boldsymbol{A}) = n$.

If the basis functions are simply taken as a linear combination of snapshots of the concerned term $\boldsymbol{s}$ evaluated at the solution, satisfaction of (33) becomes contingent on the particular form of the right-hand side $\boldsymbol{D}$. The worst conceivable scenario for its fulfillment is when $\boldsymbol{D} = \boldsymbol{0}$ for any input parameter $\boldsymbol{\mu} \in \mathcal{D}$—which, incidentally, is the case at hand. Indeed, in such circumstances, any $\boldsymbol{f}(\boldsymbol{\mu}^j)$ lies in the kernel of $\boldsymbol{A}$, and therefore, the basis functions themselves will reside in such a subspace; since $\mathcal{N}(\boldsymbol{A})$ and $\mathcal{R}(\boldsymbol{A})$ are orthogonal subspaces, it follows that $\mathcal{R}(\boldsymbol{A}) \cap \text{span}\{\boldsymbol{\Psi}_i\}_{i=1}^{p} = \emptyset$—hence the ill-posedness.

The proposal advocated by the authors in [17] to guarantee that condition (33) is invariably observed – regardless of the particular form of the right-hand side $\boldsymbol{D}$ – is to construct the set of basis matrix by *expanding* the reduced basis set for the nonlinear term $\boldsymbol{s}$ with a basis for $\mathcal{R}(\boldsymbol{A})$:

$$\{\overbrace{\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_p}^{\text{Basis for } \boldsymbol{s}}, \overbrace{\boldsymbol{\Upsilon}_1, \boldsymbol{\Upsilon}_2, \ldots, \boldsymbol{\Upsilon}_n}^{\text{Basis for } \mathcal{R}(\boldsymbol{A})}\}. \tag{35}$$

However, in the problem addressed in Ref. [17], the right-hand side term is invariably zero (self-equilibrium problem), and under such conditions, the set of vectors in Eq. (35) is linearly independent (since $\boldsymbol{\Psi}_i \cdot \boldsymbol{\Upsilon}_j = \boldsymbol{0}$). To extend the method for cases in which $\boldsymbol{D} \neq \boldsymbol{0}$, we propose here to exploit the orthogonal decomposition induced by the operator and construct the basis set as

$$\{\overbrace{\hat{\boldsymbol{\Psi}}_1, \hat{\boldsymbol{\Psi}}_2, \ldots, \hat{\boldsymbol{\Psi}}_p}^{\text{Basis for } \hat{\boldsymbol{s}}}, \overbrace{\boldsymbol{\Upsilon}_1, \boldsymbol{\Upsilon}_2, \ldots, \boldsymbol{\Upsilon}_n}^{\text{Basis for } \mathcal{R}(\boldsymbol{A})}\} \tag{36}$$

[6] As argued in Ref. [17], in interpolatory approaches (when the coefficients c_i are determined by interpolation), failure to meet this necessary condition is conducive to rank deficient Jacobian matrices.

[7] It is assumed that Eq. (30) is well-posed, and therefore, $\boldsymbol{A}$ has full rank, $\text{rank}(\boldsymbol{A}) = n$.

where $\{\hat{\boldsymbol{\Psi}}\}_{i=1}^{p}$ is a reduced basis for the orthogonal projection of $\boldsymbol{s}$ onto the kernel of $\boldsymbol{A}$ (denoted by $\hat{\boldsymbol{s}}$). In doing so, $\hat{\boldsymbol{\Psi}}_j \in \mathcal{N}(\boldsymbol{A})$, and hence $\langle \hat{\boldsymbol{\Psi}}_j, \boldsymbol{\Upsilon}_i \rangle = 0$ (here, $\langle \cdot, \cdot \rangle$ symbolizes the scalar product employed in the orthogonal decomposition of the pertinent space in $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{N}(\boldsymbol{A})$).

4.3.3. Application of the EBA to the cubature problem

Let us now discuss the application of the method described in the foregoing to the problem at hand, which is to approximate $f_I : \mathcal{D} \times \Omega \to \mathbb{R}$ in equation

$$\int_\Omega f_I \, d\Omega = F_I, \qquad I = 1, 2, \ldots, n. \tag{37}$$

Comparison of the preceding equation with Eq. (30) indicates that, in this case, it is the *integral* itself that plays the role of the linear operator $\boldsymbol{A}$. Thus, for purposes of constructing the reduced basis of f_I, we shall regard the integral as an operator $\mathcal{I}$ that maps elements of the space of integrable functions on the domain $\Omega \subset \mathbb{R}^{\mathrm{d}}$ (denoted by $\mathcal{V}$) to $\mathbb{R}$. Next we introduce the orthogonal decomposition, in the L_2-norm, of $\mathcal{V}$ into the range and kernel of $\mathcal{I}$, i.e.: $\mathcal{V} = \mathcal{N}(\mathcal{I}) \oplus \mathcal{R}(\mathcal{I})$. The kernel of the integral operator is formed by all integrable functions whose integral is zero:

$$\mathcal{N}(\mathcal{I}) = \left\{ g \neq 0 \in \mathcal{V} \,|\, \int_\Omega g \, d\Omega = 0 \right\}. \tag{38}$$

On the other hand, the range of $\mathcal{I}$ is formed by those functions of $\mathcal{V}$ that are orthogonal to $\mathcal{N}(\mathcal{I})$, that is

$$\mathcal{R}(\mathcal{I}) = \left\{ h \neq 0 \in \mathcal{V} \,|\, \forall g \in \mathcal{N}(\mathcal{I}), \int_\Omega g\, h \, d\Omega = 0 \right\}. \tag{39}$$

It readily follows from the preceding definition that the *range of the integral operator is the space of constant functions in* Ω; thus, the dimension of $\mathcal{R}(\mathcal{I})$ is equal to 1, and a basis set for this space is formed by any constant function over Ω. Likewise, it is immediate to see that the projection of any sample of the integrand f_I^j ($I = 1, 2, \ldots, n$, $j = 1, 2 \ldots, P$) onto $\mathcal{N}(\mathcal{I})$ can be calculated by subtracting to f_I^j its *average* value over Ω, i.e.:

$$\hat{f}_I^j = f_I^j - \frac{1}{V} \int_\Omega f_I^j \, d\Omega, \qquad I = 1, 2, \ldots, n; \;\; j = 1, 2, \ldots, P \tag{40}$$

where $V = \int_\Omega d\Omega$ is the volume of the domain. Therefore, the desired set of basis functions for the integrand is to be constructed as the union of a basis set for the zero-average snapshots (40) (obtained via the POD, for instance) and a constant function (for instance, $g(\boldsymbol{x}) = 1, \forall \boldsymbol{x} \in \Omega$)

$$\{\overbrace{\Lambda_1, \Lambda_2, \ldots, \Lambda_p}^{\text{Basis for } \hat{f}_I^j}, \overbrace{1}^{\text{Basis for } \mathcal{R}(\mathcal{I})}\}. \tag{41}$$

Let us now discuss the implications of using the above basis set in the formulation of the minimization problem (23). Introducing (41) in the expressions (27) of $\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}}$ and $\boldsymbol{b}$, we arrive at

$$\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}} = \begin{bmatrix} \hat{\boldsymbol{J}}_{\boldsymbol{\mathcal{Z}}} \\ \mathbf{1}^T \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} \hat{\boldsymbol{b}} \\ V \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ V \end{bmatrix} \tag{42}$$

where $\hat{\boldsymbol{J}}_{\boldsymbol{\mathcal{Z}}} \in \mathbb{R}^{p \times m}$ and $\hat{\boldsymbol{b}} \in \mathbb{R}^p$ are defined by

$$\hat{\boldsymbol{J}}_{\boldsymbol{\mathcal{Z}}} := \begin{bmatrix} \Lambda_1(\bar{\boldsymbol{x}}_1) & \Lambda_1(\bar{\boldsymbol{x}}_2) & \cdots & \Lambda_1(\bar{\boldsymbol{x}}_m) \\ \Lambda_2(\bar{\boldsymbol{x}}_1) & \Lambda_2(\bar{\boldsymbol{x}}_2) & \cdots & \Lambda_2(\bar{\boldsymbol{x}}_m) \\ \cdots & \cdots & \vdots & \cdots \\ \Lambda_p(\bar{\boldsymbol{x}}_1) & \Lambda_p(\bar{\boldsymbol{x}}_2) & \cdots & \Lambda_p(\bar{\boldsymbol{x}}_m) \end{bmatrix} \qquad \hat{\boldsymbol{b}} = \begin{bmatrix} \int_\Omega \Lambda_1 \, d\Omega \\ \int_\Omega \Lambda_2 \, d\Omega \\ \vdots \\ \int_\Omega \Lambda_p \, d\Omega \end{bmatrix} = \mathbf{0} \tag{43}$$

whereas $\mathbf{1}^T = [1, 1, \ldots, 1]$ (an all-ones row vector of dimension m). Substitution of the new definitions of $\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}}$ and $\boldsymbol{b}$ in the expression of the objective function in problem (23) leads to

$$\|\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}}\boldsymbol{\omega} - \boldsymbol{b}\|^2 = \|\hat{\boldsymbol{J}}_{\boldsymbol{\mathcal{Z}}}\boldsymbol{\omega}\|^2 + \|\mathbf{1}^T\boldsymbol{\omega} - V\|^2 = \|\hat{\boldsymbol{J}}_{\boldsymbol{\mathcal{Z}}}\boldsymbol{\omega}\|^2 + \left(\sum_{g=1}^{m}\omega_g - V\right)^2 . \tag{44}$$

Notice that the second term in the above objective function vanishes when the sum of the weights equal the volume of the domain. Therefore, aside from eliminating the ill-posedness of the standard approach, the proposed strategy leads to a minimization problem in which *the requirement that the sum of weights has to be as close as possible to the total volume of the domain is explicitly imposed*. It should be mentioned that neither the original cubature scheme of An. et al. [19], nor later related proposals (e.g. [20]), exhibits this remarkable feature.

Remark 4.1. It is interesting to note that the format exhibited by $\boldsymbol{J}_{\boldsymbol{\mathcal{Z}}}$ and $\boldsymbol{b}$ in the proposed approach (Eq. (42)) bears a close resemblance to their counterparts in classical Gaussian quadrature (see, for instance, page 143 in Ref. [38]). The reason of such a similarity is that they share a common attribute: the use of orthogonal basis functions. However, Gaussian formulae were originally conceived for one-dimensional domains (quadrature) and polynomial functions, while our approach can be applied to any multidimensional domain (cubature), and, furthermore, can deal with general, empirical basis functions, the qualifier empirical meaning "derived from computational experiments" [39]—hence the employed appellation Empirical Cubature Method.

4.4. Discrete formulation

In formulating the minimization problem (29), it was tacitly assumed that a continuous description of the integrand is available. However, in common finite element implementations, the value of the integrand is only calculated at the integration points of each finite element. For practical reasons, thus, it proves convenient to retrace the analysis described in the foregoing from a "discrete" perspective, i.e., treating the integrand as a spatially discrete variable.

The most immediate implication of the integrand being defined only at the integration points of the mesh is that the integral we wish to approximate can no longer be regarded as a linear operator that maps continuous functions into $\mathbb{R}$; rather, it should be viewed as an operator that maps vectors of $\mathbb{R}^M$ (recall that M is the number of FE integration points) to $\mathbb{R}$. The matrix representation of this operator can be inferred from the expression for the approximated integral of f_I as follows:

$$\begin{aligned} F_I = \int_\Omega f_I \, d\Omega \approx \sum_{g=1}^{M} W_g f_I(\boldsymbol{x}_g) = \sum_{g=1}^{M} \sqrt{W_g}(\sqrt{W_g} f_I(\boldsymbol{x}_g)) \\ = \sqrt{\boldsymbol{W}}^T \boldsymbol{\mathcal{F}}_I, \qquad\qquad I = 1, 2, \ldots, n \end{aligned} \tag{45}$$

where $\sqrt{\boldsymbol{W}} \in \mathbb{R}^M$ is defined by

$$\sqrt{\boldsymbol{W}} := \begin{bmatrix} \sqrt{W_1} & \sqrt{W_2} & \cdots & \sqrt{W_M} \end{bmatrix}^T , \tag{46}$$

and $\boldsymbol{\mathcal{F}}_I \in \mathbb{R}^M$ is formed by gathering in a single column vector the values of the integrand at all FE points (multiplied by the square root of each finite element integration weight)

$$\boldsymbol{\mathcal{F}}_I := \begin{bmatrix} \sqrt{W_1} f_I(\boldsymbol{x}_1) \\ \sqrt{W_2} f_I(\boldsymbol{x}_2) \\ \vdots \\ \sqrt{W_M} f_I(\boldsymbol{x}_M) \end{bmatrix} . \tag{47}$$

It readily follows from Eq. (45) that $\sqrt{\boldsymbol{W}}^T$ is but the *matrix representation* of the integral operator when the domain space $\mathbb{R}^M$ is endowed with the standard scalar product.

Remark 4.2. Notice that the definition of operator $\sqrt{\boldsymbol{W}}^T$ tacitly presupposes that all FE weights are positive. In view of this apparent limitation of the method (it excludes FE integration rules with negative weights), it may be wondered why not to define the operator by putting all the weights in the vector $\boldsymbol{\mathcal{F}}_I$ (that is, $(\boldsymbol{\mathcal{F}}_I)_g = W_g f_I(\boldsymbol{x}_g)$), or in the operator itself (so that $(\boldsymbol{\mathcal{F}}_I)_g = f_I(\boldsymbol{x}_g)$). The reason is that the discrete formulation (45) is, strictly speaking, the only one consistent with its continuous counterpart:

$$\|f_I\|^2_{L_2(\Omega)} = \int_\Omega f_I f_I \, d\Omega \approx \sum_{g=1}^{M} W_g f_I(\boldsymbol{x}_g) f_I(\boldsymbol{x}_g) = \|\boldsymbol{\mathcal{F}}_I\|^2 \tag{48}$$

i.e., the euclidean norm of $\boldsymbol{\mathcal{F}}_I$ is the actual approximation of the L_2 to norm of $f_I(\boldsymbol{x})$. Nevertheless, the procedure explained in the following discussion can be also applied when the other two (less restrictive) definitions of the discrete integral operator are adopted (in fact, in the event of meshes with fairly uniform element size, the three alternatives should yield similar results).

4.4.1. Basis matrices

Having defined the discrete version of the integral operator, the next step consists in determining the basis matrix for the nonlinear term $\boldsymbol{\mathcal{F}}_I$ $(I = 1, 2, \ldots, n)$. According to the previously described expanded basis approach, the basis matrix for $\boldsymbol{\mathcal{F}}_I$ is to be constructed as the union of a basis matrix for the range of $\sqrt{\boldsymbol{W}}^T$ (which is simply the space spanned by $\sqrt{\boldsymbol{W}}$) and a basis matrix, $\boldsymbol{\Lambda} \in \mathbb{R}^{M\times p}$, for the projection of $\boldsymbol{\mathcal{F}}_I$ onto the kernel of $\sqrt{\boldsymbol{W}}^T$, i.e:

$$\text{Expanded basis matrix} = \left[\boldsymbol{\Lambda}_1 \quad \boldsymbol{\Lambda}_2 \quad \cdots \quad \boldsymbol{\Lambda}_p \quad \sqrt{\boldsymbol{W}}\right]. \tag{49}$$

The starting point for computing $\boldsymbol{\Lambda}$ is the snapshot matrix of $\boldsymbol{\mathcal{F}}_I^j$ for all components $I = 1, 2, \ldots, n$ and all training configurations $j = 1, 2, \ldots, P$:

$$\boldsymbol{X}_{\mathcal{F}} = \left[\boldsymbol{\mathcal{F}}_1^1 \quad \cdots \quad \boldsymbol{\mathcal{F}}_n^1 \quad \boldsymbol{\mathcal{F}}_1^2 \quad \cdots \quad \boldsymbol{\mathcal{F}}_n^2 \quad \boldsymbol{\mathcal{F}}_1^P \quad \cdots \quad \boldsymbol{\mathcal{F}}_n^P\right]. \tag{50}$$

The projection of each column of $\boldsymbol{X}_{\mathcal{F}}$ onto $\mathcal{N}(\sqrt{\boldsymbol{W}}^T)$ can be calculated by subtracting its orthogonal projection onto $\mathcal{R}(\sqrt{\boldsymbol{W}}^T)$; this operation yields

$$\hat{\boldsymbol{\mathcal{F}}}_I^j := \boldsymbol{\mathcal{F}}_I^j - \frac{\sqrt{\boldsymbol{W}}}{\|\sqrt{\boldsymbol{W}}\|}\left(\frac{\sqrt{\boldsymbol{W}}^T}{\|\sqrt{\boldsymbol{W}}\|}\boldsymbol{\mathcal{F}}_I^j\right) = \begin{bmatrix} \sqrt{W_1}(f_I^j(\boldsymbol{x}_1) - F_I^j/V) \\ \sqrt{W_2}(f_I^j(\boldsymbol{x}_2) - F_I^j/V) \\ \vdots \\ \sqrt{W_M}(f_I^j(\boldsymbol{x}_M) - F_I^j/V) \end{bmatrix}, \tag{51}$$

which is essentially the same expression obtained in the continuous case (40) (the integrand at a given point minus the volume average), but with each entry multiplied by the square root of the corresponding FE integration weight. The required basis matrix can be finally determined by collecting all these vectors in a single matrix $\hat{\boldsymbol{X}}_{\mathcal{F}} \in \mathbb{R}^{M\times nP}$

$$\hat{\boldsymbol{X}}_{\mathcal{F}} = \left[\hat{\boldsymbol{\mathcal{F}}}_1^1 \quad \cdots \quad \hat{\boldsymbol{\mathcal{F}}}_n^1 \quad \hat{\boldsymbol{\mathcal{F}}}_1^2 \quad \cdots \quad \hat{\boldsymbol{\mathcal{F}}}_n^2 \quad \hat{\boldsymbol{\mathcal{F}}}_1^P \quad \cdots \quad \hat{\boldsymbol{\mathcal{F}}}_n^P\right] \tag{52}$$

and then applying a dimensionality reduction technique such as the SVD[8] to obtain an approximated basis matrix of rank p for the column space of $\hat{\boldsymbol{X}}_{\mathcal{F}}$, i.e.:

$$\hat{\boldsymbol{X}}_{\mathcal{F}} \approx \boldsymbol{\Lambda}\boldsymbol{\Sigma}_\Lambda \boldsymbol{V}_\Lambda^T \tag{53}$$

[8] It is interesting to note that, since each row of $\hat{\boldsymbol{X}}_{\mathcal{F}}$ is multiplied by the square root of the corresponding finite element weight, the minimization problem underlying the SVD is actually posed in the norm defined by the diagonal matrix of finite element weights (which, in turn, as we saw earlier, is the discrete counterpart of the L_2 norm).

where $\boldsymbol{\Sigma}_\Lambda$ and $\boldsymbol{V}_\Lambda^T$ are the matrices of singular values and right singular vectors, respectively, associated to the selected, dominant left singular vectors[9] $\boldsymbol{\Lambda} \in \mathbb{R}^{M\times p}$.

Remark 4.3. The standard SVD is a RAM-intensive operation, and, consequently, manipulating relatively large snapshot matrices may easily exhaust the memory capabilities of the computer at hand. To overcome this difficulty, we propose to adopt two complementary strategies. Firstly, in problems in which the geometry is input-parameter independent, the nonlinearity between the internal forces and displacements is concentrated on the stresses and, therefore, one may obtain the basis matrix for the internal forces from a basis matrix for the stresses themselves. This strategy is specially advantageous for relatively high number of reduced displacement modes (n), as the size of the snapshot stress matrix is independent of this reduced dimension.

For high number of training configurations, however, the above strategy may prove insufficient to lower the offline computational requirements to affordable levels. For such cases, we have devised a partitioned version of the SVD that precludes the necessity of manipulating the whole matrix, hence significantly diminishing memory requirements.

To preserve the continuity of the presentation, the detailed description of these two complementary methods is relegated to Appendices A and B, respectively.

4.4.2. Minimization problem

The statement of the discrete minimization problem that allows one to determine the optimized location of integration points and corresponding weights is quite similar to the continuous problem (23): find $\boldsymbol{\alpha} \in \mathbb{R}_+^m$ and $\mathbf{z} \in \mathbb{N}^m \subset \{1, 2, \ldots, M\}$ such that

$$(\boldsymbol{\alpha}, \mathbf{z}) = \arg \min_{\tilde{\boldsymbol{\alpha}}\geq \mathbf{0}, \tilde{\mathbf{z}}} \|\boldsymbol{J}_{\tilde{\mathbf{z}}}\tilde{\boldsymbol{\alpha}} - \boldsymbol{b}\|^2. \tag{54}$$

Note that, in this case, the goal is to choose a set of m points *among* the M FE integration points of the underlying finite element mesh $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M\}$ —that is, we have now a *combinatorial optimization problem*. Hereafter, the gth optimal point will be denoted by $\bar{\boldsymbol{x}}_g := \boldsymbol{x}_{\mathbf{z}_g}$.

Following the same logic that led to Eq. (42), $\boldsymbol{J}_\mathbf{z} \in \mathbb{R}^{p+1\times m}$ and $\boldsymbol{b} \in \mathbb{R}^{p+1}$ can be decomposed into the following block matrices

$$\boldsymbol{J}_\mathbf{z} = \begin{bmatrix} \boldsymbol{\Lambda}_\mathbf{z} & \sqrt{\boldsymbol{W}}_\mathbf{z} \end{bmatrix}^T, \qquad \boldsymbol{b} = \begin{bmatrix} \mathbf{0}^T & V \end{bmatrix}^T \tag{55}$$

($\boldsymbol{\Lambda}_\mathbf{z}$ and $\sqrt{\boldsymbol{W}}_\mathbf{z}$ denotes the block matrices of $\boldsymbol{\Lambda}$ and $\sqrt{\boldsymbol{W}}$, respectively, formed by the rows corresponding to the indices of the selected points $\mathbf{z} \subset \{1, 2, \ldots, M\}$). With this decomposition, the objective function is expressible as

$$\|\boldsymbol{J}_\mathbf{z}\boldsymbol{\alpha} - \boldsymbol{b}\|^2 = \|\boldsymbol{\Lambda}_\mathbf{z}^T\boldsymbol{\alpha}\|^2 + \|\sqrt{\boldsymbol{W}}_\mathbf{z}^T\boldsymbol{\alpha} - V\|^2. \tag{56}$$

Introduction of the preceding expression into problem (54) leads to the following minimization problem: find $\boldsymbol{\alpha} \in \mathbb{R}_+^m$ and $\mathbf{z} \in \mathbb{N}^m \subseteq \{1, 2 \ldots M\}$ such that

$$(\boldsymbol{\alpha}, \mathbf{z}) = \arg \min_{\tilde{\boldsymbol{\alpha}}\geq \mathbf{0}, \tilde{\mathbf{z}}} \left(\|\boldsymbol{\Lambda}_{\tilde{\mathbf{z}}}^T\tilde{\boldsymbol{\alpha}}\|^2 + \left(\sum_{g=1}^m \sqrt{W_{\tilde{\mathbf{z}}_g}}\tilde{\alpha}_g - V \right)^2 \right). \tag{57}$$

After solving for $\boldsymbol{\alpha}$, the sought-after weights can be computed as

$$\omega_g := \sqrt{W_g}\alpha_g, \qquad g = 1, 2, \ldots, m. \tag{58}$$

Observe that, as in the continuous case, the condition that the sum of weights should be as equal as possible to the volume of the domain appears explicitly in the objective function (57).

[9] Alternatively, instead of directly using as basis matrix the left singular vectors, one may use as basis matrix $\boldsymbol{\Lambda} \leftarrow \boldsymbol{\Lambda}\boldsymbol{\Sigma}_\Lambda$. This is equivalent to define the objective function in (57) in terms of the norm defined by such singular values.

4.5. Solution of the optimization problem

We turn now to the solution of the *combinatorial* optimization problem (54). In the field of combinatorial optimization, this type of problem is termed a *hard regressor selection* [40]: we are given a matrix $\boldsymbol{J} \in \mathbb{R}^{p+1\times M}$, whose columns are potential regressors, and a vector $\boldsymbol{b} \in \mathbb{R}^{p+1}$ is to be fit by a linear combination of $m < M$ columns of $\boldsymbol{J}$. The problem is to choose the subset of m columns to be used as well as the associated coefficients. One straightforward approach would be to check every possible combination, considering that, for a fixed set of columns $\mathbf{z}$, the optimal $\boldsymbol{\alpha}$ can be found by solving a nonnegative least-squares problem—using, for instance, the active set algorithm developed by Lawson and Hanson [41]. In principle, this should be done for each of the $\binom{M}{m}$ feasible combinations. Needless to say, since $M \gg m$ (for otherwise the whole reduction procedure would prove pointless), this brute force approach is not viable and recourse to *heuristic* methods, able to at least determine *sub-optimal solutions*, is to be made.

4.5.1. Greedy selection method

The heuristic method employed in the present work is described in the flowchart of Box IV. This method is based on the *greedy* algorithm put forward by An and co-workers in their seminal paper [19]. Two features distinguishes our algorithm from that in Ref. [19], namely: (1) the definition of the matrices $\boldsymbol{J}$ and $\boldsymbol{b}$ appearing in the objective function ($\boldsymbol{J}$ is a row-wise orthogonal matrix and all entries of $\boldsymbol{b}$ are zero but the last one, which is equal to the volume of the domain); and (2) the unrestricted least-squares appearing in Step 3. The reason why $\boldsymbol{\alpha}$ is tentatively computed by an unrestricted least-squares (LS) fit in Step 3 is that we have empirically observed that, in almost all iterations, this operation furnishes vectors $\boldsymbol{\alpha}$ with positive entries, hence precluding the need to solve the expensive nonnegative least-squares problem of Step 5.

4.5.2. Convergence to absolute minimum

Numerical experience shows that the algorithm in Box IV *invariably drives the objective function* $\|\boldsymbol{J}_{\mathbf{z}}\boldsymbol{\alpha} - \boldsymbol{b}\|$ *to zero when the number of selected points*[10] (i.e. colums) *equals the number of rows of* $\boldsymbol{J}$, that is, when

$$m_+ = p + 1. \tag{59}$$

This convergence property can be attributed to one of the distinguishing feature of the proposed cubature scheme, namely, that $\boldsymbol{J}$ is the transpose of an orthogonal basis matrix, and hence, it has full rank. When $m_+ = p+1$, the block matrix $\boldsymbol{J}_{\mathbf{z}}$ becomes square and, furthermore, since the algorithm selects linearly independent columns, *invertible*. This means that $\boldsymbol{\alpha} = \boldsymbol{J}_{\mathbf{z}}^{-1}\boldsymbol{b}$ and, consequently, $\|\boldsymbol{J}_{\mathbf{z}}\boldsymbol{\alpha} - \boldsymbol{b}\| = 0$.

The linear independence of the selected columns follows from the fact that, at Step 1, the algorithm identifies the new index $\mathbf{z}_k$ as that whose associated column is the *most positively parallel* to the current residual, for this is the column that will reduce the residual norm the most. This necessarily implies that $\boldsymbol{J}_{\mathbf{z}_k}$ must be linearly independent of the previously selected columns $\{\boldsymbol{J}_{\mathbf{z}_1}, \boldsymbol{J}_{\mathbf{z}_2}, \ldots, \boldsymbol{J}_{\mathbf{z}_{k-1}}\}$—a linearly dependent column carries redundant information and would not, thus, contribute to lower the residual.

5. Hyperreduced-order model

For completeness, the *offline steps* required for determining the set of integration points and their associated weights are summarized[11] in Box V. Likewise, the statement of the resulting "hyper-reduced" order problem is set forth in Box VI. It should be stressed that the only difference between this problem and the reduced-order problem shown previously in Section 3 (see Box III) is the evaluation of the integral of the reduced internal forces—now it requires computation of the integrand at solely $m \ll M$ points. As a consequence, the solving effort associated to the constitutive equations also diminishes by a factor M/m, as well as memory requirements for storing the history of internal variables.

[10] Notice that, at Step 6, points whose associated weights are identically zero are removed from $\mathbf{z}$. Hence, the cardinality of $\mathbf{z}$ at the end of each iteration coincides with the number of points with strictly positive weights (m_+).

[11] The reduced-order operators $\boldsymbol{\Phi}_0$, $\boldsymbol{M}$, $\boldsymbol{M}_0$, $\boldsymbol{R}_{ext}$ and $\boldsymbol{B}_g = \boldsymbol{B}_g^h\boldsymbol{\Phi}$ ($g = 1, 2, \ldots, M$) were already determined in the offline phase of the first reduction stage, see Box II, are, hence, need not be computed again.

DATA: $\boldsymbol{J} = \begin{bmatrix} \boldsymbol{\Lambda} & \sqrt{\boldsymbol{W}} \end{bmatrix}^T \in \mathbb{R}^{(p+1)\times M}$, $\boldsymbol{b} = \begin{bmatrix} \boldsymbol{0} & V \end{bmatrix}^T \in \mathbb{R}^{p+1}$, TOL, m
Initializations
– Set of integration points: $\mathbf{z} \leftarrow \emptyset$, Set of candidate points: $\mathbf{y} \leftarrow \{1, 2, \ldots, M\}$
– Nonzero components of $\boldsymbol{\alpha}$: $m_+ \leftarrow 0$, number of iterations $k \leftarrow 1$
– Residual vector: $\boldsymbol{r} \leftarrow \boldsymbol{b}$
while $\|\boldsymbol{r}\|/\|\boldsymbol{b}\| > TOL$ **AND** $m_+ <= m$ **do**
1. Compute new point i as: $i = \arg \max_{i \in \mathbf{y}} \tilde{\mathbf{J}}_{\mathbf{y}}^T \boldsymbol{r}/\|\boldsymbol{r}\|$, where $\tilde{\mathbf{J}}_j := \boldsymbol{J}_j/\|\boldsymbol{J}_j\|$
2. Move i from set $\mathbf{y}$ to set $\mathbf{z}$ ($\mathbf{z} \leftarrow \mathbf{z} \cup i$ and $\mathbf{y} \leftarrow \mathbf{y} \setminus i$)
3. Determine $\boldsymbol{\alpha}$ by (*unrestricted least-squares*):
$$\boldsymbol{\alpha} \leftarrow \left(\boldsymbol{J}_{\mathbf{z}}^T \boldsymbol{J}_{\mathbf{z}}\right)^{-1} \boldsymbol{J}_{\mathbf{z}}^T \boldsymbol{b}$$
4. If all entries of $\boldsymbol{\alpha}$ are nonnegative, go to step 7.
5. Determine $\boldsymbol{\alpha}$ by solving the *nonnegative least-squares problem*
$$\boldsymbol{\alpha} \leftarrow \arg \min_{\tilde{\boldsymbol{\alpha}} \geq \boldsymbol{0}} \|\boldsymbol{J}_{\mathbf{z}} \tilde{\boldsymbol{\alpha}} - \boldsymbol{b}\|^2$$
6. Set $\mathbf{z} \leftarrow \mathbf{z} \setminus \mathbf{z}_0$, (where $\mathbf{z}_0 \subset \mathbf{z} \mid \boldsymbol{\alpha}(\mathbf{z}_0) = 0$), $\boldsymbol{\alpha} \leftarrow \boldsymbol{\alpha}(\mathbf{z})$ and $\mathbf{y} \leftarrow \mathbf{y} \cup \mathbf{z}_0$.
7. Update the residual: $\boldsymbol{r} \leftarrow (\boldsymbol{b} - \boldsymbol{J}_{\mathbf{z}} \boldsymbol{\alpha})$.
8. Set $k \leftarrow k + 1$, $m_+ \leftarrow \text{card}(\mathbf{z})$
end
9. Compute the desired **integration weights** as
$$\omega_g = \sqrt{W}_{\mathbf{z}_g} \alpha_g, \qquad g = 1, 2, \ldots, m$$

Box IV. Empirical cubature method (offline stage). Greedy algorithm for computing an optimal set of integration points $z \subset \{1, 2, \ldots, M\}$ and . corresponding weights $\omega \in \mathbb{R}_+^m$.

1. Solve the reduced-order problem (without approximation of internal forces, see Box III) for the set of representative input parameters $\{\boldsymbol{\mu}^i\}_{i=1}^P$ (the same employed for the problem presented in Box III).
2. Store the components of the reduced internal forces at all integration points and all training configurations in the snapshot matrix $\boldsymbol{X}_{\mathcal{F}} \in \mathbb{R}^{M \times nP}$ (see Eq. (50)). Alternatively, one may store the stresses at the integration points, and then calculate a "compressed" $\boldsymbol{X}_{\mathcal{F}} \in \mathbb{R}^{M \times nq}$ by means of formula (A.5) in Appendix A. Likewise, store the FE integration weights (including the Jacobian) in a vector $\boldsymbol{W} \in \mathbb{R}^M$.
3. Compute the matrix of zero-integral snapshots $\hat{\boldsymbol{X}}_{\mathcal{F}}$ by applying formula (51) to each column of matrix $\boldsymbol{X}_{\mathcal{F}}$.
4. Determine an orthogonal basis matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{M \times p}$ for the column space of $\hat{\boldsymbol{X}}_{\mathcal{F}}$ as the p leading left singular vectors arising from the SVD of $\hat{\boldsymbol{X}}_{\mathcal{F}}$.
5. Construct the matrix $\boldsymbol{J} \in \mathbb{R}^{p+1 \times M}$ and $\boldsymbol{b} \in \mathbb{R}^{p+1}$ appearing in the cubature optimization problem as $\boldsymbol{J} = \begin{bmatrix} \boldsymbol{\Lambda} & \sqrt{\boldsymbol{W}} \end{bmatrix}^T$ and $\boldsymbol{b} = \begin{bmatrix} \boldsymbol{0}^T & V \end{bmatrix}^T$, where $V = \sum_{i=1}^M W_i$.
6. Determine the set of integration points $\mathbf{z} \in \mathbb{N}^m$ and their associated weights $\boldsymbol{\omega} \in \mathbb{R}_+^m$ by means of the algorithm described in Box IV.

Box V. Offline operations (second reduction stage).

5.1. Reconstruction of displacement, stress and strain fields

In practice, the output of interest in the hyper-reduced-order problem described in Box VI is rarely the vector of reduced displacements $\boldsymbol{d}$ *per se*, but rather a derived quantity of either the nodal displacements or the stresses at the Gauss points of the finite element mesh (or both). The vector of nodal displacements can be recovered by multiplying the reduced displacement vectors[12] $\boldsymbol{d}$ by the corresponding basis matrix, i.e., $\boldsymbol{d}^h \approx \boldsymbol{\Phi} \boldsymbol{d}$.

On the other hand, stresses (or any other related variable, such as the internal variables $\boldsymbol{\xi}$) at any integration point $\boldsymbol{x}_g$ ($g = 1, 2, \ldots, M$) of the finite element mesh can be, in principle, recovered by locally integrating in time the

[12] Recall that $\boldsymbol{d}$ corresponds to the unknown nodal displacements. Prescribed displacements are given by the expression $\boldsymbol{d}_0^h = \boldsymbol{\Phi}_0 \boldsymbol{d}_0$, where $\boldsymbol{\Phi}_0$ is defined in Eq. (10).

Given $\boldsymbol{\mu} \in \mathcal{D}$, $\boldsymbol{g}^h_{\mathbf{b}}(\boldsymbol{\mu}) : [0, T] \to \mathbb{R}^{n_0}$, $\boldsymbol{u}_0, \boldsymbol{v}_0 \in \mathbb{R}^n$, $\boldsymbol{F}^h_{ext}\big|_{\mathbf{c}}(\boldsymbol{\mu}) : [0, T] \to \mathbb{R}^{n_f}$, find $\boldsymbol{d} : [0, T] \to \mathbb{R}^n$ such that

$$\boldsymbol{M}\ddot{\boldsymbol{d}} + \boldsymbol{F} = \boldsymbol{F}_{ext} - \boldsymbol{M}_0\ddot{\boldsymbol{d}}_0$$

where

$$\boldsymbol{F} = \int_\Omega \boldsymbol{f}\, d\Omega \approx \sum_{g=1}^{m} \omega_g \boldsymbol{f}(\boldsymbol{x}_{\mathbf{z}_g}; \cdot), \qquad (\boldsymbol{f} = \boldsymbol{\Phi}^T \boldsymbol{f}^h)$$

and $\boldsymbol{F}_{ext} = \boldsymbol{R}_{ext}\, \boldsymbol{F}^h_{ext}\big|_{\mathbf{c}}$, subject to the Dirichlet boundary and initial conditions:

$$\boldsymbol{d}_0 = \boldsymbol{g}^h_{\mathbf{b}}, \qquad \boldsymbol{d}(0) = \boldsymbol{u}_0, \qquad \dot{\mathbf{d}}(0) = \boldsymbol{v}_0,$$

and to the constitutive equations

$$\mathcal{H}(\boldsymbol{\sigma}, \boldsymbol{d}, \boldsymbol{d}_0, \boldsymbol{\xi}; \boldsymbol{\mu})|_{\boldsymbol{x}_{\mathbf{z}_g}} = \mathbf{0}, \quad g = 1, 2, \ldots, m$$

Box VI. Statement of the hyper-reduced order problem (*with* approximation of internal forces).

corresponding constitutive equations. In general, these recovery operations are carried out when the simulation is completed, as part of the postprocess – to display, for instance, contour plot of stresses – and therefore do not affect the online cost of the reduced-order simulation.

However, there are certain problems in which the output of interest is a function of the whole stress field and, in addition, it has to be computed online, at the end of each time step. In such cases, an *efficient* recovery of the stress field is a must. One route for efficiently recovering or *reconstructing* the stresses at all Gauss points is via *least-squares fitting* (the so-called Gappy Data reconstruction, introduced by Everson and Sirovich in [8]). Indeed, let $\boldsymbol{\Psi} \in \mathbb{R}^{s \cdot M \times q}$ be a stress basis matrix (calculate by the SVD, for instance),[13] and let $\hat{\boldsymbol{\mathcal{S}}}_{\mathbf{z}} \in \mathbb{R}^{s \cdot m}$ denote the vector containing the stresses computed by the hyper-reduced order model at the selected integration points $\mathbf{z}$, that is:

$$\hat{\boldsymbol{\mathcal{S}}}_{\mathbf{z}} = \left[\boldsymbol{\sigma}^T(\boldsymbol{x}_{\mathbf{z}_1}) \quad \boldsymbol{\sigma}^T(\boldsymbol{x}_{\mathbf{z}_2}) \quad \cdots \quad \boldsymbol{\sigma}^T(\boldsymbol{x}_{\mathbf{z}_m})\right]^T. \tag{60}$$

Least-squares fitting yields the following approximated, global stress vector $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{s \cdot M}$

$$\boldsymbol{\mathcal{S}} \approx \boldsymbol{\Psi}\, \boldsymbol{\Psi}^+_{\mathbf{z}} \hat{\boldsymbol{\mathcal{S}}}_{\mathbf{z}} \tag{61}$$

where $\boldsymbol{\Psi}^+_{\mathbf{z}} = (\boldsymbol{\Psi}^T_{\mathbf{z}}\, \boldsymbol{\Psi}_{\mathbf{z}})^{-1}\, \boldsymbol{\Psi}^T_{\mathbf{z}}$ stands for the pseudo-inverse of $\boldsymbol{\Psi}_{\mathbf{z}} \in \mathbb{R}^{s \cdot m \times q}$ ($\boldsymbol{\Psi}_{\mathbf{z}}$ is the matrix formed by the rows of $\boldsymbol{\Psi}$ corresponding to the set of indices $\mathbf{z} \in \mathbb{N}^m \subset \{1, 2, \ldots, M\}$).[14]

Recovery or reconstruction by least-squares fitting can be also applied to strains. However, to compute a basis matrix (designated by $\boldsymbol{\Upsilon}$) for strains, one need not to store snapshots as in the case of stresses. Rather, a "compressed" snapshot matrix $\boldsymbol{E}_\varepsilon$ (featuring $n + n_0$ columns) can be obtained by multiplying the displacement basis matrices $\boldsymbol{\Phi} \in \mathbb{R}^{N \times n}$ and $\boldsymbol{\Phi}_0 \in \mathbb{R}^{N_0 \times n_0}$ by the corresponding FE strain–displacement matrices. This can be done by defining the global (sparse) displacement basis matrix

$$\underline{\boldsymbol{\Phi}} = \begin{bmatrix} \boldsymbol{\Phi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi}_0 \end{bmatrix} \tag{62}$$

and then computing the rows of $\boldsymbol{E}_\varepsilon$ associated to the FE integration point $\boldsymbol{x}_g$ ($g = 1, 2, \ldots, M$) as

$$\boldsymbol{E}_\varepsilon(\boldsymbol{x}_g) = \boldsymbol{B}^{(e)}(\boldsymbol{x}_g)\, \underline{\boldsymbol{\Phi}}^{(e)}, \tag{63}$$

$\boldsymbol{B}^{(e)}(\boldsymbol{x}_g)$ being the FE strain–displacement matrix of element e at integration point $\boldsymbol{x}_g$, and $\underline{\boldsymbol{\Phi}}^{(e)}$ the rows of $\underline{\boldsymbol{\Phi}}$ corresponding to the degrees of freedom of element e (e is the index of the finite element containing the integration

[13] Notice that this stress basis matrix is to be calculated anyway when internal forces are computed by the method described in Appendix A.

[14] For $\boldsymbol{\Psi}^+_{\mathbf{z}}$ to exist, $\boldsymbol{\Psi}_{\mathbf{z}}$ must be full rank, and this, in turn, requires the number of rows of $\boldsymbol{\Psi}_{\mathbf{z}}$ be greater than the number of columns ($m \cdot s \geq q$). In Appendix A, we show that $p \sim n \cdot q$, which, by virtue of the findings of Section 4.5.2, implies that $m \sim n \cdot q$. Thus, it follows that the necessary condition for $\boldsymbol{\Psi}^+_{\mathbf{z}}$ to exist is largely met.

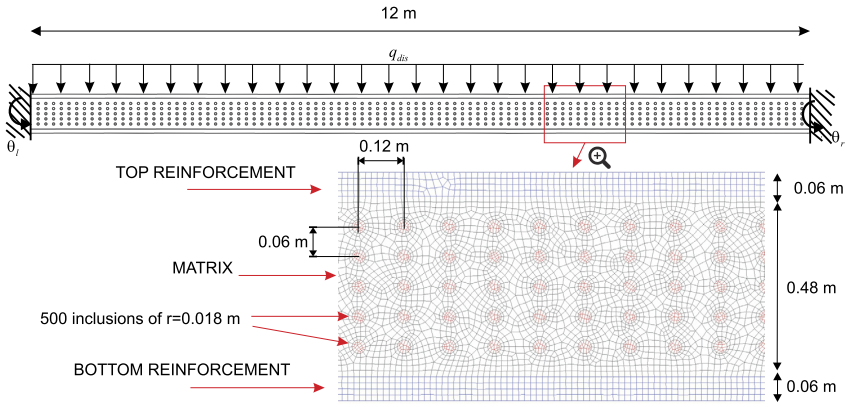


Fig. 1. Geometry of the plate (plane strain), along with the employed FE discretization and boundary conditions.

point $\boldsymbol{x}_g$). Once $\boldsymbol{E}_\varepsilon$ is at one's disposal, an orthonormal basis matrix for strains can be determined[15] by an SVD. It is worth noting that, in cases in which strains admit an additive decomposition into elastic and inelastic parts, one may reconstruct with this basis matrix, not only the total strains, but also each of these contributions. This capability is illustrated in the numerical example discussed in the ensuing section (see Fig. 10).

6. Numerical results

In this section, the efficiency of the proposed model-order reduction strategy is assessed in two representative structural examples, namely, a quasistatic (cylindrical) bending problem, on the one hand, and a dynamic, forced vibration problem, on the other hand (both in composite plates undergoing infinitesimal, elasto-plastic deformations).

6.1. Bending of a composite plate

6.1.1. Problem set-up

The composite plate is made of three distinct materials: matrix, reinforcements, and foam. The mechanical behavior of both the matrix and reinforcements materials is modeled by a rate-independent, Von Mises elastoplastic model endowed with a linear, isotropic hardening law (consult Ref. [42] for details on the implementation of this elastoplastic model). The material properties for the matrix material are: Young's modulus $E^m = 70 \cdot 10^3$ MPa, Poisson's ratio $\nu^m = 0.3$, yield stress $\sigma_y^m = 60$ MPa and hardening modulus $H^m = 5$ MPa. For the reinforcement material, on the other hand, these constants take the following values: $E^r = 200 \cdot 10^3$ MPa, $\nu^r = 0.3$, $\sigma_y^r = 110$ MPa and $H^r = 10$ MPa. Lastly, the foam inclusions are assumed to behave elastically, with $E^f = 20$ MPa and $\nu^f = 0.3$.

The finite element mesh can be also seen in Fig. 1. The number of (four-node bilinear) elements is $N_{elem} = 45\,349$, and the number of nodes $N_{node} = 46\,163$. The employed quadrature formula, on the other hand, is the standard 2×2 Gauss rule, the total number of Gauss points amounting thus to $M = 4\,N_{elem} = 181\,396$. To overcome incompressibility issues while maintaining the displacement-based formulation presented in the preceding sections, the commonly known as "B-bar" approach is adopted [42]. The constitutive differential equations are integrated in time using the classical (fully implicit) backward-Euler scheme.

6.1.2. First reduction stage

The goal of the hyper-reduced order model (HROM) we wish to develop is to predict the *bending moment on the left edge* of the plate for *any prescribed rotation on the left and right edges* (θ_l *and* θ_r*, respectively) and any transverse, uniformly distributed load* (q_{dis}). The set of *input parameters* $\boldsymbol{\mu}$ in the problem, thus, can be symbolically

[15] Normally, $\boldsymbol{E}_\varepsilon$ is full rank and can be thereby directly taken as a basis matrix, i.e., $\boldsymbol{\Upsilon} = \boldsymbol{E}_\varepsilon$.

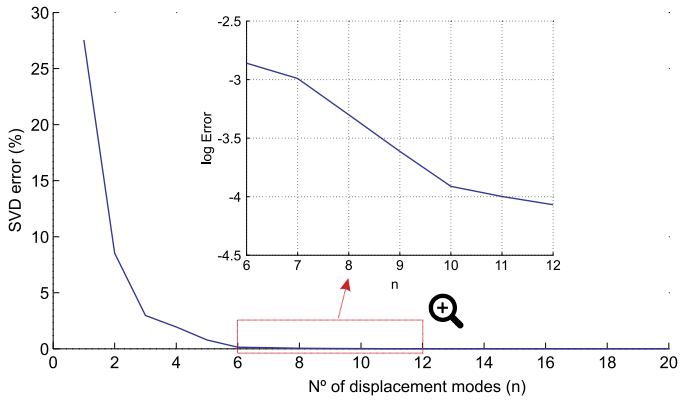


Fig. 2. SVD truncation error versus number of basis vectors employed in the approximation (n). The portion between 6 and 12 modes is shown in magnified form in logarithmic scale.

represented by $\boldsymbol{\mu} = \{\theta_l(t), \theta_r(t), q_{dis}(t)\}$. Notice that the time variable t appears implicitly also in the set of input parameters—due to plastic yielding, the response of the plate at a particular time depends on the *history* of these three parameters.

As explained in the preceding sections, the construction of the desired HROM involves two sequential dimensionality reduction stages. The first reduction stage consists in the creation of the reduced-order model with no approximation of internal forces (henceforth labeled ROM). To arrive at the ROM, we follow the steps outlined in Box II. The first step is to run finite element analysis for representative values of such input parameters (the *training* inputs). We have chosen three sets of such training inputs: rotation of increasing magnitude (linear with time) on the left edge while the right edge remains fixed ($\boldsymbol{\mu}^{tr1} = \{\theta_l^0 t/T, 0, 0\}$), rotation of increasing magnitude on the right edge while the left edge remains fixed ($\boldsymbol{\mu}^{tr2} = \{0, \theta_r^0 t/T, 0\}$), and transverse load of increasing magnitude while the right and the left edges remain fixed ($\boldsymbol{\mu}^{tr3} = \{0, 0, q_{dis}^0 t/T\}$), where $t \in [0, T]$, with $T = 10$ s. The values of the constants appearing in the preceding expression are set to $\theta_l^0 = -\theta_r^0 = 0.025$ rad, and $q_{dis}^0 = -0.85$ MN/m^2—these values ensure that the plate is loaded well into the inelastic range and, consequence, plastic hinges develop. The time domain for each input history is discretized into 200 equally spaced steps, resulting in a total number of $P = 3 \cdot 200 = 600$ snapshots.

Three matrices are to be stored in memory and processed by dimensionality reduction in the first reduction stage (Step 2 of Box II): the matrix of prescribed displacements $\boldsymbol{X}_{d0} \in \mathbb{R}^{N_0 \times P}$, the matrix of external forces $\boldsymbol{X}_{fext} \in \mathbb{R}^{N \times P}$, and the matrix of unrestricted displacements $\boldsymbol{X}_d \in \mathbb{R}^{N \times P}$ (the number of restricted and unrestricted DOFs is $N_0 = 156$ and $N = 2N_{node} - N_0 = 92\,170$, respectively). Spatial variation of boundary conditions can be (exactly) described by just two parameters (θ_l and θ_r), and therefore, $n_0 = \text{rank}(\boldsymbol{X}_{d0}) = 2$; by the same token, for external forces, $n_f = \text{rank}(\boldsymbol{X}_{fext}) = 1$. Using the SVD, the corresponding basis matrices $\boldsymbol{\Xi} \in \mathbb{R}^{N_0 \times n_0}$ and $\boldsymbol{\Theta} \in \mathbb{R}^{N \times n_f}$ (see Section 3.2) can be easily obtained; the interpolation indices $\mathbf{b}$ and $\mathbf{c}$, on the other hand, can be either automatically determined by means of the Empirical Interpolation Method [3,4], or by manually selecting two horizontal degrees of freedom (DOFs) (one on each edge) for interpolation of prescribed displacements, and one vertical DOF on the top boundary for interpolation of external force.

To obtain the basis matrix $\boldsymbol{\Phi}$ for the unrestricted displacements, we follow the SVD-like elastic/inelastic factorization proposed by the authors in Ref. [17]—and sketched, for completeness, in Appendix C. Since the problem is quasi-static in the small strains regime, the number of elastic modes[16] is equal to the spatial dimensionality of the set of input parameters, i.e., $n^{el} = n_0 + n_f = 3$. Therefore, we only have to elucidate how many inelastic modes $n^{in} = n - n^{el}$ can be deemed as dominant or "essential". To this end, we plot in Fig. 2 the (dimensionless) SVD truncation error estimates defined as $e^* := \|\boldsymbol{X}_d - \boldsymbol{X}_d^*(n)\|/\|\boldsymbol{X}_d\|$, where $\boldsymbol{X}_d^*(n)$ denotes the SVD approximation of rank n.

It can be appreciated in Fig. 2 that the first three modes – the elastic modes – contains more than 95% of the information, while the remaining 5% only corresponds to pure inelastic modes. With $n = 10$ (7 inelastic modes), the

[16] The number of displacement modes necessary to exactly reproduce the FE results in the elastic range.

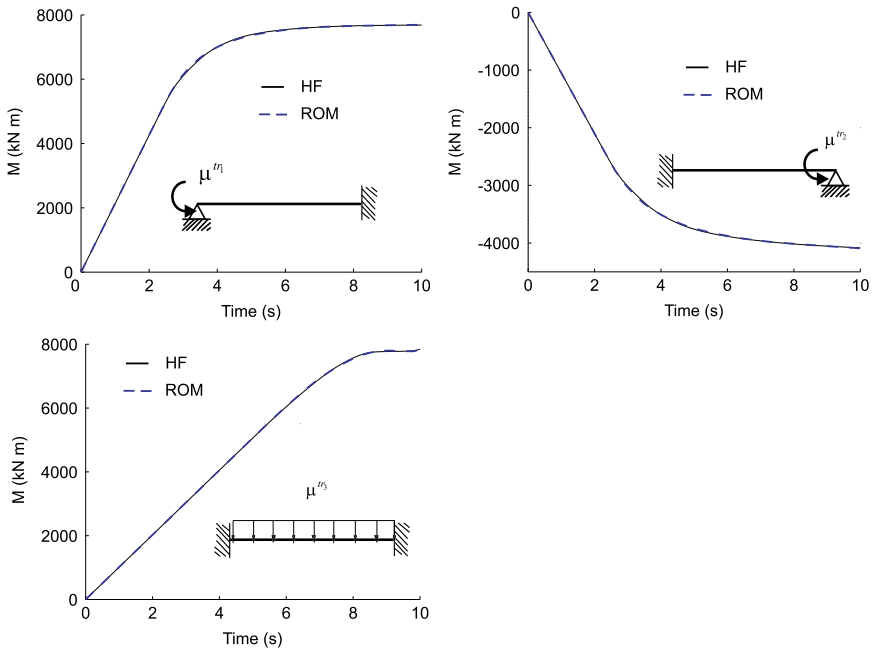
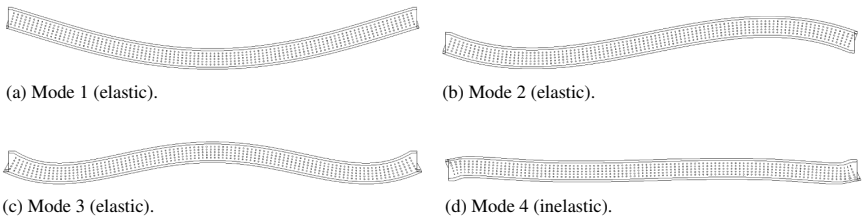


Fig. 3. Bending moment on the left edge versus time for the three training trajectories. Results computed with the FE model and the ROM using $n = 10$ and the full set of FE integration points. The relative L2 errors for the three cases are 0.20%, 0,19% and 0.16%.



(a) Mode 1 (elastic).

(b) Mode 2 (elastic).

(c) Mode 3 (elastic).

(d) Mode 4 (inelastic).

Fig. 4. Deformed shapes corresponding to the first 4 dominant displacement modes ($\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \boldsymbol{\Phi}_3, \boldsymbol{\Phi}_4\}$).

error level is around 0.01%. In Fig. 3, we plot the output of interest (evolution of bending moment on the left edge) for the three trajectories using the FE model and the ROM constructed with such a number of displacement modes ($n = 10$); as expected, differences between FEM response and the one predicted by the ROM are negligible (below 0.2%). Finally, by way of illustration, Fig. 4 displays the deformed shape corresponding to the first 4 displacement modes ($\boldsymbol{\Phi}_i$, $i = 1, 2, 3, 4$).

6.1.3. Second reduction stage

In the first reduction stage, discussed in the foregoing, the number of displacement unknowns have been reduced from $N = 92\,170$ to $n = 10$ with almost no loss in accuracy. Yet, this reduction in the number of DOFs only provides modest speedup of around 2. This fact highlights that the actual bottleneck for fast online computation is not the solution of the discrete balance equations, but rather the determination of the stresses, internal forces and stiffness matrices at all the integration points of the underlying finite element mesh (in this case, $M = 181\,396$). The second reduction stage, summarized in Box V, is aimed at choosing among these FE points a reduced set of optimized cubature points (as well as their associated positive weights).

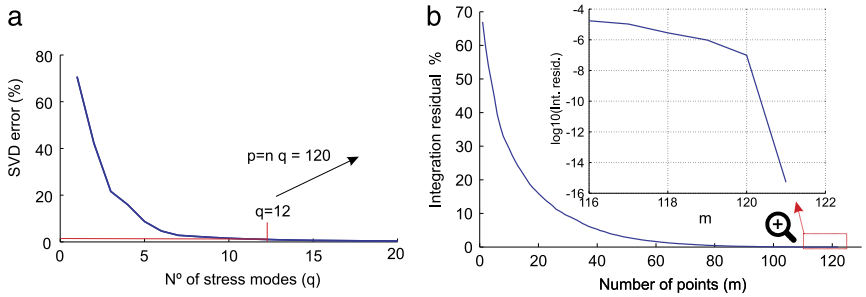


Fig. 5. (a) SVD truncation error for stresses versus number of modes q. (b) Dimensionless residual $\|\boldsymbol{r}\|/V$ versus number of integration points for the case $q = 12$ (hence $p = q \cdot n = 12 \cdot 10 = 120$). The portion between 116 and 122 points is shown in magnified form in logarithmic scale.

We begin by solving the reduced-order equations (Step 1) for the same input parameters employed in the first reduction stage (that is, $\{\boldsymbol{\mu}^{tr1}(t_i), \boldsymbol{\mu}^{tr2}(t_i), \boldsymbol{\mu}^{tr3}(t_i)\},\ i = 1, 2, \ldots, 200$). For computing and processing the reduced internal forces in Step 2, we adopt the stress-based procedure described in Appendix A: the stresses at all FE Gauss points and all time steps for the three trajectories are stored column-wise in a matrix $\boldsymbol{X}_\sigma \in \mathbb{R}^{sM \times P}$ (here $s = 4$); then, the elastic/inelastic factorization (the same employed for displacements) is performed on $\boldsymbol{X}_\sigma$, taking as number of elastic stress modes[17] $q^{el} = 3$. The corresponding SVD error graph versus number of stress modes q – the analogous of Fig. 2 – is displayed in Fig. 5(a). It can be gleaned from this Figure that $q = 12$ gives reasonably low stress approximation error (below 1%). Accordingly, we construct the "compressed" internal force matrix $\boldsymbol{X}_\mathcal{F}$ (see formula (A.5) in Appendix A) using $q = 12$ stress modes—hence, $\boldsymbol{X}_\mathcal{F}$ features $p = q \cdot n = 12 \cdot 10 = 120$ columns.

Next, we determine the zero-average internal force matrix $\hat{\boldsymbol{X}}_\mathcal{F}$ (Step 3 in Box V), and use the SVD to determine a set of orthogonal basis vectors $\boldsymbol{\Lambda}$ for the reduced internal forces (Step 4). With $\boldsymbol{\Lambda}$ at our disposal, we construct the matrices $\boldsymbol{J}$ and $\boldsymbol{b}$ appearing in the objective function of the cubature optimization problem (Step 5), and finally solve this problem by means of the greedy algorithm of Box IV (Step 6).

6.1.4. Empirical cubature

To examine the convergence of this greedy algorithm, we plot in Fig. 5(b) the dimensionless residual $\|\boldsymbol{r}\|/\|\boldsymbol{b}\| = \|\boldsymbol{r}\|/V$ (the one employed as termination criterion in Box IV) versus the number of points m (for the case $p = n \cdot q = 10 \cdot 12 = 120$). It can be readily seen that the residual decreases monotonically as the number of integration points increases, and at the threshold $m = p + 1 = 121$, it drops sharply to a negligible value ($\sim 10^{-15}$), indicating that, as theoretically anticipated in Section 4.5.2, the algorithm has converged to the absolute minimum.

Concerning the computation time, the partitioned SVD of the stress snapshot matrix took approximately 70 s, while the selection of $m = 121$ points among $M = 181\,396$ FE points was carried out in approximately 45 s (both operations in an serial *Matlab* program, running at 2.9 GHz with 8 GB of RAM and 4 *Intel Core-i7* processors, in Linux). This is as a relatively low computation time when compared with previous cubature optimization methods. When using the original[18] greedy algorithm by An et al. [19], the CPU time for selecting $m = 121$ points rises up to 1014 s (approximately 8 times slower). Difference in performance becomes more pronounced as the number of selected points increases. For $m = 300$ points, for instance, our algorithm needed 233 s ($\approx$4 min), while the algorithm in [19] employed 4053 s (1 h and 10 min). As pointed out in Section 4.5.1, the reasons why our method appears to outperform the original scheme of [19] are two: firstly, memory requirements are drastically reduced; and, secondly, and most importantly, our approach avoids in almost all iterations the expensive nonnegative least-squares problem. Specifically, in the selection of $m = 300$ points, our algorithm only needed to solve the nonnegative least squares problem in 5 iterations (of a total of 307 iterations).

[17] In order not to process the whole matrix $\boldsymbol{X}_\sigma$, we employ the partitioned SVD strategy, described in Appendix B, for both the elastic and inelastic matrices. In particular, we employ a partition of 10 block matrices with 60 snapshots each.

[18] We implemented the greedy algorithm proposed by [19] in its *raw* form, i.e., without the heuristics aimed at accelerating the algorithm (namely, subset strategy for choosing new candidates, and subset training).

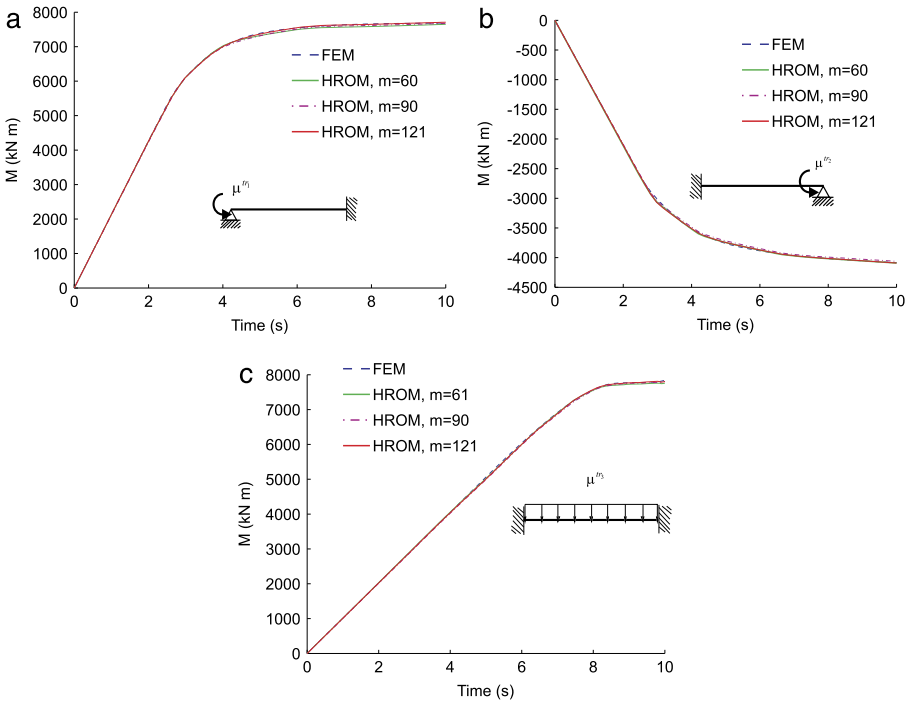


Fig. 6. Bending moment on the left edge versus time for the three training trajectories. Results computed with the FE model ($M = 181\,396$ integration points) and the HROM using $m = 60$, 90 and 121 integration points.

Table 1
Relative L2 error between the predictions of the HROM for the three training trajectories – using $m = 60$, 90 and 121 integration points – and the reference FE model (that employs $M = 181\,396$ integration points).

	Traj. 1	Traj. 2	Traj. 3	Max.
$m = 60$	0.48%	0.30%	0.78%	0.78%
$m = 90$	0.51%	0.68%	0.50%	0.68%
$m = 121$	0.29%	0.31%	0.17%	0.31%

6.1.5. HROM results

Next we study the extent to which the integration error affects the quality of the response predicted by the HROM (in terms of the output of interest). To this end, we analyze in Fig. 6 the evolution of the bending moment on the left edge versus time, for the three training trajectories, using both the finite FE model and the HROM with $q = 12$ stress modes and varying number of integration points. Specifically, we set $m = 60$, 90 and 121. The relative L2 errors for each case are displayed in Table 1. Inspection of Fig. 6 and Table 1 shows that, remarkably, deviations between FE response and the HROM graphs are practically imperceptible for the three cases. These results also suggest that there is no need to exactly integrate all the internal force modes: gains in accuracy achieved by passing from $m = 60$ to $m = p + 1 = 121$ points are practically negligible (from 0.78% to 0.31%).

Fig. 7. Location of the $m = 60$ integration points chosen by the greedy algorithm (for $p = 12$ stress modes).

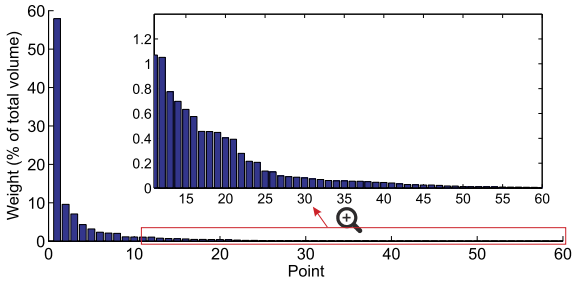


Fig. 8. Weight (in % of the total volume V, and sort in descending order) associated to each integration point (for $m = 60$).

Fig. 7 depicts the position of the elements[19] containing the $m = 60$ points selected by the greedy algorithm. Observe that around 50 points are equally distributed along regions close to the right and left edges, while the remaining points are scattered over the middle portion of the beam. Likewise, the weights (divided by the total volume) associated to these $m = 60$ points are plotted in Fig. 8. Notice that the distribution of the total volume among the chosen points is far from being uniform: the sum of the 10 highest weights amounts to 90% of the total volume, while the remaining weights only contribute with a 9.99% (hence, the error in integrating the volume with these 60 points is only 0.01%).

We pointed out in Section 5.1 that, although the HROM only tracks the evolution of stresses and (plastic and elastic) strains at the reduced set of integration points, one can "reconstruct", for post-processing purposes, such fields by using least-square fitting (see Eq. (61)). To illustrate this capability, we plot in Figs. 9 and 10 the contour plots of effective stresses and effective plastic strains, respectively, at the end of the first training configuration computed by the FE model and the HROM model (using $q = 12$ stress modes, $n + n_0 = 12$ strain modes and $m = 60$ points). The qualitative resemblance between the HROM and FE patterns is startling in both cases—despite the fact that the number of integration points has been reduced by a factor over 3000. Obviously, upon closer inspection, some "inconsistencies" become apparent. For instance, in Fig. 10(b), it can be appreciated that the HROM predicts that plastic yielding occurs at some of the (elastic) foam inclusions. This is because least-squares fitting is a purely data driven approach and, consequently, the reconstructed variables are not consistent, in general, with the constitutive behavior assumed for each material.

6.1.6. Testing trajectory

We examine now the error incurred by the HROM in predicting the response of the plate for input parameters different from those employed in the "training" process. The plate is subjected to a loading/unloading cycle of prescribed rotation on the right edge in tandem with transverse load, while the left edge remains fixed, i.e., $\boldsymbol{\mu} = \{0, \theta_r^{max} g(t), q_{dis}^{max} g(t)\}$. The constants are set to $\theta_r^{max} = -0.01$ rad and $q_{dis}^{max} = -0.4$ M N/m^2; on the other hand, $g = g(t) : [0, T] \to \mathbb{R}$ is the piecewise constant function shown in Fig. 11(a). The analysis required 300 equally spaced time steps. The plots of bending moment on the left edge versus time for the FE model and the HROM ($m = 60$

[19] It should be stressed here that the selection of a Gauss point within a given finite element does not imply that the rest of Gauss points of the element are included in the reduced set of integration points. Hence, as distinct from the "mesh-sampling" method advocated by Farhart and co-workers [20,21], the HROM proposed here completely ignores the finite element origin of the integration points.

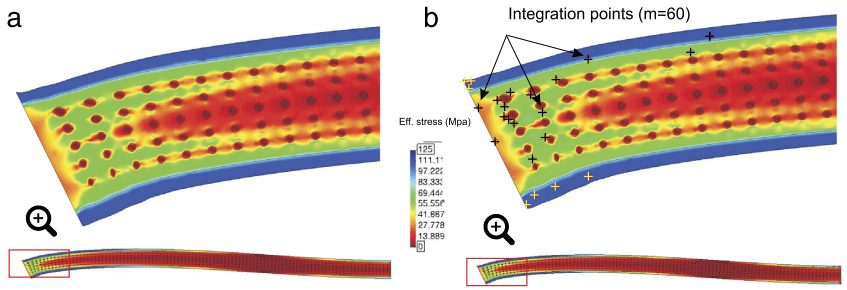


Fig. 9. Contour plot of effective stresses at the end of the first training configuration (deformed shape amplified by a factor of 10). (a) Finite element model ($M = 181\,396$ integrations points). (b) Hyper-reduced order model ("reconstruction" using $q = 12$ stress modes and $m = 60$ integration points).

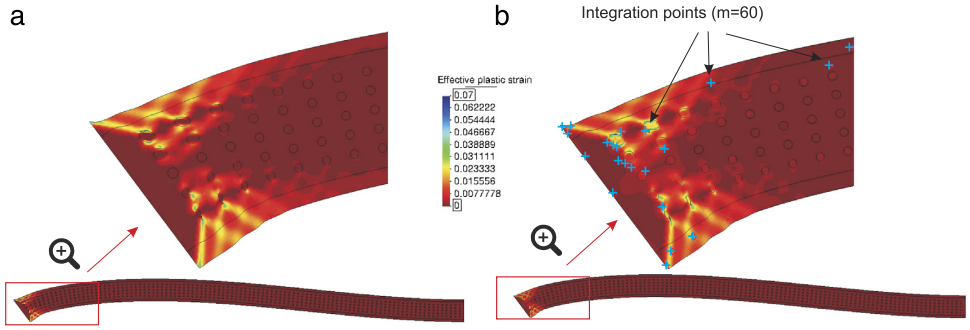


Fig. 10. Contour plot of effective plastic strains at the end of the first training configuration (deformed shape amplified by a factor of 10). (a) Finite element model ($M = 181\,396$ integrations points). (b) Hyperreduced order model ("reconstruction" using $n + n_0 = 10 + 2 = 12$ strain modes and $m = 60$ integration points).

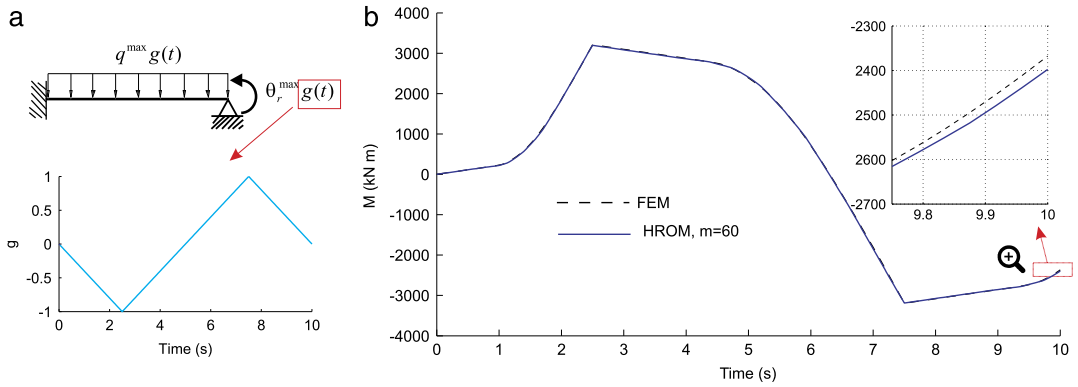


Fig. 11. (a) Input parameters employed to test the HROM. (b) Bending moment on the left edge versus time for these input parameters. Results computed with the FE model and the HROM using $m = 60$ integration points.

integration points) are displayed in Fig. 11(b). Remarkably, the two curves are practically indistinguishable at the scale of plot; the maximum error is observed at the end of the process (residual bending moment, see enlarged region), and it is below 1.5%.

Table 2
Comparison of the dimensions and computation time of the finite element problem and the hyper-reduced order problem, along with the corresponding "compression" ratios.

	FEM	HROM	"Compression" ratio
Number of unrestricted DOFs	$N = 92170$	$n = 10$	9217
Number of restricted DOFs	$N_0 = 156$	$n_0 = 2$	78
Number of parameters defining external forces	$N_f = 776$	$n_f = 1$	776
Number of integration points	$M = 181396$	$m = 60$	3023.2
Computation time	6180 s	2.5 s	2472

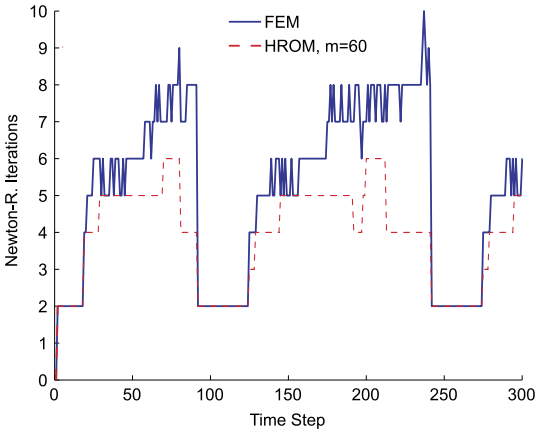


Fig. 12. Number of iterations required for convergence of the accompanying Newton–Raphson algorithm for both the FE model and the HROM (with $m = 60$ points).

6.1.7. Compression ratios and speedup

To summarize, we present in Table 2 the dimensions characterizing both the high-fidelity FE model and the employed HROM, along with the corresponding "compression" or dimensionality reduction factors. Observe that both the number of displacement unknowns and the number of integration points have been reduced by three orders of magnitude. These startling compression ratios are reflected in equally astonishing speedup factors (≈2400): the 300 time steps of this example were computed in 2.5 s by the HROM, while the FE model employed 1 h and 43 min (both in a vectorized[20] *Matlab* program, running at 2.9 GHz with 8 GB of RAM and 4 *Intel Core-i7* processors, in Linux).

Lastly, to assess the robustness of the HROM, we show in Fig. 12 the number of iterations required for convergence of the accompanying Newton–Raphson algorithm in the case of the FE model and the HROM. Observe that, at all time steps, the number of iterations employed by the HROM is less or equal than in finite element model (for the same convergence tolerance).

6.2. Forced vibration of a composite plate

6.2.1. Problem set-up

We now turn our attention to the development of an HROM for the free vibration of the composite plate shown in Fig. 13. The procedure to arrive at such a HROM is essentially the same as the one explained in the previous example; thereby, in the interest of brevity, some details will be omitted (such as the SVD error analyses).

The plate is made of two materials: matrix and foam. The mechanical behavior of the matrix is modeled by a rate-independent, Von Mises elastoplastic model endowed with a linear, isotropic hardening law ($E^m = 100{\cdot}10^3$ MPa,

[20] It should be remarked that the speedup factors are even more spectacular when standard, non-vectorized FE Matlab codes are used; in our case, the speedup for this example rises above 18 000.

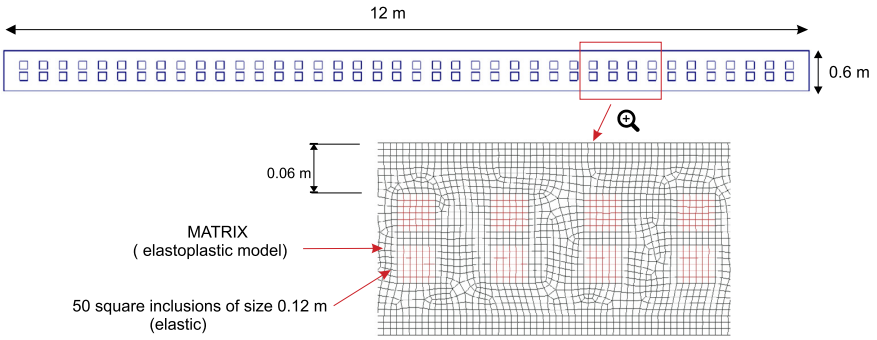


Fig. 13. Geometry of the plate (plane strain), along with the employed FE discretization.

$\nu^m = 0.3$, $\sigma_y^m = 60$ MPa and $H^m = 5$ MPa), while the foam inclusions are assumed to behave elastically ($E^f = 20$ MPa and $\nu^f = 0.3$). The density of the matrix and the foam, on the other hand, are $\rho^m = 7.75 \cdot 10^3$ kg/m^3 and $\rho^f = 0.775$ kg/m^3, respectively.

The finite element mesh can be also seen in Fig. 13. The number of (four-node bilinear) elements is $N_{elem} = 17\,829$ (hence, $M = 4 \cdot N_{elem} = 71\,316$), and the number of nodes $N_{node} = 18\,460$. The constitutive differential equations are integrated in time using the classical (fully implicit) backward-Euler scheme, while the momentum equation is integrated using a Newmark β-method [43], with parameters $\beta = 1/4$ and $\gamma = 1/2$.

6.2.2. Training process

The goal of the reduced-order model in this problem is to predict the vibration behavior of the plate when the left edge is subjected to a vertical, oscillatory displacement of the form

$$v_l(t) = \sum_{i=1}^{Q} A_i \sin 2\pi \frac{t}{T_i^v} \tag{64}$$

while the other edge remains free (hence, $\boldsymbol{\mu} = \{v_l(t)\}$). In particular, we are interested in capturing the vibration behavior when the periods of the prescribed displacement T_i^v ($i = 1, 2, \ldots, Q$) range between the first and fourth largest *natural* period of the plate: $\bar{T}_1 \leq T_i^v \leq \bar{T}_4$, including the possible development of plastic yielding when $T_i^v \approx \bar{T}_1$ and $T_i^v \approx \bar{T}_2$, that is, when resonance occurs at the two largest natural periods of the structure (an eigenvalue analysis shows that the natural periods of this structure are: $\bar{T}_1 = 0.760$ s, $\bar{T}_2 = 0.125$ s, $\bar{T}_3 = 0.047$ s and $\bar{T}_4 = 0.029$ s).

The training process is made according to the above specifications: we set $Q = 2$, $A_1 = A_2 = 2.1 \cdot 10^{-3}$ m, and $T_1^v = 0.98\,\bar{T}_1 = 0.745$ s and $T_2^v = 1.08\,\bar{T}_2 = 0.135$ s. The resulting graph of prescribed displacement versus time is displayed in Fig. 14 (blue line).

6.2.3. First reduction stage

We run a finite element analysis in which the time domain is discretized into 6000 equally spaced steps and store in memory the matrix of prescribed displacements $\boldsymbol{X}_{d0} \in \mathbb{R}^{N_0 \times P}$ and the matrix of unrestricted displacements $\boldsymbol{X}_d \in \mathbb{R}^{N \times P}$ (the number of restricted and unrestricted DOFs is $N_0 = 77$ and $N = 2N_{node} - N_0 = 36843$, respectively). Notice that spatial variation of Dirichlet boundary conditions can be described by just one parameters (v_l), and therefore, $n_0 = \text{rank}(\boldsymbol{X}_{d0}) = 1$; by the same token, for external forces, $n_f = 0$.

To obtain the basis matrix $\boldsymbol{\Phi}$ for unrestricted displacements, we follow the approach described in Appendix C.1; in this approach, the first n^{vib} columns of $\boldsymbol{\Phi}$ are the dominant natural vibration modes of the structure, while the remaining modes are determined by a weighted SVD of the orthogonal complement of the snapshot matrix $\boldsymbol{X}_d$. Here, we set $n^{vib} = 4$ (to meet of the requirement outlined earlier), and $n^{\perp} = 6$. These $n = 10$ modes proved to be sufficient for predicting the displacement response of the structure with an error level below 1%. This can be appreciated in

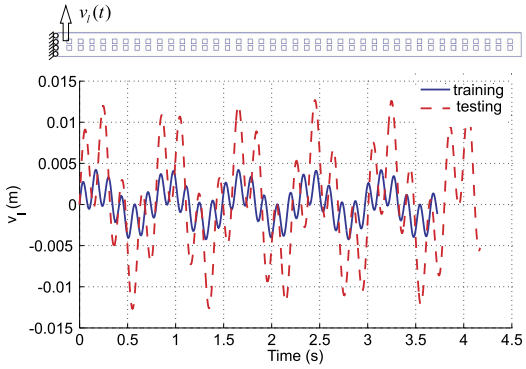


Fig. 14. Prescribed displacements versus time employed for training and testing the HROM. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

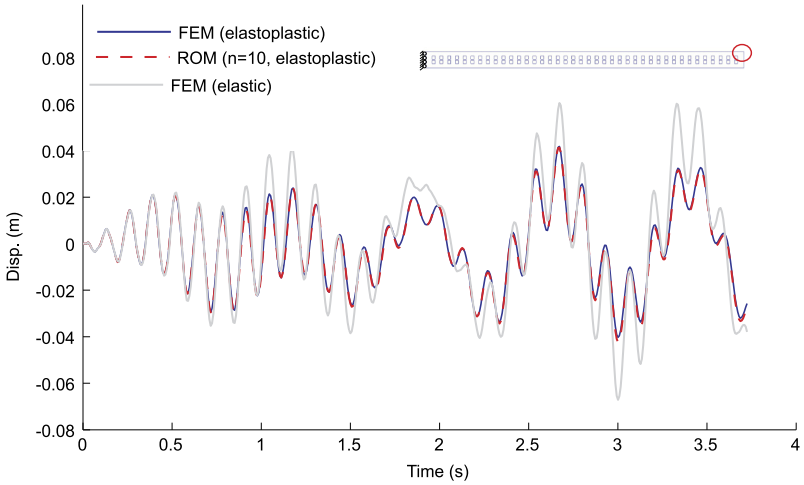


Fig. 15. Displacement of the right, top corner node versus time computed with the FE model and the ROM (using n^{vib} vibration modes and $n^{\perp} = 6$ SVD modes).

Fig. 15, where we show the vertical displacement of the top, rightmost corner node versus time computed by the FE model and the ROM with these $n = 10$ modes. To illustrate the above mentioned "attenuation" effect due to plastic yielding, we show also in this Figure the displacement computed by the FE model had the matrix material been entirely elastic.

6.2.4. *Second reduction stage*

For the second reduction stage, we adopt (as in the quasi-static example) the stress-based procedure outlined in Appendix A: the stresses at all FE Gauss points and all time steps are stored column-wise in a snapshot matrix, and then, the SVD-like elastic/inelastic factorization is performed on such a matrix. An error analysis revealed that, by using 4 elastic modes and 6 inelastic modes, the truncation error is below 1%. Accordingly, we construct the required internal force matrix $\hat{\boldsymbol{X}}_{\mathcal{F}}$ using these modes (hence, $\hat{\boldsymbol{X}}_{\mathcal{F}}$ has $p = q \cdot n = 10 \cdot 10 = 100$ columns). With $\hat{\boldsymbol{X}}_{\mathcal{F}}$ at our disposal, we construct the matrices $\boldsymbol{J}$ and $\boldsymbol{b}$ appearing in the objective function of the cubature optimization problem, and finally solve such a problem by means of the greedy algorithm of Box IV.

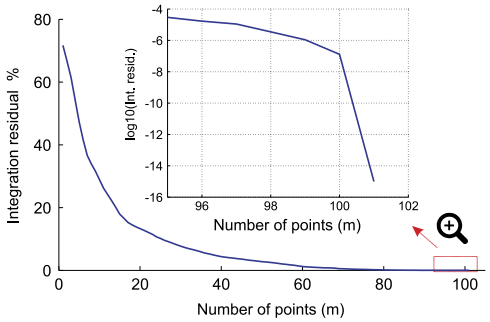


Fig. 16. (a) Dimensionless residual $\|\boldsymbol{r}\|/V$ versus number of integration points for the case $q = 10$ (hence $p = q \cdot n = 10 \cdot 10 = 100$ internal force modes). The portion between 95 and 102 points is shown in magnified form in logarithmic scale.

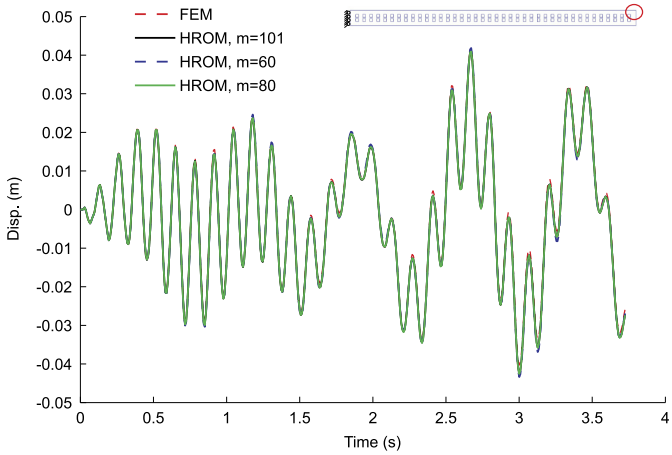


Fig. 17. Displacement of the right, top corner node versus time computed with the FE model and the HROM (using $q = 10$ stress modes and $m = 60, 80$ and 101 integration points).

6.2.5. *Empirical cubature*

Fig. 16(b) contains the graph of the dimensionless residual $\|\boldsymbol{r}\|/\|\boldsymbol{b}\| = \|\boldsymbol{r}\|/V$ versus the number of points m (for the case $p = n \cdot q = 100$). As expected, the algorithm converges to the absolute minimum (zero integration error) at $m = p + 1 = 101$ points. To assess a posteriori integration errors, we compare in Fig. 17 the plots of displacement of the top, right node versus time for the FE model and the HROM with varying number of integration points ($m = 60, 80$ and 101). Inspection of these plots shows that the level of accuracy at $m = 60$ (residual error around 1%) and $m = 101$ (residual error around 10^{-15}) are quite similar. This observation reinforces the conclusion made in the previous example that there is no need to exactly integrate all the internal force modes appearing in the objective function. On the other hand, Fig. 18(a) shows the location of the elements containing the selected points (for $m = 60$). Observe that around 60% of such points are located in the region near the left edge, wherein plastic yielding is more likely to occur due to resonance-induced bending. Likewise, the value of the associated weights (as a percentage of the total volume) is displayed in Fig. 18(b). The sum of these $m = 60$ weights furnishes a volume only 10^{-5}% below the total volume.

6.2.6. *Testing trajectory*

It only remains to assess the capability of the HROM to predict the vibrational behavior of the plate under conditions different from those used in the "calibration" (training) process. To this end, we set in Eq. (64) $Q = 2$, $A_1 = A_2 = 6.33 \cdot 10^{-3}$ m, and in order to induce resonance, $T_1^v = 1.02\,\bar{T}_1 = 0.758$ s (2% above the first natural

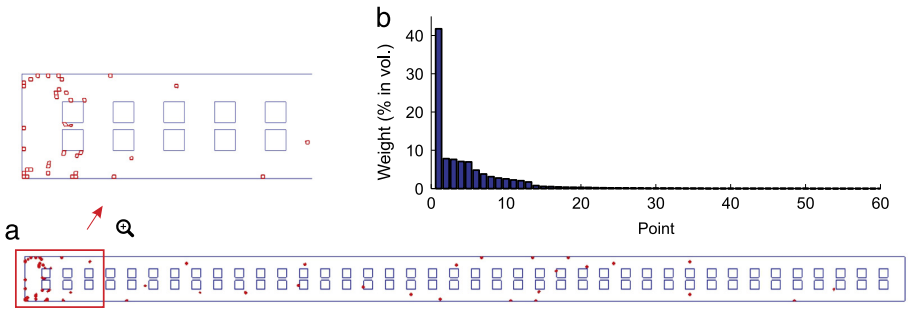


Fig. 18. (a) Location of the $m = 60$ integration points chosen by the greedy algorithm (for $p = 10$ stress modes). (b) Weights (sort in descending order) in % of total volume.

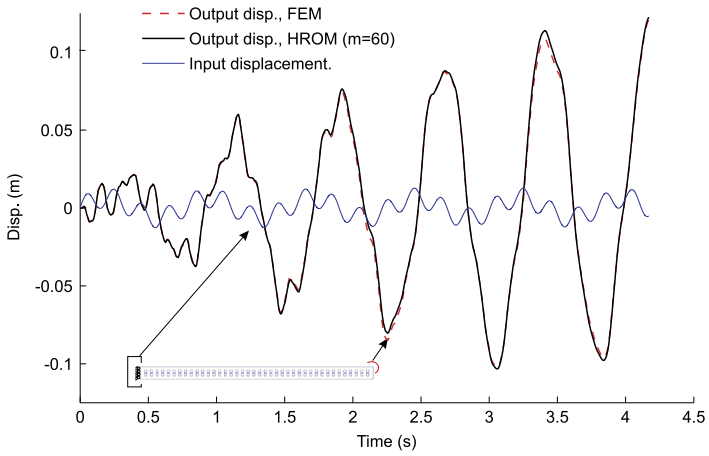


Fig. 19. Displacement of the right, top corner node versus time computed with the FE model and the HROM with $m = 60$ points. The input displacement (prescribed on the left edge) is also shown.

period). The other constant, T_2^v, is set 60% above the second natural period ($T_2^v = 1.6\,\bar{T}_2 = 0.20$ s). The resulting prescribed displacement versus time graph is shown in Fig. 14 (red, dotted curve).

In Fig. 19, we show the vertical displacement of the top, rightmost corner node versus time computed by the FE model and the HROM (with $n = 10$ displacement modes and $m = 60$ integration points). To illustrate the resonance phenomenon, this graph is accompanied by the plot of the input displacement versus time (displacement prescribed at the left edge). Discrepancies between the predictions of the high-fidelity, FE model ($M = 71\,316$ integration points) and the hyperreduced-order model ($m = 60$ integration points) are barely perceptible until time $t \approx 1.75$ s; thereafter, small drifts (below 5%) are detected at the peaks of the curve. Remarkably, at the end of the simulation, the HROM error is below 1.5%. The remarkable accuracy of the HROM can be also appreciated in term of stresses, in Fig. 20, where we display the contour plots of effective stresses (at the end of the simulation) obtained by the FEM and the HROM (using $q = 10$ stress modes and $m = 60$ points).

7. Concluding remarks

- *Robustness*. One of the most attractive features of the proposed hyper-reduced order model – and in general, of all cubature-based ROMs – is that it preserves the spectral properties of the Jacobian matrix of the finite element motion equations. This has been corroborated by the examples shown in the preceding section. The materials of the studied composite beams obey small strains, elastoplastic constitutive equations endowed with strain hardening,

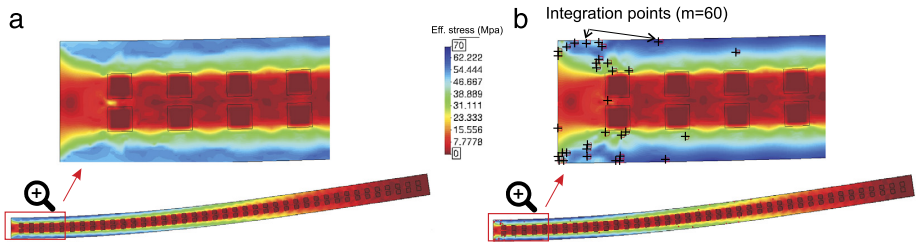


Fig. 20. Contour plot of effective stresses at the end of the simulation. (a) Finite element model ($M = 71\,316$ integrations points). (b) Hyperreduced order model ("reconstruction" using $q = 10$ stress modes and $m = 60$ integration points). Deformed shaped amplified 10 times in both cases.

and therefore, the finite element and the reduced order stiffness matrices at all Gauss points are symmetric and positive definite. Since the stiffness matrices in the HROM are positive combination of these matrices, it follows that the HROM matrices are also symmetric and positive definite. As the graphs shown in Fig. 12 demonstrate, this translates into equally fast or even faster convergence rates of the accompanying solution algorithm (Newton–Raphson).

- *Number of integration points*. We have theoretically demonstrated and numerically confirmed that the number of points needed to *exactly* integrate p internal force modes (plus the volume) is $m = p + 1$. It turn, it is shown in Appendix A that the number of internal force modes is equal to the product of numbers of displacement and stress modes: $p = n\,q$. Thus, it may be concluded that

$$n \leq m \leq n\,q + 1 \tag{65}$$

(the lower bound n is dictated by well-posedness considerations, see inequality (32)). Furthermore, in most problems (specially those in which the input parameters are variations of boundary conditions), the number of stress modes is of the same order than the number of displacement modes ($q = \mathcal{O}(n)$); thus, we can say that:

$$n \leq m \leq \mathcal{O}(n^2). \tag{66}$$

For instance, in the quasistatic example of Section 6.1, $n = 10$ and $q = 12$, while in the vibration problem of Section 6.2, $n = q = 10$; this amounts to $m = 121$ and $m = 101$ points, respectively, for exactly integrating the corresponding internal force modes.

Nevertheless, the a posteriori accuracy assessments for varying number of integration points carried out for the two examples (Figs. 6 and 17) suggest that *exactly* integrating the chosen p modes amount to "overkill": in both cases cases, using around 50% and 60% of the threshold $p + 1$ was sufficient to deliver reasonably accurate predictions of the pertinent outputs of interest.

- *Computer implementation*. Three distinct simulation codes are involved in the overall model reduction process, namely, (1) the finite element code, employed for generating displacement snapshots (see Box I); (2) the intermediate reduced-order code, employed for generating the reduced internal force snapshots (see Box III); (3) and the final hyperreduced-order code, Box VI, which is the one used in the "online" computations. It is noteworthy that the three underlying problems are nothing but particular cases of Galerkin approximation method, the basic differences being the basis functions employed for the approximation and the integration rule for the evaluation of internal forces. Indeed, the finite element model uses classical shape functions with local support, while both the ROM and the HROM seek the solution in spaces spanned by Ritz (globally supported) basis functions. Likewise, internal forces (and therefore stiffness matrices) are evaluated by elementwise Gauss rules in the FEM and the ROM, whereas, in the HROM, this operation is carried by a tailored cubature scheme—the proposed Empirical Cubature Method.
- *Empirical Cubature Method*. A distinguishing feature of the proposed Empirical Cubature Method with respect to similar cubature schemes is that the weights at almost all iterations of the greedy selection algorithm are calculated with a standard, unconstrained least-squares—in fact, the nonnegative least squares problem of step 5, see Box IV, is included to filter out small negative weights caused by roundoff errors. We have not given a formal proof of this salient feature – it was discovered by "serendipity" –, but numerical experiments seem to suggest that it may be attributed to the fact that basis vectors in the objection function are mutually orthogonal, and furthermore, their volumetric averages are zero. Research in this aspect of the method is currently in progress and will be reported in a forthcoming publication.

Acknowledgments

The research leading to these results has received funding from, on the one hand, the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement no. 320815, Advanced Grant Project COMP-DES-MAT, and, on the other hand, the Spanish Ministry of Economy and Competitiveness through the National Research Plan 2014: MAT2014-60919-R.

Appendix A. Basis matrix for internal forces

A key step in the development of the proposed reduced-order integration scheme (see Section 4.4) is the determination of an orthogonal basis matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{M\times p}$ for the column space of matrix $\hat{\boldsymbol{X}}_{\mathcal{F}}$. The direct route to determine this basis matrix is to store the reduced internal forces at all FE integration points and all training configuration in the snapshot matrix $\boldsymbol{X}_{\mathcal{F}} \in \mathbb{R}^{nM\times P}$, then use Eq. (51) to transform $\boldsymbol{X}_{\mathcal{F}}$ into $\hat{\boldsymbol{X}}_{\mathcal{F}}$, and finally apply the SVD to $\hat{\boldsymbol{X}}_{\mathcal{F}}$ ($\hat{\boldsymbol{X}}_{\mathcal{F}} \approx \boldsymbol{\Lambda}\boldsymbol{\Sigma}_{\Lambda}\boldsymbol{V}_{\Lambda}^{T}$).

Alternatively, in problems in which the geometry is input-parameter independent, one may obtain the basis matrix for the internal forces from a basis matrix for the stresses. Indeed, the reduced internal force at a given FE integration point $\boldsymbol{x}_g$ is given by

$$f_I(\boldsymbol{x}_g;\boldsymbol{\mu}) = \boldsymbol{B}_I^T(\boldsymbol{x}_g)\boldsymbol{\sigma}(\boldsymbol{x}_g,\boldsymbol{\mu}). \tag{A.1}$$

Here, $\boldsymbol{B}_I \in \mathbb{R}^{s\times 1}$ denotes the reduced strain–displacement matrix for the displacement mode $\boldsymbol{\Phi}_I$, and $\boldsymbol{\sigma} \in \mathbb{R}^s$ the Cauchy stress vector ($s = 4$ or 6 for 2D or 3D problems, respectively). In turn, the expression for $\boldsymbol{B}_I$ reads

$$\boldsymbol{B}_I(\boldsymbol{x}_g) = \boldsymbol{B}^{(e)}(\boldsymbol{x}_g)\,\boldsymbol{\Phi}_I^{(e)} \tag{A.2}$$

$\boldsymbol{B}^{(e)}(\boldsymbol{x}_g)$ being the FE strain–displacement matrix of element e at point $\boldsymbol{x}_g$, and $\boldsymbol{\Phi}_I^{(e)}$ the entries of $\boldsymbol{\Phi}_I$ corresponding to the degrees of freedom of element e (e denotes the index of the finite element containing the integration point $\boldsymbol{x}_g$). Since the geometry is input-independent, $\boldsymbol{B}_I$ is also independent[21] of $\boldsymbol{\mu}$; thus, *it is solely the stresses that depend on the input parameters*. The dimensionality reduction effort can be thereby concentrated on the matrix of stress snapshots, constructed by collecting the stress vectors for all FE integrations points and all training configurations as follows:

$$\boldsymbol{X}_{\sigma} := \begin{bmatrix}\boldsymbol{\mathcal{S}}^1 & \boldsymbol{\mathcal{S}}^2 & \cdots & \boldsymbol{\mathcal{S}}^P\end{bmatrix}, \tag{A.3}$$

where $\boldsymbol{\mathcal{S}}^j \in \mathbb{R}^{sM}$ is formed by stacking the stress vector at all FE integration points in a single column vector:

$$\boldsymbol{\mathcal{S}}^j := \begin{bmatrix}\boldsymbol{\sigma}^j(\boldsymbol{x}_1)\\ \boldsymbol{\sigma}^j(\boldsymbol{x}_2)\\ \vdots\\ \boldsymbol{\sigma}^j(\boldsymbol{x}_M)\end{bmatrix}. \tag{A.4}$$

Suppose that the SVD is applied to $\boldsymbol{X}_{\sigma}$, and let $\boldsymbol{\Psi} \in \mathbb{R}^{sM\times q}$ and $\{\bar{\lambda}_i\}_{i=1}^{q}$ denote the corresponding left singular vector matrix (of rank q) and their associated singular values, respectively. $\boldsymbol{\Psi}$ (with each column multiplied by its corresponding singular values) can be construed as a "compressed" matrix of stress snapshots, and, accordingly, one may calculate the required internal forces matrix $\boldsymbol{X}_{\mathcal{F}}$ using this compressed version; a generic column of $\boldsymbol{X}_{\mathcal{F}}$ calculated this way is given by

$$\boldsymbol{\mathcal{F}}_I^j = \begin{bmatrix}\sqrt{W_1}\boldsymbol{B}_I^T(\boldsymbol{x}_1)\bar{\lambda}_j\,\boldsymbol{\Psi}_j(\boldsymbol{x}_1)\\ \sqrt{W_2}\boldsymbol{B}_I^T(\boldsymbol{x}_2)\bar{\lambda}_j\,\boldsymbol{\Psi}_j(\boldsymbol{x}_2)\\ \vdots\\ \sqrt{W_M}\boldsymbol{B}_I^T(\boldsymbol{x}_M)\bar{\lambda}_j\,\boldsymbol{\Psi}_j(\boldsymbol{x}_M)\end{bmatrix}, \qquad I = 1,2,\ldots,n, \quad j = 1,2,\ldots,q. \tag{A.5}$$

[21] This assertion is only true in small strains. To derive a similar decomposition for large strains problems, one has to replace the Cauchy stress vector $\boldsymbol{\sigma}$ by the first Piola–Kirchhoff stress vector.

Notice that the size of the "compressed" matrix $\boldsymbol{X}_{\mathcal{F}}$ is $M \times q\, n$, i.e., the number of columns of $\boldsymbol{X}_{\mathcal{F}}$ is equal to the product of two reduced dimensions: the number of displacement modes n and the number of stress modes q.

The benefits of using this alternative method based on determining the dominant stress modes are twofold. Firstly, memory requirements are drastically reduced, since one need not allocate additional memory for storing the reduced internal forces at all FE points and for all training configurations. Secondly, as opposed to the snapshot matrix of reduced internal forces, the snapshot matrix of stresses is independent of the reduced dimensions of the problem. As a consequence, for problems in which $n > s$, it proves more advantageous in terms of computational cost to apply the SVD to the stress matrix $\boldsymbol{X}_{\sigma}$ rather than to the internal force matrix $\hat{\boldsymbol{X}}_{\mathcal{F}}$. Besides, in problems in which the output of interest depends on stress quantities, the stress basis matrix $\boldsymbol{\Psi}$ is needed for reconstructing the stress field anyway, and therefore, this stress-based strategy would not involve additional computational effort.

Appendix B. Partitioned SVD

For large snapshot matrices, attempting to directly calculate the Singular Value Decomposition (SVD) may exceed the memory capabilities of the computer at hand. In what follows, we propose an (approximate) method for calculating the SVD that precludes the necessity of manipulating the whole matrix, and, therefore, can help in diminishing memory requirements. The method is based on the observation that, in general, contiguous columns within the snapshot matrix exhibit high degree of linear correlation.

Suppose we wish to *approximate* the truncated SVD of rank n of a matrix $\mathbf{X} \in \mathbb{R}^{N\times P}$ (with $N \geq P$, and $n \ll P$). The expression for the *exact* factorization reads

$$\mathbf{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T + \boldsymbol{E}, \tag{B.1}$$

where $\boldsymbol{U} \in \mathbb{R}^{N\times n}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{V} \in \mathbb{R}^{P\times n}$ are the truncated matrices of left singular vectors, singular values, and right singular vectors, respectively; $\boldsymbol{E}$, on the other hand, stands for the contribution of the discarded singular values (if any). We begin by partitioning $\mathbf{X}$ into Q block matrices as follows

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_Q\end{bmatrix}. \tag{B.2}$$

Next, we apply the SVD on each block matrix $\mathbf{X}_i$ $(i = 1, 2, \ldots, Q)$, retaining only the first $r_i = \min(\text{rank}(\mathbf{X}_i), n)$ singular values:

$$\mathbf{X}_i = \boldsymbol{U}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i^T + \bar{\boldsymbol{E}}_i \quad (i = 1, 2, \ldots, Q) \tag{B.3}$$

$\boldsymbol{U}_i \in \mathbb{R}^{N\times r_i}$, $\boldsymbol{\Sigma}_i \in \mathbb{R}^{r_i\times r_i}$ and $\boldsymbol{V}_i \in \mathbb{R}^{P\times r_i}$ being the truncated matrices of left singular vectors, singular values, and right singular vectors, respectively (and $\bar{\boldsymbol{E}}_i$ the matrix corresponding to the discarded, if any, trailing singular values). Substitution of Eq. (B.3) into Eq. (B.2) yields

$$\mathbf{X} = \begin{bmatrix}\boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^T & \boldsymbol{U}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^T & \cdots & \boldsymbol{U}_Q\boldsymbol{\Sigma}_Q\boldsymbol{V}_Q^T\end{bmatrix} + \bar{\boldsymbol{E}}, \tag{B.4}$$

where $\bar{\boldsymbol{E}} = \begin{bmatrix}\bar{\boldsymbol{E}}_1 & \bar{\boldsymbol{E}}_2 & \cdots & \bar{\boldsymbol{E}}_Q\end{bmatrix}$. The above expression can be rearranged as follows

$$\begin{aligned}\mathbf{X} &= \overbrace{\begin{bmatrix}\boldsymbol{U}_1\boldsymbol{\Sigma}_1 & \boldsymbol{U}_2\boldsymbol{\Sigma}_2 & \cdots & \boldsymbol{U}_Q\boldsymbol{\Sigma}_Q\end{bmatrix}}^{\bar{\boldsymbol{X}}} \overbrace{\begin{bmatrix}\boldsymbol{V}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{V}_2^T & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{V}_Q^T\end{bmatrix}}^{\breve{\boldsymbol{V}}^T} + \bar{\boldsymbol{E}} \\ &= \bar{\boldsymbol{X}}\breve{\boldsymbol{V}}^T + \bar{\boldsymbol{E}},\end{aligned} \tag{B.5}$$

where $\bar{\boldsymbol{X}} \in \mathbb{R}^{N\times r}$ and $\breve{\boldsymbol{V}} \in \mathbb{R}^{P\times r}$, with $r = \sum_{i=1}^{Q} r_i \leq n\,Q$. The final step consists in obtaining the truncated SVD of $\bar{\boldsymbol{X}}$ (of rank n)

$$\bar{\boldsymbol{X}} = \bar{\boldsymbol{U}}\bar{\boldsymbol{S}}\bar{\boldsymbol{V}}^T + \breve{\boldsymbol{E}}, \tag{B.6}$$

where $\bar{\boldsymbol{U}} \in \mathbb{R}^{N\times n}$, $\bar{\boldsymbol{S}} \in \mathbb{R}^{n\times n}$ and $\bar{\boldsymbol{V}} \in \mathbb{R}^{r\times n}$. Introducing the above factorization into Eq. (B.5), and moving the error terms to the left-hand side, we finally get

$$\mathbf{X} - (\bar{\boldsymbol{E}} + \check{\boldsymbol{E}}) = \bar{\boldsymbol{U}}\bar{\boldsymbol{S}}\boldsymbol{V}^{*T}, \tag{B.7}$$

where $\boldsymbol{V}^{*T} := \bar{\boldsymbol{V}}^T\check{\boldsymbol{V}}^T$. Note that both $\check{\boldsymbol{V}}$ and $\bar{\boldsymbol{V}}$ are columnwise orthonormal matrices; it follows thus that $\boldsymbol{V}^*$ exhibits also this property; indeed

$$\boldsymbol{V}^{*T}\boldsymbol{V}^* = \bar{\boldsymbol{V}}^T(\check{\boldsymbol{V}}^T\check{\boldsymbol{V}})\bar{\boldsymbol{V}} = \bar{\boldsymbol{V}}^T(\boldsymbol{I})\bar{\boldsymbol{V}} = \boldsymbol{I} \tag{B.8}$$

($\boldsymbol{I} \in \mathbb{R}^{n\times n}$ is the identity matrix). Since both $\bar{\boldsymbol{U}}$ and $\bar{\boldsymbol{S}}$ arise from an SVD, and therefore, are columnwise orthonormal and diagonal with positive entries, respectively, it follows from the uniqueness of such a decomposition that the factorization $\bar{\boldsymbol{U}}\bar{\boldsymbol{S}}\boldsymbol{V}^{*T}$ appearing in Eq. (B.7) is indeed the SVD of the matrix $\mathbf{X} - (\bar{\boldsymbol{E}} + \check{\boldsymbol{E}})$. Hence, $\bar{\boldsymbol{U}}\bar{\boldsymbol{S}}\boldsymbol{V}^{*T}$ can be regarded as an *approximation (of order $\mathcal{O}(\|\bar{\boldsymbol{E}} + \check{\boldsymbol{E}}\|)$) to the truncated SVD of rank n of matrix* $\mathbf{X}$. Notice that, if $n = \operatorname{rank}(\mathbf{X})$, then $\bar{\boldsymbol{E}} = \check{\boldsymbol{E}} = \mathbf{0}$, and the described method would provide the exact factorization.

Concerning the memory saves provided by this approximated SVD, notice that it requires Q independent SVDs of matrices of size, on average, $N \times \frac{P}{Q}$, and one additional SVD over a matrix of size $N \times r$, with $r \leq nQ$. So, the larger matrix to be manipulated has only $r \leq nQ$ columns—recall that $n \ll P$ by hypothesis.

Appendix C. Elastic/inelastic dimensionality reduction

Let $\mathbf{X} \in \mathbb{R}^{N\times n}$ be a snapshot matrix, of either displacement, internal forces, or stresses, and let us decompose this matrix as $\mathbf{X} = \begin{bmatrix}\mathbf{X}_e & \mathbf{X}_i\end{bmatrix}$, where $\mathbf{X}_e$ and $\mathbf{X}_i$ stand for the block matrices corresponding to solutions in the elastic and inelastic ranges, respectively. We seek an SVD-like factorization $\mathbf{X} = \bar{\boldsymbol{U}}\bar{\boldsymbol{S}}\bar{\boldsymbol{V}}^T$ (with $\bar{\boldsymbol{U}}$ and $\bar{\boldsymbol{V}}$ orthonormal matrices, and $\bar{\boldsymbol{S}}$ diagonal with positive entries) such that the first n^{el} columns of $\bar{\boldsymbol{U}}$ form a basis matrix for the column space of the elastic matrix $\mathbf{X}_e$.

To this end, we first compute a generic orthogonal basis matrix $\mathbf{D} \in \mathbb{R}^{N\times n^{el}}$ for the column space of $\mathbf{X}_e$ (using, for instance, the SVD itself), and then obtain the matrix arising from projecting $\mathbf{X}$ onto the space spanned by $\mathbf{D}$:

$$\bar{\boldsymbol{X}}_e := \mathbf{D}(\mathbf{D}^T\mathbf{X}). \tag{C.1}$$

In doing so, we can write

$$\mathbf{X} = \bar{\boldsymbol{X}}_e + \bar{\boldsymbol{X}}_i, \tag{C.2}$$

where $\bar{\boldsymbol{X}}_i := \mathbf{X} - \bar{\boldsymbol{X}}_e$. Finally, introducing the SVD of both $\bar{\boldsymbol{X}}_e$ and $\bar{\boldsymbol{X}}_i$ in the preceding equation, we arrive at the desired factorization[22]

$$\begin{aligned}\mathbf{X} &= \bar{\boldsymbol{U}}_e\bar{\boldsymbol{S}}_e\bar{\boldsymbol{V}}_e + \bar{\boldsymbol{U}}_i\bar{\boldsymbol{S}}_i\bar{\boldsymbol{V}}_i = \overbrace{\begin{bmatrix}\bar{\boldsymbol{U}}_e & \bar{\boldsymbol{U}}_i\end{bmatrix}}^{\bar{\boldsymbol{U}}}\overbrace{\begin{bmatrix}\bar{\boldsymbol{S}}_e & \mathbf{0}\\ \mathbf{0} & \bar{\boldsymbol{S}}_i\end{bmatrix}}^{\bar{\boldsymbol{S}}}\overbrace{\begin{bmatrix}\bar{\boldsymbol{V}}_e^T\\ \bar{\boldsymbol{V}}_i^T\end{bmatrix}}^{\bar{\boldsymbol{V}}^T}\\ &= \bar{\boldsymbol{U}}\bar{\boldsymbol{S}}\bar{\boldsymbol{V}}^T.\end{aligned} \tag{C.3}$$

This strategy proves specially advantageous in those small strains, quasi-static problems in which the set of input parameters solely embodies variations of boundary conditions and external forces. Indeed, in such cases, the elastic response can be captured exactly by[23] $n^{el} = n_0 + n_f$ elastic modes and, therefore, by taking $n \geq n^{el}$, the reduced-order model is guaranteed to deliver elastic solutions with the same accuracy as the underlying (full-order) finite element model—in other words, only the solution in the inelastic range is subject to approximation.

[22] Alternatively, when the sole variable of interest is the matrix of left singular vectors $\bar{\boldsymbol{U}}$, the inelastic modes may be obtained by simply calculating the SVD of $\mathbf{X}_i - \mathbf{D}(\mathbf{D}^T\mathbf{X}_i)$.

[23] Here, n_0 and n_f denote the number of parameters used to characterize the Dirichlet boundary conditions and external forces, respectively (see Section 3.2).

C.1. Modal analysis combined with SVD

In small strains, nonlinear vibration problems, a similar decomposition may offer some benefits when seeking dominant displacement modes; however, instead of employing as elastic modes the dominant left singular vectors arising from an SVD of the elastic snapshots, it proves more consistent to use the first n^{vib} *natural vibrational modes* (those with lowest frequencies), and determine the remaining modes using a "weighted SVD" (one that uses as minimization norm the mass matrix $\boldsymbol{M}^h$). The procedure can be sketched as follows: suppose that, after solving the corresponding eigenproblem, we have at our disposal the first n^{vib} natural vibration modes, denoted henceforth by $\boldsymbol{\Phi}^{vib}$ (these modes are assumed to be $\boldsymbol{M}^h$-orthogonal). To determine the remaining $n^{\perp} = n - n^{vib}$ displacement modes, designated by $\boldsymbol{\Phi}^{\perp}$, we obtain first the Cholesky decomposition of the mass matrix $\boldsymbol{M}^h$, i.e., $\boldsymbol{M}^h = \bar{\boldsymbol{M}}^{hT}\bar{\boldsymbol{M}}^h$, and then compute the truncated SVD (of rank $n^{\perp}$) of the matrix defined as

$$\bar{\boldsymbol{X}} = \bar{\boldsymbol{M}}^h \left(\boldsymbol{X}_d - \boldsymbol{\Phi}^{vib}(\boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h\boldsymbol{X}_d) \right). \tag{C.4}$$

The desired basis matrix $\boldsymbol{\Phi}^{\perp}$ finally emerges from making

$$\boldsymbol{\Phi}^{\perp} = \bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{U}}, \tag{C.5}$$

$\bar{\boldsymbol{U}} \in \mathbb{R}^{N\times n^{\perp}}$ being the matrix of left singular vectors arising from the above mentioned SVD. It is easily seen that the resulting basis matrix for displacements

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}^{vib} & \boldsymbol{\Phi}^{\perp} \end{bmatrix} \tag{C.6}$$

is $\boldsymbol{M}^h$-orthonormal. Indeed, by definition, $\boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h\boldsymbol{\Phi}^{vib} = \boldsymbol{I}$. Likewise, from Eq. (C.5), it follows that

$$\begin{aligned} \boldsymbol{\Phi}^{\perp^T}\boldsymbol{M}^h\boldsymbol{\Phi}^{\perp} &= \left(\bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{U}}\right)^T \left(\bar{\boldsymbol{M}}^{hT}\bar{\boldsymbol{M}}^h\right)\left(\bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{U}}\right) \\ &= \bar{\boldsymbol{U}}^T \left(\bar{\boldsymbol{M}}^{h-T}\bar{\boldsymbol{M}}^{hT}\right)\left(\bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{M}}^h\right)\bar{\boldsymbol{U}} \\ &= \bar{\boldsymbol{U}}^T\bar{\boldsymbol{U}} = \boldsymbol{I}. \end{aligned} \tag{C.7}$$

To complete the proof, it only remains to demonstrate that $\boldsymbol{\Phi}^{\perp^T}\boldsymbol{M}^h\boldsymbol{\Phi}^{vib} = \boldsymbol{0}$. Since the column space of $\boldsymbol{\Phi}^{\perp}$ is but a subspace of the column space of $\bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{X}}$, this demonstration boils down to showing that $\boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h(\bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{X}}) = \boldsymbol{0}$:

$$\begin{aligned} \boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h(\bar{\boldsymbol{M}}^{h-1}\bar{\boldsymbol{X}}) &= \boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h \left(\boldsymbol{X}_d - \boldsymbol{\Phi}^{vib}(\boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h\boldsymbol{X}_d)\right) \\ &= \boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h\boldsymbol{X}_d - \overbrace{\boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h\boldsymbol{\Phi}^{vib}}^{\boldsymbol{I}}(\boldsymbol{\Phi}^{vib^T}\boldsymbol{M}^h\boldsymbol{X}_d) \\ &= \boldsymbol{0}. \end{aligned} \tag{C.8}$$

The advantage of this way of determining the displacement modes is that it ensures that the free, elastic vibration behavior corresponding to the first n^{vib} vibration modes is exactly captured—independently of the input parameters employed in the training process.

References

[1] K. Carlberg, C. Bou-Mosleh, C. Farhat, Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations, Internat. J. Numer. Methods Engrg. 86 (2) (2011) 155–181.

[2] D. Ryckelynck, Hyper-reduction of mechanical models involving internal variables, Internat. J. Numer. Methods Engrg. 77 (1) (2009) 75–89.

[3] M. Barrault, Y. Maday, N. Nguyen, A. Patera, An empirical interpolation'method: application to efficient reduced-basis discretization of partial differential equations, C. R. Math. 339 (9) (2004) 667–672.

[4] S. Chaturantabut, D. Sorensen, Discrete empirical interpolation for nonlinear model reduction, in: Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference, CDC/CCC 2009, Proceedings of the 48th IEEE Conference on, IEEE, 2010, pp. 4316–4321.

[5] N. Nguyen, A. Patera, J. Peraire, A best points interpolation method for efficient approximation of parametrized functions, Internat. J. Numer. Methods Engrg. 73 (2008) 521–543.

[6] J. Baiges, R. Codina, S. Idelsohn, Explicit reduced-order models for the stabilized finite element approximation of the incompressible Navier–Stokes equations, Internat. J. Numer. Methods Fluids 72 (12) (2013) 1219–1243.
[7] P. Astrid, S. Weiland, K. Willcox, T. Backx, Missing point estimation in models described by proper orthogonal decomposition, IEEE Trans. Automat. Control 53 (10) (2008) 2237–2251.
[8] R. Everson, L. Sirovich, Karhunen–Loeve procedure for gappy data, J. Opt. Soc. Amer. A 12 (8) (1995) 1657–1664.
[9] K. Carlberg, J. Cortial, D. Amsallem, M. Zahr, C. Farhat, The GNAT nonlinear model reduction method and its application to fluid dynamics problems, in: 6th AIAA Theoretical Fluid Mechanics Conference, vol. 2730, Honolulu, Hawaii, June, 2011, pp. 2011–3112.
[10] S. Chaturantabut, D.C. Sorensen, Application of POD and DEIM on dimension reduction of non-linear miscible viscous fingering in porous media, Math. Comput. Model. Dyn. Syst. 17 (4) (2011) 337–353.
[11] P. Kerfriden, O. Goury, T. Rabczuk, S.P.-A. Bordas, A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics, Comput. Methods Appl. Mech. Engrg. 256 (2013) 169–188.
[12] M. Drohmann, B. Haasdonk, M. Ohlberger, Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation, SIAM J. Sci. Comput. 34 (2) (2012) A937–A969.
[13] D. Galbally, K. Fidkowski, K. Willcox, O. Ghattas, Non-linear model reduction for uncertainty quantification in large-scale inverse problems, Internat. J. Numer. Methods Engrg. 81 (12) (2010) 1581–1608.
[14] A. Radermacher, S. Reese, POD-based model reduction with empirical interpolation applied to nonlinear elasticity, Internat. J. Numer. Methods Engrg. (2015). URL http://dx.doi.org/10.1002/nme.5177.
[15] M. Grepl, Y. Maday, N. Nguyen, A. Patera, Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations, Math. Model. Numer. Anal. 41 (3) (2007) 575–605.
[16] H. Antil, S.E. Field, F. Herrmann, R.H. Nochetto, M. Tiglio, Two-step greedy algorithm for reduced order quadratures, J. Sci. Comput. 57 (3) (2013) 604–637.
[17] J.A. Hernández, J. Oliver, A. Huespe, M. Caicedo, J. Cante, High-performance model reduction techniques in computational multiscale homogenization, Comput. Methods Appl. Mech. Engrg. 276 (2014) 149–189.
[18] T. Aanonsen, Empirical Interpolation with Application to Reduced Basis Approximations, 2009.
[19] S. An, T. Kim, D. James, Optimizing cubature for efficient integration of subspace deformations, ACM Trans. Graph. 27 (5) (2009) 165.
[20] C. Farhat, P. Avery, T. Chapman, J. Cortial, Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency, Internat. J. Numer. Methods Engrg. 98 (9) (2014) 625–662.
[21] C. Farhat, T. Chapman, P. Avery, Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models, Internat. J. Numer. Methods Engrg. 102 (5) (2015) 1077–1110.
[22] C. von Tycowicz, C. Schulz, H.-P. Seidel, K. Hildebrandt, An efficient construction of reduced deformable objects, ACM Trans. Graph. 32 (6) (2013) 213.
[23] T. Kim, J. Delaney, Subspace fluid re-simulation, ACM Trans. Graph. 32 (4) (2013) 62.
[24] Y. Teng, M.A. Otaduy, T. Kim, Simulating articulated subspace self-contact, ACM Trans. Graph. 33 (4) (2014).
[25] C. Von-Tycowicz, C. Schulz, H.-P. Seidel, K. Hildebrandt, Real-time nonlinear shape interpolation, ACM Trans. Graph. 34 (3) (2015) 34.
[26] J. Chadwick, S. An, D. James, Harmonic shells: a practical nonlinear sound model for near-rigid thin shells, ACM Trans. Graph. 28 (5) (2009).
[27] S. Li, J. Huang, F. de Goes, X. Jin, H. Bao, M. Desbrun, Space-time editing of elastic motion through material optimization and reduction, ACM Trans. Graph. 33 (4) (2014) 108.
[28] K. Carlberg, R. Tuminaro, P. Boggs, Preserving lagrangian structure in nonlinear model reduction with application to structural dynamics, SIAM J. Sci. Comput. 37 (2) (2015) B153–B184.
[29] S. Boyaval, Reduced-basis approach for homogenization beyond the periodic setting, 2007. Arxiv preprint math/0702674.
[30] J. Yvonnet, Q. He, The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains, J. Comput. Phys. 223 (1) (2007) 341–368.
[31] E. Monteiro, J. Yvonnet, Q. He, Computational homogenization for nonlinear conduction in heterogeneous materials using model reduction, Comput. Mater. Sci. 42 (4) (2008) 704–712.
[32] N. Nguyen, A multiscale reduced-basis method for parametrized elliptic partial differential equations with multiple scales, J. Comput. Phys. 227 (23) (2008) 9807–9822.
[33] Y. Efendiev, J. Galvis, E. Gildin, Local–global multiscale model reduction for flows in high-contrast heterogeneous media, J. Comput. Phys. 231 (24) (2012) 8100–8113.
[34] Y. Efendiev, J. Galvis, F. Thomines, A systematic coarse-scale model reduction technique for parameter-dependent flows in highly heterogeneous media and its applications, Multiscale Model. Simul. 10 (4) (2012) 1317–1343.
[35] A. Abdulle, Y. Bai, Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems, J. Comput. Phys. 231 (21) (2012) 7014–7036.
[36] A. Abdulle, Y. Bai, Adaptive reduced basis finite element heterogeneous multiscale method, Comput. Methods Appl. Mech. Engrg. 257 (2013) 203–220.
[37] L. Hogben, Handbook of Linear Algebra, Chapman & Hall/CRC, 2006.
[38] W. Press, B. Flannery, S. Teukolsky, W. Vetterling, Numerical Recipes in Fortran 77: The Art of Scientific Computing 1, Cambridge University Press, Cambridge, 1993.
[39] P. Krysl, S. Lall, J. Marsden, Dimensional model reduction in non-linear finite element dynamics of solids and structures, Internat. J. Numer. Methods Engrg. 51 (4) (2001) 479–504.
[40] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge Univ Pr., 2004.
[41] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, Vol. 161, SIAM, 1974.
[42] J.C. Simo, T.J.R. Hughes, Computational Inelasticity, Springer, New York, 1998.
[43] T. Belytschko, W.K. Liu, B. Moran, Nonlinear Finite Elements for Continua and Structures, John Wiley and Sons Ltd., New York, 2001.

B.5 Paper #5

Title: Continuum Approach to Computational Multi-Scale Modeling of Fracture.
Authors:

- **J. Oliver**: Professor of Continuum Mechanics and Structural analysis at the Escola Tecnica Superior d'Enginyers de Camins, Canals i Ports (Civil Engineering School) of the Universitat Politècnica de Catalunya (Technical University of Catalonia BarcelonaTech). Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **E. Roubin**: Maître de conférence at the 3SR (Sols, Solides, Structures et Risques) and the IUT DGGC in Grenoble.
- **A. E. Huespe**: Professor of Mechanics at the Faculty of Chemical Engineering, Dept. of Materials, National University of Litoral, Santa Fe, Argentina. Independent researcher of Conicet at CIMEC (Centro de Investigaciones en Mecánica Computacional), National University of Litoral (UNL).

Key Engineering Materials Vol. 627
Advances in Fracture and Damage Mechanics XIII
Editors: J. Alfaiate and M.H. Aliabadi
ISSN: 1662-9795
Trans Tech Publications
DOI: 10.4028/www.scientific.net/KEM.627.349
Link to Publisher

Continuum approach to computational multi-scale modeling of fracture

J. Oliver[1,a], M. Caicedo[1,b], E. Roubin[1,c] and A.E. Huespe[1,2,d]

[1]Technical University of Catalonia (UPC)/International Center for Numerical Methods in Engineering (CIMNE), Campus Nord UPC, Edifici C-1, 08034 Barcelona, Spain

[2]CIMEC-Universidad del Litoral (UNL), Santa Fe, Argentina

[a]xavier.oliver@upc.edu, [b]mcaicedo@cimne.upc.edu, [c]eroubin@cimne.upc.edu, [d]ahuespe@intec.unl.edu.ar

Keywords: Fracture, computational multi-scale modeling, strain injection, crack path field, Continuum Strong Discontinuity Approach

Abstract. This paper presents a FE^2 multi-scale framework for numerical modeling of the structural failure of heterogeneous quasi-brittle materials. The model is assessed by application to cementitious materials. Using the Continuum Strong Discontinuity Approach (CSD), innovative numerical tools, such as strain injection and crack path field techniques, provide a robust, and mesh-size, mesh-bias and RVE-size objective, procedure to model crack onset and propagation at the macro-scale.

Introduction

In view of the significant impact of low-scale material morphology in regards to macroscopic responses, numerical strategies has been wildly developed in order to both analyze and model the complex relationship between scales. Among the different issues that yet remain to be explored, this paper focuses on softening behavior modeling, i.e. non-smooth behaviors, and complex small scale failure mechanisms. These problematic find their applications, inter alia, in multi-scale modeling of the quasi-brittle fracture of cementitious materials.

The theoretical framework used in this work is based on the so-called (FE^2) methods. A coupled FEM is used in both scales. Homogenized quantities at the lower scale, represented by a so-called failure-cell, are therefore transferred, in a one-way fashion, to material points (Gauss points) of the macroscopic structure. Such a direct homogenization procedure eventually links this failure-cell to the notion of statistical Representative Volume Element (hereafter referred to as RVE). However, in case of softening behavior, the non-smooth kinematical description of the thin scale raises the problematic of existence of an RVE and the lack of objectivity of the response with respect to its size. Solutions have been provided, for example, in the context of regularized non-local damage models [1], or in the context of homogenized discrete models [4].

In contrast, this work is an attempt to address this issue in the context of local and continuum models and the Continuum Strong Discontinuity Approach (hereafter referred to as CSDA) [2]. The different key ingredients of the theoretical and numerical framework, which unlock the different issues and flaws mentioned above are: a) at the macroscopic level, the non-smooth kinematics representation is captured using strain-localization and FE with embedded regularized strong discontinuities, b) at the mesoscopic level, the failure-cell fracture behavior is fitted with cohesive-bands with predefined position, even though there is no “a priori” limitation on the considered failure mechanisms, c) from the standard homogenization at the RVE the usual set of macroscopic quantities, in terms of a strain/stress relationship, emerges, together with, and specifically for this approach, *a physically meaningful internal length linked with the size of the RVE and the activated microscopic fracture mechanism*. This internal length is exported to the upper scale and used as the macroscopic strain localization bandwidth, this providing both an objective regularization parameter ---leading to a macroscopic response completely insensitive to RVE size--- and a consistent up-scaling scheme of the failure material properties ---such as the fracture energy.

Model description

Mesoscopic scale description and homogenization procedure. The dissipative processes that occur at the mesoscale are modeled using cohesive bands (that are much thinner than the cell dimensions) scattered within the matrix, the aggregates and the interfaces between them. In this way, they can model a set of predefined crack patterns including several mechanisms such as percolation of the crack through the matrix (necessary for softening behavior), mortar/aggregate decohesion and rupture within the aggregates. Eventually, as the damage stage increases, a dominant mechanism naturally prevails, thus representing the final pattern of the micro-crack. This mechanism is now referred to as *mesoscopic failure mechanism.*

The meso-structure constitutive behavior is considered elastic outside these cohesive bands, whereas the material exhibits strain softening inside the bands (using a classical isotropic damage model), thus, eventually leading to strain localization. From the CSDA [2] it is well known that this kind of narrow bands represents, in a consistent way, a strong discontinuity in which the regularized parameter is given by its bandwidth, ensuring the equivalence of the fracture energy in both scales.

The approach is based on a first order homogenization of strains under a small strain framework in both scales, ensuring energetic balance through the Mandel-Hill principle). Though the mathematical framework is not developed here, attention is focused on the outcome of this homogenization procedure within the overall framework of the strong discontinuity kinematics. The macroscopic constitutive response is point-wise equivalent to an inelastic law (in an incremental fashion) as a function of the homogenized elastic tangent tensor, $\boldsymbol{C^{hom}}$, and the incremental homogenized inelastic strain rate $\dot{\boldsymbol{\varepsilon}}^{(i)}$ i.e.

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{C^{hom}} : (\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x}) - \dot{\boldsymbol{\varepsilon}}^{(i)}). \tag{1}$$

The main interest of this result lies in the evolution of the latter inelastic strain tensor, as it can be expressed in terms of failure mechanism variables at the meso-scale:

$$\dot{\boldsymbol{\varepsilon}}^{(i)} = \frac{1}{l_\mu} \overline{(\boldsymbol{n} \otimes \dot{\boldsymbol{\beta}})^s}_{S_k}, \tag{2}$$

where the barked notation represents the average value the symmetrical tensor product between the strong discontinuity normal, $\boldsymbol{n}$, and the rate of the displacement jump, $\dot{\boldsymbol{\beta}}$, of each cohesive band, belonging to the manifold of the mesoscopic failure mechanism S_k, i.e. the mesoscopic crack. In addition, and defined as the ration between the measure (volume or area) of the RVE and the measure (surface or length) of S_k, l_μ stands the aforementioned *material internal length*.

Numerical aspects at the meso-cell. The cohesive bands are modeled by quadrilateral elements of very small thickness. Even though more general phenomenological models can be set, herein a constitutive *only tension constitutive model* [2] governs the softening behavior in order, first, to represent the typical tensile failure mechanisms of cementitious materials and second, to exhibit the structural impact of the heterogeneities that can lead to complex macroscopic responses.

The external force for the equilibrium equation is the macroscopic strain tensor since, consistently with the considered homogenization theory, no external force comes into play. Furthermore, and following the usual minimal boundary restrictions at the RVE, Dirichlet boundary conditions, avoiding rigid body motions, are applied.

Finally, it is worth noticing that the mesoscopic failure mechanism propagates naturally through the strain localization that takes place within the cohesive bands. At each time step of the problem, the in-loading cohesive bands define the active failure mechanism and therefore they define the set of elements used to upscale non-linear features such as the characteristic length of Eq. (2).

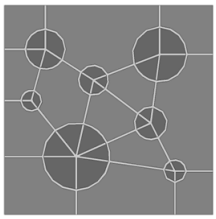
Material failure propagation at the structural level. The *strain injection technique* [5] is used in order to provide a robust and efficient model that can capture failure propagation even with high

strain localization states. Two specific strain-rate modes (*weak and strong discontinuity modes*) are injected within two evolving sets of elements, following criteria based on up-scaled non-linear features, such as bifurcation analysis of the homogenized localization tensor, or based on the dissipated energy. In addition, the crucial matter of positioning strong discontinuities is tackled by a parallel technique termed *crack path field* [5]. This technique uses a directional derivative of a scalar field, based on the internal strain-like variables, whose zero level set defines the crack path.

Based on physical considerations of the crack characteristics ---both in terms of path and kinematics---, the coupling of these methods leads to a numerical framework capable of drastically reducing the classical drawbacks of propagating fracture approaches, such as stress locking, hourglass modes, mesh bias, etc. More details can be found, in a one-scale framework in [5].

Numerical results: application to classical concrete-like simulations

Mesoscopic material properties. The description of the meso-cell is taken to be the same for both tests. The FE discretization is depicted in Fig. 1a where the morphology and the considered pattern of cohesive bands can be seen (from dark to light gray: elastic aggregates, elastic mortar matrix and cohesive bands). The material elastic properties are taken to be the same for elastic and cohesive band elements.



	Mortar	**Aggregates**	
Young Modulus	21 220	100 000	[MPa]
Poisson ratio	0.15	0.15	[-]
Yield stress	3	-	[MPa]
Fracture energy	20	-	[N/m]

Fig. 1. Mesoscopic FE discretization and considered material properties

Nooru-Mohamed test. This test has been simulated following the detail in [3] (load-path 1) where prior to a tension loading, the specimen is subjected to shear load up to 5kN and 10kN. First, a typical mesoscopic failure mechanism is depicted in Fig. 2a, where the activated cohesive bands are clearly shown. Figs. 2b and 2c represent the injection state (light gray for weak and dark gray for strong injection) for two different time steps, thus showing its evolving aspect. By enhancing the kinematics only in a narrow zone around the strain localization, the displacement discontinuity can be well capture and yet, instable modes (hourglass) are avoided. It can be checked that the injection pattern follows the crack path field shown Fig. 2d.

a) b) c) d)

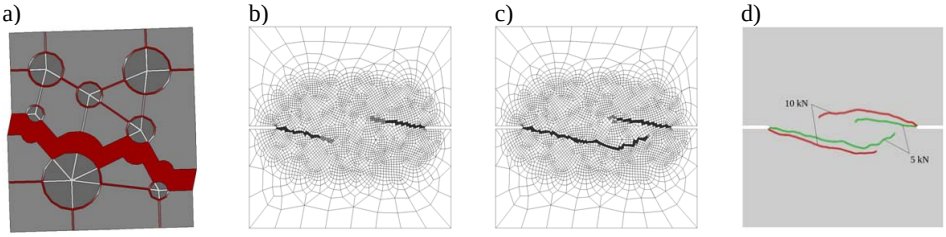


Fig. 2a: a) Damage in cell, b) Strain-injection pattern (5kN), c) Strain-injection pattern (10kN) and d) Crack path fields for 5kN and 10kN.

The latter figure also represents the crack path field in the case of a shear loading of 10kN. Matching the results of [3], a larger spacing between the two cracks along with a larger curvature can be noticed.

Four point bending test. This example shows the impact of the meso-cell failure mechanisms onto the macroscopic response. The meso-cell (Fig. 1) has been slightly modified in order to perform four different tests whose results are depicted in Figs. 3a and 3b, in terms of the crack path field and the macroscopic responses, respectively. The first two meso-cells (top left and top right in Fig. 3a) allow and preclude, respectively, intra-aggregate failure. Therefore, they could correspond to typical failure mechanisms for hard and light-aggregate concrete. The other two tests (bottom right and bottom left) have the cohesive bands *locked for failure* excepting one horizontal and one vertical path, respectively, therefore restricting the set of possible mesoscopic failure mechanisms to a predetermined one. In Fig. 3a, the link of the mesoscopic failure mechanisms and the obtained macroscopic crack path can clearly be observed. Furthermore, in Fig. 3b, the translation of the considered failure-mechanisms into structural responses can be observed.

a) b)

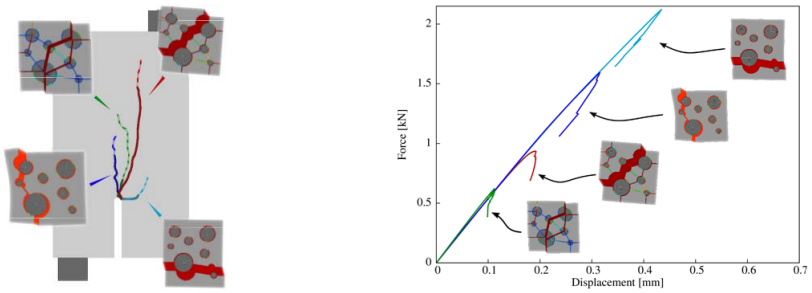


Fig. 3. Four-point bending test: a) crack path fields for different activated mesoscopic failure mechanisms (crack patterns), b) macroscopic (structural) responses

Conclusions

The result of this work is a multi-scale framework that preserves the correct dissipation and objectivity with respect to the size and bias of the FE mesh. Furthermore, the two presented simulations show the, physically meaningful, effects of the activated meso-scale failure mechanisms on the structural response. The proposed framework allows extending the classical FE^2 computational homogenization procedures from smooth problems to fracture mechanics analyzes.

Acknowledgments

The research leading these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP/2007-2013) / ERC Grant Agreement n. 320815, Advanced Grant Project COMP-DES-MAT. The third author would also like to thanks the École Normale Supérieur de Cachan (FR) for their financial support.

References

[1] V.P. Nguyen, O. Lloberas-Valls, M. Stroeven and L. J. Sluys, in: *Computer Methods in Applied Mechanics and Engineering*, (2010)
[2] J. Oliver, A.E. Huespe, M.D.G Pulido and E. Chaves, in: *Engineering Fracture Mechanics*, 2002
[3] M.B Nooru-Mohamed, E. Schlangen and J.G.M. Van Mier, in: *Cement Based Materials*, 1993
[4] P.J. Sánchez, P.J. Blanco, A.E. Huespe, and R.A. Feijóo in: *Comp. Meth. In Appl. Mech. And Eng.*, 2013.
[5] J. Oliver, A. Huespe and I.F. Dias, in: *Comp. Meth. In Appl. Mech. And Eng.*, 2014

B.6 Paper #6

Title: Model Order Reduction in computational multiscale fracture mechanics.
Authors:

- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **J. Oliver**: Professor of Continuum Mechanics and Structural analysis at the Escola Tecnica Superior d'Enginyers de Camins, Canals i Ports (Civil Engineering School) of the Universitat Politècnica de Catalunya (Technical University of Catalonia BarcelonaTech). Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **A. E. Huespe**: Professor of Mechanics at the Faculty of Chemical Engineering, Dept. of Materials, National University of Litoral, Santa Fe, Argentina. Independent researcher of Conicet at CIMEC (Centro de Investigaciones en Mecánica Computacional), National University of Litoral (UNL).
- **O. Lloberas-Valls**: Postdoctoral researcher at the International Center for Numerical Methods in Engineering (CIMNE).

Key Engineering Materials Vol. 713
Advances in Fracture and Damage Mechanics XV
Editors: Jesús Toribio, Vladislav Mantic, Andrés Sáez, M.H. Ferri Aliabadi
ISSN: 1662-9795
Trans Tech Publications
DOI: 10.4028/www.scientific.net/KEM.713.248

Model Order Reduction in computational multiscale fracture mechanics

M. Caicedo[1,a], J. Oliver[1,b], A.E. Huespe[2,c], O. Lloberas-Valls[1,d]

[1] E.T.S. d'Enginyers de Camins, Canals i Ports, Technical University of Catalonia (BarcelonaTech)/International Center for Numerical Methods in Engineering (CIMNE) Campus Nord UPC, Edifici C-1, c/ Jordi Girona 1-3, 08034 Barcelona, SPAIN.

[2] CIMEC-UNL-CONICET, Gemes 3450, Santa Fe, Argentina.

[a]mcaicedo@cimne.upc.edu,[b]oliver@cimne.upc.edu,[c]ahuespe@intec.unl.edu.ar,[d]olloberas@cimne.upc.edu

Keywords: Model Order Reduction, Reduced-Order Cubature, Fracture, Computational Multi-scale modeling, Continuum Strong Discontinuituy Approach.

Abstract. Nowadays, the model order reduction techniques have become an intensive research field because of the increasing interest in the computational modeling of complex phenomena in multi-physic problems, and its consequent increment in high-computing demanding processes; it is well known that the availability of high-performance computing capacity is, in most of cases limited, therefore, the *model order reduction* becomes a novelty tool to overcome this paradigm, that represents an immediately challenge in our research community. In computational multiscale modeling for instance, in order to study the interaction between components, a different numerical model has to be solved in each scale, this feature increases radically the computational cost. We present a reduced model based on a multi-scale framework for numerical modeling of the structural failure of heterogeneous quasi-brittle materials using the Strong Discontinuity Approach (CSD). The model is assessed by application to cementitious materials. The *Proper Orthogonal Decomposition* (POD) and the *Reduced Order Integration Cubature* are the proposed techniques to develop the reduced model, these two techniques work together to reduce both, the complexity and computational time of the *high-fidelity* model, in our case the FE^2 standard model.

1 Introduction

The present model departs from the multiscale framework developed in [2] for the numerical modeling of failure via hierarchical multi-scale models, taking advantage of the reduced order techniques developed in [1], the theoretical framework used in this work is based on the so-called (FE^2) methods via *first order computational homogenization* for the coupling between scales, in which homogenized quantities at the lower scale, represented by a so-called failure-cell, are therefore transferred, in a one-way fashion, to material points (Gauss points) of the macroscopic structure. The formulation is presented in terms of strains in a non-conventional format imposing the *natural* multiscale boundary conditions via Lagrange multipliers.

This work attempts to solve the problematic of excessive computational time in multi-scale models, in our case an additional complexity is induced by the discontinuous displacement field produced by the strain localization at both scales. Nonetheless, the methodology can also be straightforward extended to problems with continuous fields.

2 Model description

2.1 Generalities of the FE^2 method applied to multiscale fracture problems

This approach is developed under a small strain framework, the equality of internal power at both scales is guaranteed via *Hill-Mandell Macro-Homogeneity principle*. In virtue of the finite element method, the dissipative processes that occur at the meso-scale are modeled using *cohesive bands*, represented by quadrilateral elements endowed with a regularized continuum damage model. These bands are characterized by a *high aspect ratio* (width smaller than its length and, in turn, that width being much thinner than the representative cell dimensions). In addition, scattered within the matrix, the aggregates and the interfaces between them are also included. In this way, they can model a set of predefined crack patterns including several mechanisms such as percolation of the crack through the matrix (necessary for softening behavior), mortar/aggregate decohesion and rupture between aggregates. Depending on the loading process at the large scale, these crack patterns are loading and unloading until the full consolidation, finally, a dominant mechanism naturally prevails, thus representing the final pattern of the micro-crack. This mechanism is now referred to as *mesoscopic failure mechanism*. That mechanism has several features, its form and orientation will be as precise as the richness of the lower scale, and is closely related to the crack orientation obtained at the large scale.

In this approach, the macroscopic constitutive response is proven to be point-wise equivalent to an inelastic law (in an incremental fashion) as a function of the homogenized elastic tangent tensor, $\mathbf{C}^{hom}$, and the incremental homogenized inelastic strain rate $\dot{\boldsymbol{\varepsilon}}^{(i)}$ i.e.:

$$\dot{\boldsymbol{\sigma}} = \mathbf{C}^{hom} : (\dot{\boldsymbol{\varepsilon}}(x) - \dot{\boldsymbol{\varepsilon}}^{(i)}) \qquad \dot{\boldsymbol{\varepsilon}}^{(i)} = \frac{1}{l_\mu}(\mathbf{n} \otimes \dot{\boldsymbol{\beta}}) \tag{1}$$

Where, the inelastic strain component $\dot{\boldsymbol{\varepsilon}}^{(i)}$ is expressed as a function of the homogenized variables taken from the lower scale, and represent the average value of the symmetrical tensor product between the strong discontinuity normal $\mathbf{n}$, and the rate of displacement jump $\dot{\boldsymbol{\beta}}$ of each cohesive band, belonging to the manifold of the mesoscopic failure mechanism $\mathcal{S}_\mu$, i.e. the mesoscopic crack. In addition, the so-called *material characteristic length* l_μ is defined as the ratio between the measure (volume or area) of the representative volume and the measure (surface or length) of the mesoscopic failure mechanism. The equations that govern the lower scales are the next:

PROBLEM I: Given a macroscale strain $\boldsymbol{\varepsilon}$, Find $\tilde{\boldsymbol{u}}_\mu$ such that $\boldsymbol{\varepsilon}_\mu = \boldsymbol{\varepsilon} + \nabla^s \tilde{\boldsymbol{u}}_\mu$ and:

$$\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}_\mu) : \nabla^s \tilde{\boldsymbol{u}}_\mu \, d\mathcal{B}_\mu = 0 \quad ; \forall \tilde{\boldsymbol{u}}_\mu \in \mathcal{V}_\mu^u := \{\tilde{\boldsymbol{u}}_\mu \quad | \quad \int_{\mathcal{B}_\mu} \nabla^s \tilde{\boldsymbol{u}}_\mu \, d\mathcal{B}_\mu = \mathbf{0}\}; \tag{2}$$

Regarding the large scale (macro-scale), it is modeled via the finite element method. The strain injection technique [3] is used in order to provide a robust and efficient model that can capture failure propagation even in high strain localization scenarios. In addition, the crucial matter of positioning strong discontinuities is tackled by a parallel technique termed *crack-path field*. This technique uses a directional derivative of a scalar field, based on a location variable (in our case, the average of mesoscale dissipated energy) whose zero level set defines the crack path.

2.2 Model Order Reduction techniques

The reduction process is divided into two sequential stages. The first stage consists of a common Galerkin projection, via *Proper Orthogonal Decomposition* POD for the meso-scale strain field, onto a small space (*reduced-order space*). For the second stage, the main goal is to reduce the number of integration points given by the standard Gauss quadrature, by defining a new scheme that efficiently determines optimal points and its corresponding weights so that the error in the integration of the reduced model is minimized (*Reduced Order Cubature* - ROC).
In order to provide the reduced model with the input parameters and entities, the general procedure is also divided into two parts, the first one (*offline part*) in which the projection operators for the meso-scale strain field and the parameters of the new integration cubature are computed. These data, together with the material and geometrical parameters, define the set of input parameters for the first and second stage (online part).
By comparison with the standard (FE^2) scheme, the proposed model in (2) can be redefined in term of strains in a generalized fashion, imposing the kinematic constraint (2-b) in an explicit way via Lagrange multipliers.

PROBLEM IB: Given a macro-scale strain $\boldsymbol{\varepsilon}$, find $\tilde{\boldsymbol{\varepsilon}}_\mu$ and $\boldsymbol{\lambda}$ satisfying:

$$(\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{\varepsilon}, d_\mu), \boldsymbol{\lambda}(\boldsymbol{\varepsilon}, d_\mu)) = \arg\{\min_{\tilde{\boldsymbol{\varepsilon}}_\mu} \max_{\boldsymbol{\lambda}} \Pi(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda})\}; \quad \text{such that} \quad \dot{d}_\mu(\boldsymbol{y}, \boldsymbol{\varepsilon}_\mu) = g(\boldsymbol{\varepsilon}_\mu, d_\mu) \tag{3}$$

Where Π is the homogenized potential of energy in the meso-scale, expressed in the following way:

$$\Pi(\tilde{\boldsymbol{\varepsilon}}_\mu, \boldsymbol{\lambda}) = \int_{\mathcal{B}_\mu} \psi_\mu(\tilde{\boldsymbol{\varepsilon}}_\mu) \, d\mathcal{B}_\mu + \boldsymbol{\lambda} \int_{\mathcal{B}_\mu} \tilde{\boldsymbol{\varepsilon}}_\mu \, d\mathcal{B}_\mu \qquad \psi_\mu(\tilde{\boldsymbol{\varepsilon}}_\mu) = \frac{1}{2}(1 - d_\mu)(\boldsymbol{\varepsilon} + \tilde{\boldsymbol{\varepsilon}}_\mu) \cdot \mathbb{C}_\mu^{el} \cdot (\boldsymbol{\varepsilon} + \tilde{\boldsymbol{\varepsilon}}_\mu) \tag{4}$$

Being ψ_μ, d_μ and $\mathbb{C}_\mu^{el}$ the internal energy, the damage internal variable and the elastic constitutive tensor at each point $\boldsymbol{y}$ in the meso-scale, respectively.

2.2.1 Projection of strain field via POD

The reduction of the meso-scale strain field is based on the projection of the weak form of the discrete mechanical problem into a reduced *manifold* (reduced-order space), this reduced space is spanned by Ritz (globally supported) basis functions obtained via *Singular Value Decomposition* (SVD) of a set of *snapshots* taken from *training tests* computed during the offline part. Following this reasoning, the meso-scale strain fluctuation can be expressed as:

$$\tilde{\boldsymbol{\varepsilon}}_\mu(\boldsymbol{y}, t) = \sum_{i=1}^{n_{\boldsymbol{\varepsilon}}} \boldsymbol{\Phi}_i(\boldsymbol{y}) c_i(t) = \boldsymbol{\Phi}(\boldsymbol{y}) \boldsymbol{c}(t) \tag{5}$$

Where $\boldsymbol{c}(t) = \{c_1, c_2, c_3, \ldots, c_{n_{\boldsymbol{\varepsilon}}}\}$ is time dependent ($\boldsymbol{c} \in \mathbb{R}^{n_{\boldsymbol{\varepsilon}}}$) and represents the amplitude of the corresponding meso-scale strain mode updated during the online part. Now, introducing (5) and (4) into the **PROBLEM IB** and, after some straightforward manipulations, results into a new model written in terms of the reduced basis:

PROBLEM II: Given a macro-scale strain $\boldsymbol{\varepsilon}$, find $\boldsymbol{c} \in \mathbb{R}^{n_{\boldsymbol{\varepsilon}}}$ satisfying:

$$\int_{\mathcal{B}_\mu} \boldsymbol{\Phi}^T [\boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \boldsymbol{\Phi}\boldsymbol{c}) + \boldsymbol{\lambda}] \, d\mathcal{B}_\mu = \mathbf{0}; \quad \text{tal que} \quad \int_{\mathcal{B}_\mu} \boldsymbol{\Phi}(\boldsymbol{y}) \, \boldsymbol{c}(t) \, d\mathcal{B}_\mu = \mathbf{0}; \tag{6}$$

Solving the system of equations (6) for $\boldsymbol{c}$ and $\boldsymbol{\lambda}$ (Lagrange multiplier to ensure the equality of internal power at both scales via Hill-Mandel Macro-Homogeneity principle), it can be immediately noticed that this problem with $n_{\boldsymbol{\varepsilon}}+n_{\sigma}$ equations will be cheaper, (in computational cost terms), than the standard (FE^2) framework. However, the matricial form of **PROBLEM II** has to be computed (in a standard way) prior its projection onto the reduced-order space. This fact highlights that the actual bottleneck for fast online computation is not the solution of the discrete balance equations but, rather, the determination of the stresses, internal forces and stiffness matrices at all the integration points of the underlyung finite element mesh and its posterior projection. Alternatively, this approach proposes a second stage based on the **PROBLEM II**, that intends to reduce the amount of integration points in which the constitutive equation is computed.

2.2.2 Hyper-reduced model

As pointed out in the previous section, the next objective is to introduce an additional reduction step to diminish the computational burden for equation (6-a). In addition, in order to guarantee the good performance for the second stage, all possible operators have to be computed during the offline part. Particularly, the term (6-b) can be computed entirely in the offline part. To persue the main objective of the second stage, we develop a *Hyperreduced Order Model* (HPROM) via Reduced Optimized Cubature (ROC), this technique is based on a discrete minimization problem that allows determining the optimized location of integration points and the corresponding weights. Once these positions and weigths are at one's disposal, the equation (6-a) can be easily determined as:

$$\int_{\mathcal{B}_\mu} \boldsymbol{\Phi}^T[\boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon}+\boldsymbol{\Phi}\boldsymbol{c})]\,d\mathcal{B}_\mu \approx \sum_{j=1}^{n_r}(\Phi(\boldsymbol{z}_j)^T\boldsymbol{\sigma}_\mu(\boldsymbol{z}_j,\boldsymbol{c}))\,\omega_j \tag{7}$$

The success of our proposed scheme, relies on the fact that it is possible to find a set of integration points n_r, substantially smaller than the ones given by the Gauss standard quadrature, minimizing the error in the assessment of (6). Introducing the expression (7) into the **PROBLEM II**, we get:

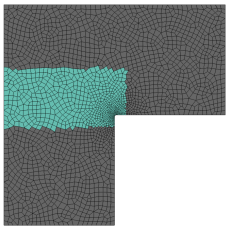
PROBLEM III: Given the macro-scale strain $\boldsymbol{\varepsilon}$, find $\boldsymbol{c}\in\mathbb{R}^{n_{\boldsymbol{\varepsilon}}}$ satisfying:

$$\sum_{j=1}^{n_r}(\boldsymbol{\Phi}(\boldsymbol{z}_j)^T\boldsymbol{\sigma}_\mu(\boldsymbol{z}_j,\boldsymbol{c}))\,\omega_j + \int_{\mathcal{B}_\mu}\boldsymbol{\Phi}^T\boldsymbol{\lambda}\,d\mathcal{B}_\mu = 0; \quad \text{tal que} \quad \int_{\mathcal{B}_\mu}\boldsymbol{\Phi}(\boldsymbol{y})\,\boldsymbol{c}(t)\,d\mathcal{B}_\mu = \boldsymbol{0}; \tag{8}$$

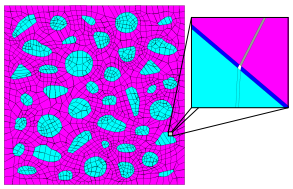
3 Numerical Results

3.1 Application to simulation of fracture in cementitious materials

The macro-scale will be splitted into two subdomains, the dark gray domain will be modeled using an elastic monoscale constitutive law, taking the elastic homogenized constitutive tangent tensor, and, in the green domain the Hiper-Reduced Order Model (HPROM). The finite element mesh of the meso-scale is also depicted in figure (1-b), where the morphology and the considered pattern of cohesive bands can be seen (in light green the aggregates, in pink the mortar matrix, in light gray the *intra-matrix* cohesive bands, and finally, in blue the interface *matrix/aggregate* cohesive bands). The mechanical behavior of the cohesive bands is modeled by a rate-independent continuum damage model endowed with a linear isotropic regularized softening, whose material properties have been taken from the experimental test in [4].



(a) Macroscale FE discretization



(b) Meso-scale FE discretization

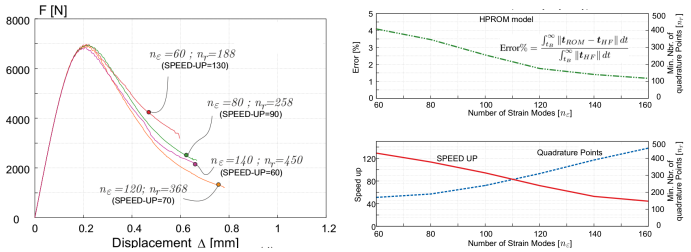
Mortar Matrix			
Young modulus	E_m	[MPa]	18500
Poisson's ratio	ν_m	[-]	0.18
Tensile strength	$f_{t,m}$	[MPa]	2.60
Fracture energy	$G_{f,m}$	[N/m]	140
Aggregates			
Young modulus	E_a	[MPa]	37000
Poisson's ratio	ν_a	[-]	0.18
(Unger & Eckard 2011)			

(c) Material properties

Figure 1: Finite element discretization and material properties

The figure (2-a) shows the structural response in terms of load-displacement ($\mathbf{P}-\delta$) curve (vertical load of the bottom, rightmost corner node versus displacement at the same place) for each set of strain modes $n_{\boldsymbol{\varepsilon}}$ and integration points n_r. It is shown the sensitivity in the convergence of the structural behavior as $n_{\boldsymbol{\varepsilon}}$ increases. As a consequence, the

amount of integration points will increase in order to guarantee the numerical stability and the exact integration of the scheme. For instance, with $n_{\boldsymbol{\varepsilon}} = 60$ and $n_r = 188$, we get a speed-up of 130 times with respect to the time consumed by the standard (FE^2) formulation. In this case the reduced model matches the peak load of the experimental test; however, the softening branch (post-peak behavior) is not as precise as expected, the improvement of this branch is achieved increasing the amount of strain modes in the online part.
Lastly, in figure (2-b) it can be observed the convergence results for the meso-scale tests using the Hiper-Reduced Order Model; fixing a number of strain modes $n_{\boldsymbol{\varepsilon}}$, we get an optimal number of integration points for the second stage, (with an error less than 5% in homogenized tractions for the CSDA taking as a reference the HF solution). In addition, it can be immediately noticed that, as the number of strain modes $n_{\boldsymbol{\varepsilon}}$ increases, the error decreases monotonically. However, the rate of decrement in the error is considerably smaller than the rate of increment of integration points, leading to an increment in the computational cost of the HPROM model, therefore, the imposition of a judicious equilibrium between error and number of integration points plays an important role in the good performance of our method, specially in low-performance computing platforms.



(a) Macro-structural response - L Shape Panel (b) Convergence analysis of the meso-scale

Figure 2: Convergence error in macro and meso scales

4 Conclusions

The result of this work is a reduced model based on a hierarchical (FE^2) multiscale approach for material failure in cementitious materials, that preserves all features of the standard FE model [2]. Furthermore, the two presented simulations show the convergence of the meso-scale and the sensitivity of the macro-structural behavior, as a function of the amount of strain modes, $n_{\boldsymbol{\varepsilon}}$, and the number of integration points, n_r. The reduced model solves the problem of unafordable computational cost widely known in multiscale hierachical (FE^2) approaches. This methodology can be straighforward extended to problems with smooth fields.

5 Acknowledgments

The research leading to these results has received funding from, on the one hand, the European Research Council under the European Union's Seventh Framework Program (FP/2007-2013) / ERC Grant Agreement n. 320815, Advanced Grant Project COMP-DES-MAT, and, on the other hand, the Spanish Ministry of Economy and Competitiveness through the National Research Plan 2014: MAT2014-60919-R. Oriol Lloberas-Valls gratefully acknowledges the funding received from the Spanish Ministry of Economy and Competitiveness through the "Juan de la Cierva" Postdoctoral Junior Grant: JCI-2012-13782 and the National Research Plan 2014: MAT2014-60919-R.

References

[1] J.A. Hernndez, J. Oliver, A.E. Huespe, M.A. Caicedo, and J.C. Cante. High-performance model reduction techniques in computational multiscale homogenization. *Computer Methods in Applied Mechanics and Engineering*, 276:149–189, 2014.

[2] J. Oliver, M. Caicedo, E. Roubin, A. E. Huespe, and J.A. Hernandez. Continuum approach to computational multi-scale modeling of propagating material failure. *Computer Methods in Applied Mechanics and Engineering*, 294:384–427, 2015.

[3] J. Oliver, I.F. Dias, and A.E. Huespe. Crack-path field and strain-injection techniques in computational modeling of propagating material failure. *Computer Methods in Applied Mechanics and Engineering*, 274:289–348, 2014.

[4] J. F. Unger and S. Eckardt. Multiscale modeling of concrete - from mesoscale to macroscale. *Archives of Computational Methods in Engineering*, 18:341–393, 2011.

B.7 Chapter in Book

Title: Multi-scale (FE^2) analysis of material failure in cement/aggregate-type composite structures
Authors:

- **J. Oliver**: Professor of Continuum Mechanics and Structural analysis at the Escola Tecnica Superior d'Enginyers de Camins, Canals i Ports (Civil Engineering School) of the Universitat Politècnica de Catalunya (Technical University of Catalonia BarcelonaTech). Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).
- **M. Caicedo**: PhD Candidate in Structural Analysis in UPC BarcelonaTech and International Center for Numerical Methods in Engineering (CIMNE).
- **A. E. Huespe**: Professor of Mechanics at the Faculty of Chemical Engineering, Dept. of Materials, National University of Litoral, Santa Fe, Argentina. Independent researcher of Conicet at CIMEC (Centro de Investigaciones en Mecánica Computacional), National University of Litoral (UNL).
- **E. Roubin**: Maître de conférence at the 3SR (Sols, Solides, Structures et Risques) and the IUT DGGC in Grenoble.
- **J. A. Hernández**: Assistant Professor of Structural Engineering and Strength of Materials at the School of Industrial and Aeronautic Engineering of Terrassa, of the Technical University of Catalonia. Senior researcher at the International Center for Numerical Methods in Engineering (CIMNE).

Computational Modelling of Concrete Structures
Proceedings of EURO–C 2014
Editors: N. Bićanić; H Mang; Gunther Meschke; Reneé de Borst
ISBN: 978-1-138-00145-9
Taylor & Francis Group: CRC Press–Balkema, 2014

Multi-scale (FE^2) analysis of material failure in cement/aggregate-type composite structures

J. Oliver, M. Caicedo, E. Roubin, J.A. Hernández
Technical University of Catalonia (UPC)/International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain

A. Huespe
CIMEC-Universidad del Litoral (UNL), Santa Fe, Argentina

ABSTRACT: The work proposes a FE^2 multiscale approach to computational modeling of material failure in concrete-like structures, made of cement/aggregate-type composite materials. Keeping the approach in a classical homogenization setting, a multiscale model is proposed, which naturally provides a microscopic length-scale to be exported to the macrostructure. There, this length scale is used as regularization parameter in the context of the Continuum Strong Discontinuity Approach to material failure, and finite elements with embedded strong discontinuities (E-FEM). The resulting technique allows robust modeling of crack propagation at the structural scale, accounting for the mesostructure morphology, supplies proper energy dissipation and solutions independent of the finite element and RVE sizes. Application to a number of examples, in the range from light-aggregate concrete to regular concrete, shows the potentiality of the method.

1 INTRODUCTION

Two-scale computational modeling of materials is a subject of increasing interest in computational mechanics. When dealing with materials displaying a spatially smooth behavior there is wide consensus, and some suitable mechanical approaches to the problem are available in the literature. For instance, the so-called *FE^2 methods*, based on the hierarchical, bottom-up one-way coupled, description of the material using the finite element method in both scales, and computational homogenization procedures at the low scale, is nowadays one of the most popular approaches. At the heart of the direct computational homogenization procedure lies the notion of representative volume element (RVE) defined as the smallest possible region representative of the whole heterogeneous media on average.

Alternatively, two-scale computational modeling for *material failure analysis* is more controversial, and exhibits additional complexity. Either if *discrete approaches* (based on non-linear softening cohesive models), or *continuum approaches* (strain localization-based or regularized models) are used at the lower scale, the kinematic description of some, or both, scales cannot be considered smooth anymore, and the existence of the RVE can be questioned arguing that, in this case, the material loses the statistical homogeneity. A crucial consequence of this issue is the lack of objectivity of the results with respect to the size of the RVE. In (Nguyen et al., 2010) a recent attempt to overcome this flaw, for regularized non-local models, can be found.

This work is an attempt to address this issue in the setting of the Continuum Strong Discontinuity Approach (CSDA) to material failure, developed by the authors in the past (Oliver, 1995, Oliver et al., 2002). The essentials of the method are:

1) At the macroscopic level, material failure is captured via *strain-localization and finite elements with embedded regularized strong discontinuities.*

2) The microstructure of the smooth-strain part of the body is represented by a classical RVE, whose size is associated to standard statistic representativeness concepts.

3) A *failure-cell* at the microscopic scale, with the same size and topological properties than the RVE is associated to material points at the strain-localizing part of the microstructure. This failure-cell is enriched with appropriated material failure mechanisms with, apparently, no restriction on their type. Though, for the sake of simplicity, cohesive-bands with a predefined position have been used in this work, there is no "a priori" limitation on using more sophisticated material failure mechanisms, e.g. arbitrarily propagating cracks or strong discontinuities (Oliver, 1995, Armero and Garikipati, 1996,

Alfaiate et al., 2003). In contrast, this failure-cell is not claimed to be a RVE, in the sense of being statistically representative of any part of the macrostructure, although standard homogenization procedures are applied to it.

4) It is proven that homogenization of the RVE and failure-cell returns a macroscopic constitutive model (stress vs. strain) with the same format than classical inelastic-strain-based phenomenological models. A set of macroscopic inelastic-strain-like internal variables emerge naturally, whose evolution equation is ruled by the activation of material non-linearities and failure mechanisms at the failure-cell. In addition, a internal-length arises from that homogenization procedure, and it is naturally determined by the size of the chosen RVE and the amount of activated material failure mechanisms at the microstructure. This internal length is of the same order then the RVE and, determines the bandwidth of the macroscopic regularized strain-localization or displacement-discontinuity bands.

5) Based on this internal length, imported from the microstructure, the macrostructure is equipped with finite elements with embedded regularized strain-localization and displacement-discontinuities. Through this method, complete insensitivity of the structural response, with respect to the RVE size, and the macroscopic and microscopic finite element meshes is achieved, and material failure properties, like the fracture energy, are consistently up-scaled.

The mechanical derivation of the model is done in section 2, in section 3 the numerical aspects are presented and, in section 4, representative simulations are presented. Finally, section 5 is devoted to some concluding remarks.

2 MULTISCALE MODEL FOR FRACTURE

Let us consider the body $\mathcal{B}$, in Figure 1. At the macroscopic scale, the body is considered to exhibit strain-localization bands, with typical bandwidth h (very small in comparison with the dimensions of the macrostructure L, $h \ll L$), which, in turn, are h-regularizations of the discontinuities (cracks) observed at the macroscopic scale.

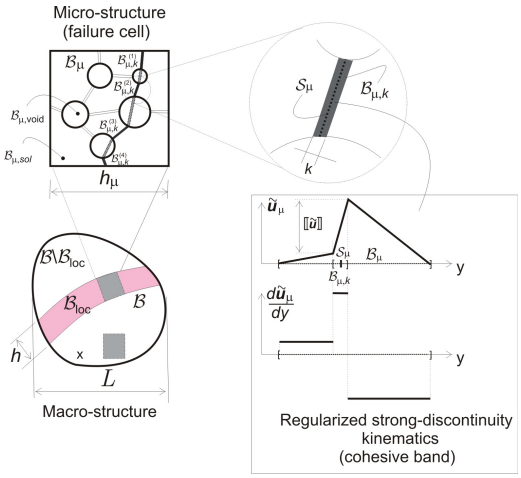


Figure 1. Multiscale failure model

2.1 *Microscopic scale: RVE/failure cell*

We assume microstructures characterized by a RVE/failure-cell like that shown in Figure 2, having cohesive bands, $\mathcal{B}_{\mu,k}$, of very small thickness ($k << h_\mu$) where the dissipative processes can take place and strain localizes.

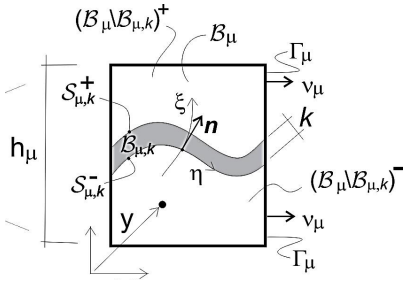


Figure 2. Multiscale model: microstructure with cohesive bands

2.1.1 *Kinematics: microscopic strain field*

Without loss of generality, we are considering cohesive bands of constant width k. Thus, we define the band length $meas(\mathcal{S}_\mu) = \mathcal{B}_{\mu,k} / k$. Then, we assume that: $meas(\mathcal{B}_{\mu,k}) << meas(\mathcal{B}_\mu)$. In $\mathcal{B}_\mu$, we define the microscopic displacement:

$$\mathbf{u}_\mu(\mathbf{y}) = \mathbf{u}(\mathbf{x}) + \varepsilon(\mathbf{x}) \cdot \mathbf{y} + \tilde{\mathbf{u}}_\mu(\mathbf{y}) \tag{1}$$

where, $\mathbf{u}(\mathbf{x})$ and $\varepsilon(\mathbf{x})$ are, respectively, the macroscopic displacement and strain of the point $\mathbf{x}$ in $\mathcal{B}$, respectively, and $\tilde{\mathbf{u}}_\mu$ is the microscopic displacement fluctuation that, for the cohesive bands is defined, in a local coordinate system (ξ, η) aligned with the band $\mathcal{B}_{\mu,k}$ (see Figure 2), as follows:

$$\tilde{\mathbf{u}}(\xi,\eta) = \hat{\tilde{\mathbf{u}}}(\xi,\eta) + H^k_{\mathcal{S}}(\xi)\, \beta(\eta) \tag{2}$$

In equation (2), the displacement $\hat{\mathbf{u}}$ is a smooth field (see Figure 3) whereas the second term of the RHS, is the k-regularized (k-ramp) function $H_{\mathcal{S}}(\xi(\mathbf{y}))$ given by the expressions:

$$H_{\mathcal{S}}(\xi(\mathbf{y})) = \begin{cases} 0 & \forall \quad \mathbf{y} \in (\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,k})^- \\ 1 & \forall \quad \mathbf{y} \in (\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,k}))^+ \\ \dfrac{\xi}{k} & \forall \quad \mathbf{y} \in \mathcal{B}_{\mu,k}) \end{cases}; \tag{3}$$

$$\boldsymbol{\beta}(\eta) \equiv \tilde{\mathbf{u}}_{\mathcal{B}_k^+}(\eta) - \tilde{\mathbf{u}}_{\mathcal{B}_k^-}(\eta).$$

where, $\boldsymbol{\beta}(\eta)$ is the displacement jump across the cohesive band. A sketch of the displacement fluctuation field along ξ is displayed in Figure 3.

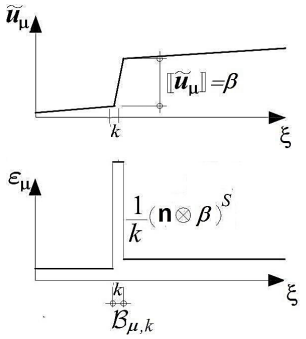


Figure 3. Multiscale model: micro-displacement fluctuations and micro-strains at the cohesive band

From equation (1), the micro strain $\boldsymbol{\varepsilon}_\mu$, assuming infinitesimal strain settings can be obtained as follows:

$$\boldsymbol{\varepsilon}_\mu(\mathbf{y}) = \nabla^s_{\mathbf{y}} \mathbf{u}_\mu = \boldsymbol{\varepsilon}(\mathbf{x}) + \tilde{\boldsymbol{\varepsilon}}_\mu(\mathbf{y}) = \boldsymbol{\varepsilon}(\mathbf{x}) + \nabla^s_{\mathbf{y}} \tilde{\mathbf{u}}_\mu(\mathbf{y}) \tag{4}$$

where $\tilde{\boldsymbol{\varepsilon}}_\mu$ is the micro strain fluctuation and supra-index $(\bullet)^s$ denotes the symmetric part of the corresponding tensor.

2.1.2 *Microscopic constitutive model*

The microstructure is considered constituted of a hardening material outside the cohesive bands, whereas this material exhibits strain-softening inside the bands. Then, the stress-strain relationship, in rate form, is given by:

$$\dot{\boldsymbol{\sigma}}_\mu(\mathbf{y}) = \begin{cases} \mathcal{C}_\mu^{hard} : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{y}); & \forall\, \mathbf{y} \in \mathcal{B}_\mu \setminus \mathcal{B}_{\mu,k} \quad (a) \\ \mathcal{C}_\mu^{soft} : \dot{\boldsymbol{\varepsilon}}_\mu(\mathbf{y}); & \forall\, \mathbf{y} \in \mathcal{B}_{\mu,k} \quad (b) \end{cases} \tag{5}$$

where $\boldsymbol{\sigma}_\mu$ are the micro-stresses, and $\mathcal{C}_\mu^{hard}$ and $\mathcal{C}_\mu^{soft}$ are the microscopic tangent constitutive operators corresponding with the (rate) constitutive behavior at domains $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,k}$ and $\mathcal{B}_{\mu,k}$, respectively (see Figure 4). At $\mathcal{B}_\mu \setminus \mathcal{B}_{\mu,k}$, $\mathcal{C}_\mu^{hard}$ corresponds with a strain-hardening constitutive behavior, so that no strain localization can take place. At $\mathcal{B}_k$, $\mathcal{C}_\mu^{soft}$ defines a strain-softening model which gives raise to strain localization. From the Continuum Strong Discontinuity Approach (CSDA) (Oliver and Huespe, 2004b, Oliver and Huespe, 2004a)) it is well known that, as this strain localization takes place in the narrow bands of bandwidth k (the cohesive bands $\mathcal{B}_{\mu,k}$), regularization of the softening parameter as $H_\mu = k\bar{H}_\mu$; $\bar{H}_\mu = 1/2(\sigma_u^2)/(EG_\mu^f)$, G_μ^f standing for the microscopic fracture energy (see Figure 4), the continuum constitutive model in equation (5)-b, degenerates in an (implicit) traction-separation law, ruling the decohesion of the modeled discrete crack.

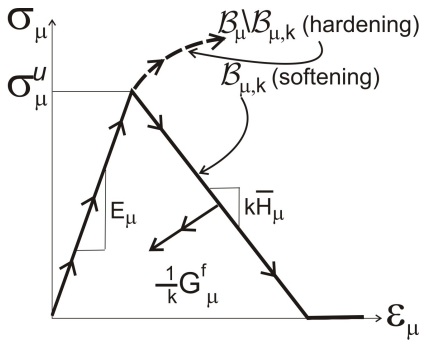


Figure 4. Constitutive model at the microstructure

2.2 *Computational homogenization elements*

Computational homogenization is based on standard procedures lying on the strain homogenization concept and the Mandell-Hill principle (de Souza Neto and Feijóo, 2006) as described next.

2.2.1 *Strain homogenization. Minimal kinematic restrictions*

The strain homogeneity condition requires that[1]:

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \frac{1}{\mathcal{B}_\mu} \int_{\mathcal{B}_\mu} \boldsymbol{\varepsilon}_\mu \, d\mathcal{B}_\mu \tag{6}$$

and replacing (4) in (6), this equation results:

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \boldsymbol{\varepsilon}(\mathbf{x}) + \frac{1}{\mathcal{B}_\mu} \int_{\mathcal{B}_\mu} \nabla^s \tilde{\mathbf{u}}_\mu \, d\mathcal{B}_\mu \tag{7}$$

which, in turn, yields:

$$\int_{\mathcal{B}_\mu} \nabla^s \tilde{\mathbf{u}}_\mu \, d\mathcal{B}_\mu = \int_{\Gamma_\mu} \big(\tilde{\mathbf{u}}_\mu \otimes^s \mathbf{v}_\mu\big) d\Gamma_\mu = \mathbf{0} \tag{8}$$

Thus, we define the space $\boldsymbol{V}_\mu$ of kinematically admissible displacements fluctuations at the micro-scale, as being:

$$\boldsymbol{V}_\mu := \{\tilde{\mathbf{u}}_\mu \quad | \quad \int_{\Gamma_\mu} \big(\tilde{\mathbf{u}}_\mu \otimes^s \mathbf{v}_\mu\big) d\Gamma_\mu = \mathbf{0}\}. \tag{9}$$

[1] From now on, and with some abuse of the notation, we will denote with the same symbol the *domain and its measure*, i.e.: $meas(\mathcal{B}_\mu) \equiv \mathcal{B}_\mu$.

Then, all micro displacement fluctuations: $\tilde{\mathbf{u}}_\mu \in \boldsymbol{V}_\mu$ should be compatible with micro strain fluctuations satisfying the homogeneity condition.

2.2.2 *Hill-Mandel variational principle*

The Hill-Mandel principle equating the macroscopic point internal power and the average value of the microscopic internal power reads

$$\boldsymbol{\sigma}(\mathbf{x}) : \delta\boldsymbol{\varepsilon}(\mathbf{x}) = \frac{1}{\mathcal{B}_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}(\mathbf{y}) : \delta\boldsymbol{\varepsilon}_\mu(\mathbf{y}) d\mathcal{B}_\mu = \\ = \frac{1}{\mathcal{B}_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}(\mathbf{y}) : (\delta\boldsymbol{\varepsilon}(\mathbf{x}) + \nabla^s_{\mathbf{y}}\delta\tilde{\mathbf{u}}_\mu(\mathbf{y})) d\mathcal{B}_\mu \\ \forall\, \delta\boldsymbol{\varepsilon}, \delta\tilde{\mathbf{u}}_\mu \in \boldsymbol{V}_\mu. \tag{10}$$

Replacement of the appropriate values of $\delta\boldsymbol{\varepsilon}, \delta\tilde{\mathbf{u}}_\mu$ in equation (9) returns the classical strain homogenization principle

$$\boldsymbol{\sigma}(\mathbf{x}) = \frac{1}{\mathcal{B}_\mu}\int_{\mathcal{B}_\mu} \boldsymbol{\sigma}_\mu d\mathcal{B}_\mu\,. \tag{11}$$

which supplies the homogenized macro stresses, $\boldsymbol{\sigma}(\mathbf{x})$, and the RVE virtual work principle (de Souza Neto and Feijóo, 2006)

$$\int_{\mathcal{B}_\mu} \nabla^s\delta\tilde{\mathbf{u}}_\mu : \boldsymbol{\sigma}_\mu(\boldsymbol{\varepsilon} + \nabla^s_{\mathbf{y}}\tilde{\mathbf{u}}_\mu)\, d\mathcal{B}_\mu = 0 \\ \forall\, \delta\tilde{\mathbf{u}}_\mu \in \boldsymbol{V}_\mu \tag{12}$$

Equation (12) provides the solution for the micro-fluctuation field $\tilde{\mathbf{u}}_\mu$.

2.3 *Macroscopic scale: homogenized constitutive model*

Substitution of the solution $\tilde{\mathbf{u}}_\mu(\mathbf{y}, \boldsymbol{\varepsilon})$ of equation (12) into equation (11) yields the macroscopic constitutive equation, whose rate form reads

$$\dot{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon}) = \frac{1}{\mathcal{B}_\mu}\int_{\mathcal{B}_\mu} \dot{\boldsymbol{\sigma}}_\mu(\boldsymbol{\varepsilon} + \nabla^s_{\mathbf{y}}\tilde{\mathbf{u}}_\mu(\mathbf{y}, \boldsymbol{\varepsilon}) d\mathcal{B}_\mu\,. \tag{13}$$

Then, algebraic elaboration on equation (13), considering equations (4) and (5), yields the following *macroscopic homogenized constitutive model*

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{\mathcal{C}}^{\text{hom}} : (\dot{\boldsymbol{\varepsilon}}(\mathbf{x}) - \dot{\boldsymbol{\varepsilon}}^{(i)}) \tag{14}$$

where $\boldsymbol{\mathcal{C}}^{\text{hom}}$ is the macroscopic homogenized value of the hardening constitutive operator $\boldsymbol{\mathcal{C}}_\mu^{hard}$ (de Souza Neto and Feijóo, 2006).

In equation (14) $\dot{\boldsymbol{\varepsilon}}^{(i)}$ are macroscopic internal variables, playing the role of inelastic strains, whose evolution is described in terms of the micro-structural variables as

$$\dot{\boldsymbol{\varepsilon}}^{(i)} = \frac{1}{l_\mu}\boldsymbol{\zeta} = \frac{1}{l_\mu}\overline{\left(\mathbf{n} \otimes^s \dot{\boldsymbol{\beta}}\right)}_{\mathcal{S}_k} \tag{15}$$

i.e.: in terms of the activated failure mechanism and the band opening process at the microscale. In equation (15) notation $(\cdot)_{\mathcal{S}_k}$ stands for the mean value of $(\cdot)$ along the microscopic activated failure mechanism, i.e.:

$$\overline{\left(\mathbf{n} \otimes^s \dot{\boldsymbol{\beta}}\right)^S}_{\mathcal{S}_k} = \frac{1}{\mathcal{S}_k}\int_{\mathcal{S}_k}\left(\mathbf{n} \otimes^s \dot{\boldsymbol{\beta}}\right)^S d\mathcal{S} \tag{16}$$

In addition, the evolution equation (15) incorporates the microscale length l_μ, defined as:

$$l_\mu = \frac{\mathcal{B}_\mu}{\mathcal{S}_k} = \mathcal{O}(h_\mu) \tag{17}$$

In summary, in the context of the description of complex materials equipped with morphological descriptors (Oliver et al., 2012, Huespe et al., 2013) equations (14) to (17) retrieve a *constitutive equation equipped with an internal length and with internal variables described by the microstructure behavior.*

3 NUMERICAL ASPECTS. FINITE ELEMENT MODEL

The multiscale formulation described above is implemented by means of a finite element model following the outlines given by the FE^2 technology (Özdemir et al., 2008). Accordingly, two finite element models are used, one for the macroscopic structure and another for the microstructure through a discretized RVE/failure-cell.

3.1 *RVE/failure-cell finite element model*

Standard quadrilateral finite elements are adopted for the numerical simulation of the RVE/failure-cell response. The cohesive bands, $\mathcal{B}_{k,\mu}$, are also modeled by quads of very small thickness ($k \ll h_\mu$) as shown in Figure 5. They are equipped with the constitutive models defined in Figure 4 and equation (5), so that only elements on the cohesive band can exhibit strain localization. The model is solved according with the RVE virtual work principle in equation (12).

Notice that the driving force (external action) for the equation solution is the macroscopic strain $\boldsymbol{\varepsilon}$ and no external force is explicitly applied.

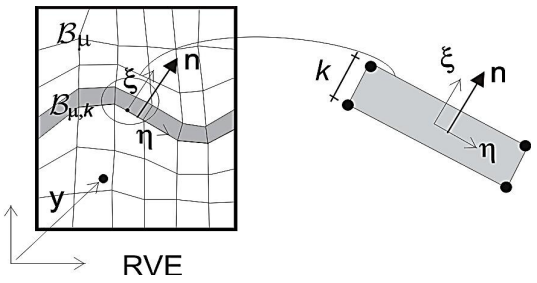


Figure 5. Micro-scale numerical model

The corresponding non-linear problem in the RVE/failure-cell is then solved for the micro-displacements fluctuation discretized as

$$\tilde{\mathbf{u}}_{\mu}(\mathbf{y},\boldsymbol{\varepsilon}) = \sum_{1}^{nnode} N_i(\mathbf{y})\mathbf{d}_{\mu,i}(\boldsymbol{\varepsilon}) \tag{18}$$

where $N_i(\mathbf{y})$ are the standard interpolation functions for quadrilaterals and $\mathbf{d}_{\mu,i}$ are the nodal values. Dirichlet boundary conditions precluding the rigid body motions and, in addition, the minimal boundary restriction in equations (9) are also imposed.

Then, material failure propagates naturally, as the softening behavior and, therefore, strain localization, is incepted in the finite elements defining the cohesive bands. At every time step of the analysis, those elements who are in in-loading state, define the active set of cohesive bands $\mathcal{B}_{\mu,k}$ (see Figure 5) or the *active failure mechanism.*

3.2 *Finite element model at the macro-structure. Material failure propagation.*

At the macroscopic scale, propagation of material failure is modeled using the *crack path field* and *strain injection* techniques recently developed by the authors (Oliver et al., 2014). They are based on using goal oriented assumed-strain fields injected in selected domains, via mixed formulations.

- In a ***first injection stage,*** embedded localization bands, of bandwidth l_{μ}, are incrementally injected (prior to development of displacement discontinuities) in an evolving subdomain of the macro-structure, $\mathcal{B}_{loc}$ (see Figure 6) were both the material discontinuous bifurcation (Willam and Sobh, 1987) an the in-loading conditions are fulfilled.

- These embedded localization bands have an isotropic character (there is no preferred direction in them), but exhibit extremely good ability to propagate material failure in the right material directions. Therefore, its evolving position and intensity is used to determine a scalar field (the crack propagation field) whose zero level set constitutes a reliable approximation to the actual crack path.

- In a ***second injection stage,*** the obtained crack path field, $\mathcal{S}$ (see Figure 6) is used to determine the correct position of an elemental embedded strong-discontinuity strain field, which is incrementally injected in the set of elements, $\mathcal{B}_{dis}$ (see Figure 6).

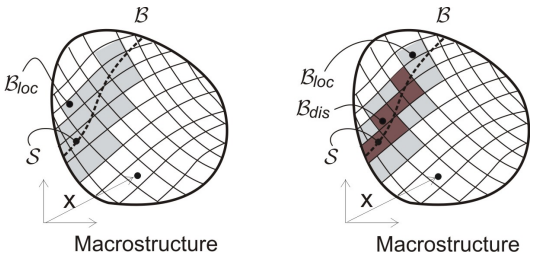


Figure 6. Strain injection procedures. Left: embedded strain localization bands injection in $\mathcal{B}_{loc}$. Right: embedded strong discontinuity injection in $\mathcal{B}_{dis}$

The resulting procedure is a robust and efficient technique to model propagating material failure in a finite element discretized body. Its intra-elemental character (one-element bandwidth band captures the regularized strain-localization) makes it especially appropriate for capturing material failure propagation in coarse meshes, in contraposition of the alternative extra elemental character techniques (e.g. phase field models) for capturing propagating material failure, where several elements have to span the localization band, this leading to, some-times unaffordably fine finite element meshes (B. Bourdin et al., 2000, Miehe et al., 2010). In addition, its implementation in an existing finite element code has a little invasive character.

4 REPRESENTATIVE SIMULATIONS. TWO-SCALE MODELING OF FAILURE PROPAGATION IN A CONCRETE-LIKE MATERIAL.

In order to explore the potential of the proposed formulation, numerical experimentation is performed on the basis of a material whose mesostructure mimics the cement-aggregate composition of concrete. This is not an attempt to obtain quantitative results, to fit some specific experiment, but, rather, to evaluate the potential of multiscale modeling, incorporating the microstructure morphology, in simulation of *complex macroscopic behavior using simple microscopic/mesoscopic models*. It should be mentioned that, unlike what is regularly done in the literature, the issue tackled here is not only modeling the macroscopic homogenized behavior at a single point of the macrostructure, but, also, *the full macroscopic structural behavior, including material failure onset and propagation* and obtaining a *complete action/response behavior during the structural failure*.

The strip in
Figure 7, incorporating a central opening to break the inhomogeneity, is subjected to horizontal stretch-

ing, induced by a constant displacement, Δ, under the action of the external force $\mathbf{F}$.

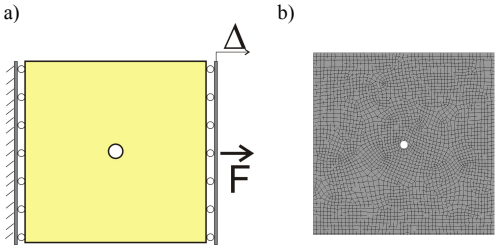


Figure 7. Strip with central opening under uniaxial stretching: a) macroscopic (structural) geometry. Dimensions (1m x1m). b) Finite element mesh (quadrilaterals)

The mesostructure morphology is sketched by the failure-cell of Figure 8 roughly characterized by some cylindrical aggregates of different sizes, immersed in a matrix of mortar. Grains and matrix are numerically modeled by quadrilateral finite elements. Additionally, the failure cell is equipped with a number of, very narrow, cohesive bands; some of them surrounding the aggregates (thus its failure representing possible aggregate/mortar decohesion) and some others crossing the matrix and connecting grains (its failure corresponding to mortar cracks).

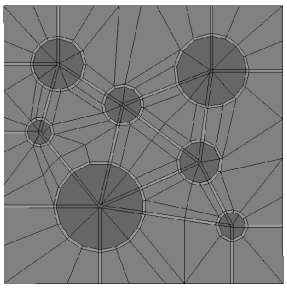


Figure 8. Strip with central opening under uniaxial stretching. 2D microscopic failure cell incorporating: a) cylindrical-shaped aggregates, immersed in a matrix and b) intra-granular, extra-granular and grain-matrix cohesive bands. Dimensions (1mm x 1mm)

Their activation allows modeling a number of failure mechanisms: those excluding the interior of the aggregates i.e.: *extra-granular failure mechanisms* or those including the grains breakage i.e.: *intra-granular failure mechanisms.*

Constitutive behavior at the matrix is assumed to be elastic; at the cohesive bands it is assumed to be ruled by a very simple model: the *only-tension-failure* constitutive model developed by the authors in the past (Oliver et al., 1990) that precludes inelastic behavior in pure compression stress states (see Figure 9).

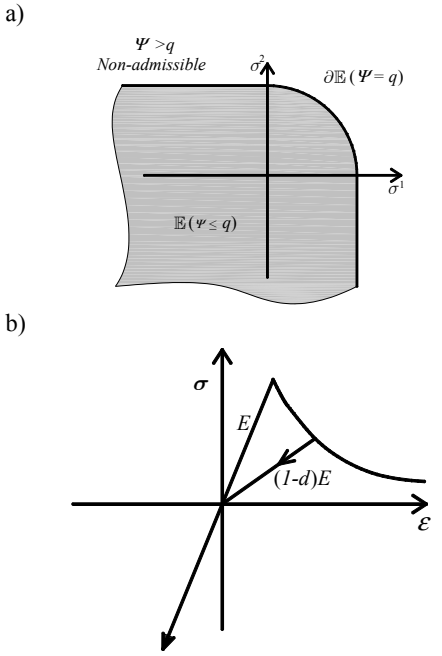


Figure 9. Only-tension-failure continuum damage model: a) elastic domain in the 2D principal stress, $\sigma^{(1)} - \sigma^{(2)}$, space, b) resulting uniaxial stress-strain constitutive model.

Then, three different types of concrete are analyzed, using the mesostructure in Figure 8, equipped with different material properties, i.e.:

- Regular-weight-aggregate (RWA) concrete: properties of matrix and aggregates are taken as those typically described in the literature for standard concrete.
- Light-weight-aggregate (LWA) concrete: properties of aggregates are taken as those typically described in the literature for light (ceramic) aggregates.
- Medium-weight-aggregate (MWA) concrete: aggregate properties are taken in between the other two cases.

The goal of the experiments is to observe how, for the same experiment, the mechanical properties of the aggregates affect the mesoscopic failure mechanism and, in turn, the macroscopic structural response.

4.1 *Regular weight aggregate (RWA) concrete modeling*

The mesostructure in Figure 8 is equipped with the material properties displayed in Table 1.

Table 1: Regular Weight Aggregate (RWA) concrete. Material properties at the mesoscopic level.

	Young modulus E	Poisson ratio ν	Ultimate stress σ_u	Fracture energy G_{μ}^{f}
Mortar (bulk)	20 Gpa	0.2	(elastic)	(elastic)
Mortar (cohesive bands)	20 Gpa	0.2	3.0 Mpa	700 J/m^2
Aggregates (bulk)	100 Gpa	0.2	(elastic)	(elastic)
Aggregates (cohesive bands)	100 Gpa	0.2	15.0 Mpa	3500 J/m^2

In Figure 10, qualitative results at the macroscopic level are presented.

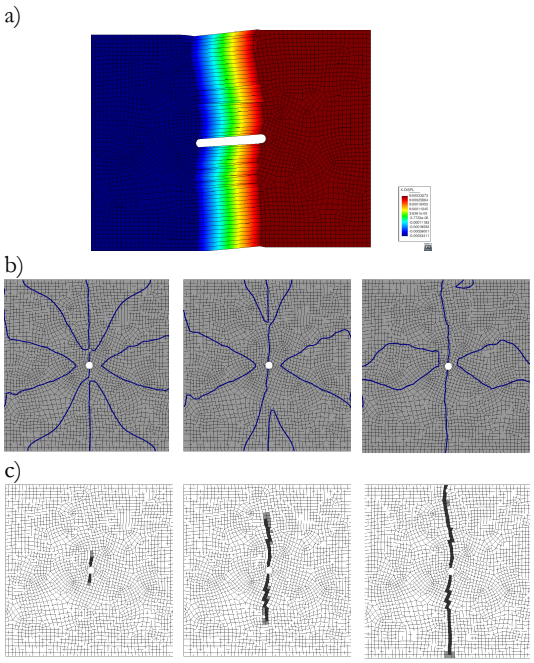


Figure 10 *Regular weight aggregate (RWA) concrete. Macroscopic results*. a) Deformed (amplified) shape showing the contours of horizontal displacements. b) Evolution, at sequential stages of the analysis, of the captured crack-path field. c) Evolution, at sequential stages of the analysis, of the strain injection pattern (darker zones correspond to strain-injected elements).

Those results show a macroscopic, approximately vertical, crack starting at the opening and exhibiting a mixed Mode I failure.

At the mesoscopic level, results for the deformation (fluctuations) of a typical failure-cell (nearby the opening in the macroscopic failure path) can be seen, conveniently amplified, in Figure 11. There it can be checked that, at an initial stage of the macroscopic localization, the failure-cell displays several (very small in amplitude) competing failure mechanisms (Figure 11-a)). Later on, an *extra-granular failure mechanism,* involving aggregate/matrix decohesion and mortar cracking, prevails (Figure 11-b)), characterizing the mesoscopic failure for this type of concrete.

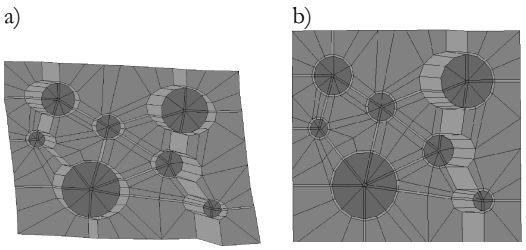


Figure 11. *Regular weight aggregate (RWA) concrete. Mesoscopic results*. Deformation of a typical failure cell displaying the failure modes: a) at the initiation of the macroscopic localization (amplification factor =1000), b) at the end of the loading process (amplification factor = 200).

In Figure 12 the evolution of the macroscopic energy along the process is presented. It can be noticed that the numerically observed dissipated energy computed as

$$E^{numerical} = \int_{t=0}^{t=\infty} \mathrm{F}(t) \cdot \dot{\Delta}(t) dt \qquad (19)$$

, where F and Δ are defined in Figure 7-a), see also Figure 16) matches very well the theoretical value computed as:

$$E^{theoretical} = L_{macro} G^{f}_{macro}$$

$$G^{f}_{macro} \equiv \overline{\left(G_{\mu}^{f}\right)}_{\mathcal{S}_{\mu}} = \frac{1}{\mathcal{S}_{\mu}} \int_{\mathcal{S}_{\mu}} G_{\mu}^{f}(\mathbf{y}) d\mathcal{S} \qquad (20)$$

i.e.: the length of the macroscopic crack, L_{macro}, times the macroscopic fracture energy G^{f}_{macro}. The theoretical value of G^{f}_{macro} is computed as the average, along the mesoscopic failure-cell failure path in Figure 11-b), of the mesoscopic fracture energy (G_{μ}^{f} in Table 1).

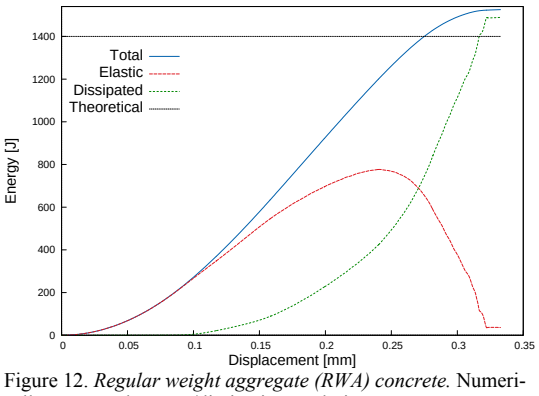


Figure 12. *Regular weight aggregate (RWA) concrete.* Numerically computed energy/dissipation evolution curves.

4.2 *Light weight aggregate (LWA) concrete modeling*

Now, light weight aggregate (LWA) concrete is considered. The material properties, for a ceramic-type aggregate are presented in Table 2.

Table 2: Light Weight Aggregate (LWA) concrete. Material properties at the mesoscopic level.

	Young modulus E	Poisson ratio ν	Ultimate stress σ_u	Fracture energy G_μ^f
Mortar (bulk)	28.6 Gpa	0.2	(elastic)	(elastic)
Mortar (cohesive bands)	28.6 Gpa	0.2	3.0 Mpa	700 J/m^2
Aggregates (bulk)	6.8 Gpa	0.2	(elastic)	(elastic)
Aggregates (cohesive bands)	6.8 Gpa	0.2	0.7 Mpa	166 J/m^2

Figure 13 displays a macroscopic failure mechanism fairly similar to the one in the previous (RWA) case.

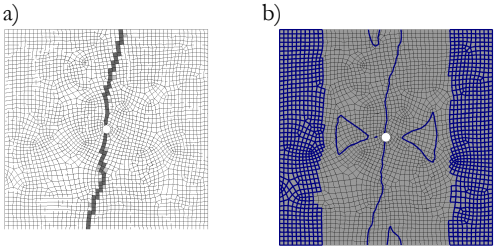


Figure 13 *Light-weight-aggregate (LWA) concrete. Macroscopic results*. a) strain injection pattern (darker zones correspond to strain-injected elements) at the end of the loading process, b) captured crack-path field.

However the mesoscopic failure mechanism is substantially different: the activated mechanism goes through the aggregates (see Figure 14-b), thus defining an intra-*granular failure mechanism*.

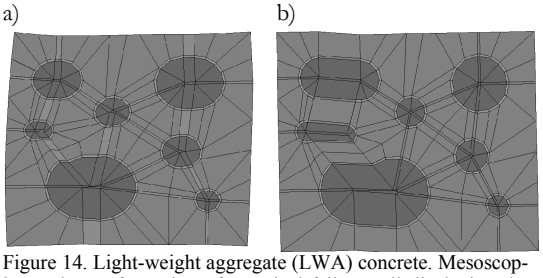


Figure 14. Light-weight aggregate (LWA) concrete. Mesoscopic results. Deformation of a typical failure cell displaying the failure modes: a) at the initiation of the macroscopic localization (amplification factor =1000), b) at the end of the loading process (amplification factor = 200).

As for the energy dissipation, in Figure 15 good agreement of the theoretical energy dissipation and the numerical one (computed according equation (19), see also Figure 16) is shown.

Notice that the energy dissipation of the process for this LWA concrete (850 J) is smaller than the one for NWA concrete in Figure 12 (1500 J). This is due to the fact that, in the present case, the mesoscopic failure mechanism, in Figure 14-b), crosses the (weaker) aggregates, which are endowed with a lower fracture energy (check G_μ^f in Table 2 versus Table 1). Therefore the macroscopic fracture energy, G_{macro}^f, in equation (20) is smaller .

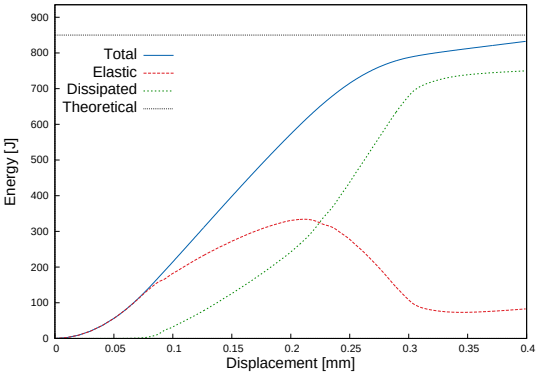


Figure 15. *Light-weight aggregate (LWA) concrete.* Numerically computed energy/dissipation evolution curves.

4.3 *Medium weight aggregate (MWA) concrete modeling*

As an intermediate case, the numerical experiment is repeated with the material properties presented in Table 3.

Table 3: Medium Weight Aggregate (MWA) concrete. Material properties at the mesoscopic level.

	Young modulus E	Poisson ratio ν	Ultimate stress σ_u	Fracture energy G_μ^f
Mortar (bulk)	28.6 Gpa	0.2	(elastic)	(elastic)
Mortar (cohesive bands)	28.6 Gpa	0.2	3.0 Mpa	700 J/m^2
Aggregates (bulk)	19.9 Gpa	0.2	(elastic)	(elastic)
Aggregates (cohesive bands)	19.9Gpa	0.2	2.1 Mpa	487 J/m^2

The macroscopic and mesoscopic failure mechanisms (not displayed) are similar to the ones, in Figure 10 and Figure 11, for RWA concrete. However, even if the qualitative effects of the different mesoscopic properties are not relevant, the quantitative structural response is very affected as it will be shown in next paragraph.

4.4 *Structural response*

Quantitative responses are now measured in terms of the force-displacement evolution curves, $F-\Delta$ in Figure 7-a), as shown in Figure 16.

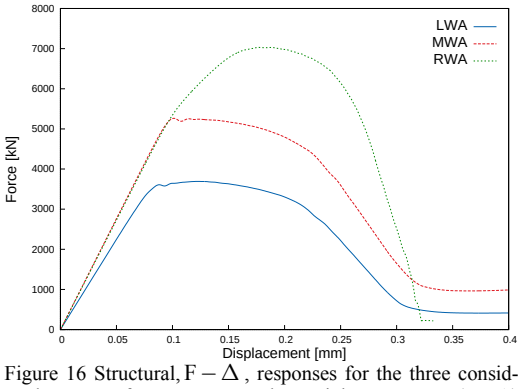


Figure 16 Structural, $F-\Delta$, responses for the three considered types of concrete. Regular weight aggregate (RWA) concrete, light weight aggregate (LWA) concrete and medium weight aggregate (MWA) concrete.

There it can be observed that:

- RWA concrete (regular concrete) behaves as observed in normal plain concrete. After an initial elastic branch, the structure exhibits some inelastic hardening and then fails abruptly (brittle failure).

- Introduction of light aggregate (MWA, LWA) concrete translates into substantial loss of the maximum carrying capacity but, also, in a more ductile failure: the non-linear hardening branch disappears and, immediately after the elastic branch, the structure softens in a smoother manner (ductile failure).

Both observations coincide with the expectations for this types of concrete.

5 CONCLUDING REMARKS

Along this work, multiscale (FE^2) techniques have been used to model material failure of concrete-like materials.

A computational homogenization setting has been proposed which, keeping the essentials of classical computational homogenization techniques, naturally introduces a length scale, related to the RVE size and the mesoscale failure mechanism, into the induced macroscopic homogenized model. This microscopic length scale can be interpreted as the actual width of the fracture process zone, defining the bandwidth of the macroscopic localization band that captures the crack at the structural level. It has been frequently claimed in modeling of concrete materials (Bazant and Jirásek, 2002) and it is automatically retrieved here through the homogenization process.

At the macroscopic scale, this microscopic length scale is locally used in the context of the Continuum Strong Discontinuity Approach to material failure (Oliver and Huespe, 2004a) , and introduced, as a regularization parameter, into a new technique (Oliver et al., 2014) for capturing the macroscopic propagation of cracks by means of finite elements with embedded discontinuities.

The result is a multiscale approach, that preserves the correct dissipation and objectivity, with respect both the macroscopic size of the finite element mesh and the size of the failure cell, which, in turn, can be readily connected with recent proposals for similar purposes (Sánchez et al., 2013).

Finally, the proposed techniques have been applied to modeling concrete like materials (i.e., materials that, at the mesoscopic level, are constituted by aggregates and mortar). Although the complexity of the mesoscale morphology is very limited by the computational cost of the combined multiscale computations, the potential of the approach is clearly stated by the presented examples.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 320815, Advanced Grant Project COMP-DES-MAT.

REFERENCES

Alfaiate, J., Simone, A. & Sluys, L. J. 2003. Non-homogeneous displacement jumps in strong embedded discontinuities. *International Journal of Solids and Structures,* 40**,** 5799-5817.

Armero, F. & Garikipati, K. 1996. An analysis of strong discontinuities in multiplicative finite strain plasticity and their relation with the numerical simulation of strain localization in solids. *Int.J. Solids and Structures,* 33**,** 2863-2885.

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

Bazant, Z. P. & Jirásek, M. 2002. Nonlocal Integral Formulations of Plasticity and Damage: Survey of Progress. *Journal of Engineering Mechanics,* 1119**,** 1125:1149.

De Souza Neto, E. A. & Feijóo, R. A. 2006. Variational foundations of multi-scale constitutive models of solid: small and large strain kinematical formulation. *LNCC Research & Development Report,* 16.

Huespe, A., Oliver, J. & Mora, D. 2013. Computational modeling of high performance steel fiber reinforced concrete using a micromorphic approach. *Computational Mechanics***,** 1-22.

Miehe, C., Welschinger, F. & Hofacker, M. 2010. Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field FE implementations. *Int. J. Numer. Meth. Engng.,* 83.

Nguyen, V. P., Lloberas-Valls, O., Stroeven, M. & Sluys, L. J. 2010. On the existence of representative volumes for softening quasi-brittle materials-a failure zone averaging scheme. *Computer Methods in Applied Mechanics and Engineering*.

Oliver, J. 1995. Continuum modelling of strong discontinuities in solid mechanics using damage models. *Computational Mechanics,* 17**,** 49-61.

Oliver, J., Cervera, M., Oller, S. & Lubliner, J. Isotropic damage models and smeared crack analysis of concrete. *In:* AL., N. B. E., ed. Proc. SCI-C Computer Aided Analysis and Design of Concrete Structures, 1990. 945-957.

Oliver, J. & Huespe, A. E. 2004a. Continuum approach to material failure in strong discontinuity settings. *Computer Methods in Applied Mechanics and Engineering,* 193**,** 3195-3220.

Oliver, J. & Huespe, A. E. 2004b. Theoretical and computational issues in modelling material failure in strong discontinuity scenarios. *Computer Methods in Applied Mechanics and Engineering,* 193**,** 2987-3014.

Oliver, J., Huespe, A. E. & Dias, I. F. 2014. Crack-path field and strain-injection techniques in computational modeling of propagating material failure. *Comp. Meth. in Appl. Mech. and Eng.,* Submitted.

Oliver, J., Huespe, A. E., Pulido, M. D. G. & Chaves, E. 2002. From continuum mechanics to fracture mechanics: the strong discontinuity approach. *Engineering Fracture Mechanics,* 69**,** 113-136.

Oliver, J., Mora, D., Huespe, A. & Weyler, R. 2012. A micromorphic model for steel fiber reinforced concrete. *International Journal of Solids and Structures,* 49**,** 2990-3007.

Özdemir, I., Brekelmans, W. & Geers, M. 2008. FE2 computational homogenization for the thermo-mechanical analysis of heterogeneous solids. *Computer Methods in Applied Mechanics and Engineering,* 198**,** 602-613.

Sánchez, P. J., Blanco, P. J., Huespe, A. E. & Feijóo, R. A. 2013. Failure-Oriented Multi-scale Variational Formulation: Micro-structures with nucleation and evolution of softening bands. *Comput. Methods Appl. Mech. Engrg.,* 257**,** 221–247.

Willam, K. & Sobh, N. 1987. Bifurcation analysis of tangential material operators. *In:* PANDE, G. N. & MIDDLETON, J. (eds.) *Transient/Dynamic Analysis and Constitutive Laws for Engineering Materials.* Martinus-Nijhoff Publishers.

This Thesis focuses on the numerical modeling of fracture and its propagation in heterogeneous materials by means of hierarchical multiscale models based on the FE2 method, addressing at the same time, the problem of the excessive computational cost through the development, implementation and validation of a set of computational tools based on reduced order modeling techniques.

For fracture problems, a novel multiscale model for propagating fracture has been developed, implemented and validated. This multiscale model is characterized by the following features:

At the macroscale level, were adapted the last advances of the Continuum Strong Discontinuity Approach (CSDA), devising a new finite element exhibiting good ability to capture and model strain localization in bands which can be intersect the finite element in random directions. At the microscale level, the use of cohesive-band elements, endowed with a regularized isotropic continuum damage model aiming at representing the material decohesion, is proposed. These cohesive-band elements are distributed within the microscale components, and their boundaries.

For model order reduction purposes, the microscale Boundary Value Problem (VBP), is rephrased using Model Order Reduction techniques. The use of two subsequent reduction techniques, known as: Reduced Order Model (ROM) and HyPer Reduced Order Model (HPROM or HROM), respectively, is proposed.

First, the standard microscale finite element model High Fidelity (HF), is projected and solved in a low-dimensional space via Proper Orthogonal Decomposition (POD). Second, two techniques have been developed and studied for multiscale models, namely: a) interpolation methods, and b) Reduced Order Cubature (ROQ) methods (An/2009). The reduced bases for the projection of the primal variables, are computed by means of a judiciously training, defining a set of pre-defined training trajectories.

Taking into account the discontinuous pattern of the strain field in problems exhibiting softening behavior. A domain separation strategy, is proposed. A cohesive domain, which contains the cohesive elements, and the regular domain, composed by the remaining set of finite elements. Each domain has an individual treatment. The microscale Boundary Value Problem (BVP) is rephrased as a saddle-point problem which minimizes the potential of free-energy, subjected to constraints fulfilling the basic hypotheses of multiscale models.

UNIVERSITAT POLITÈCNICA
DE CATALUNYA

ESCOLA TÈCNICA SUPERIOR D'ENGINYERS
DE CAMINS, CANALS I PORTS
DEPARTAMENT DE RESISTÈNCIA DE
MATERIALS I ESTRUCTURES A L'ENGINYERIA