Investigation of the interactions between the numerical and the modeling errors in the Homogenized Dirichlet Projection Method

Nicolas Moës, John Tinsley Oden, Tarek Zohdi

Investigation of the interactions between the numerical and the modeling errors in the Homogenized Dirichlet Projection Method

Nicolas Moës *, J. Tinsley Oden, Tarek I. Zohdi

The Texas Institute for Computational and Applied Mathematics, The University of Texas at Austin, Taylor Hall 2400, Austin, TX 78712, USA

The Homogenized Dirichlet Projection Method, HDPM, was developed in [1,2] as a systematic technique for analyzing highly-heterogeneous elastic structures. The method can provide an analysis of structures composed of composite materials with very complex microstructure at a fraction of the cost of solving the full fine-scale model. In the present investigation, the HDPM is revisited to take into account the unavoidable numerical errors produced in finite element approximations of the associated boundary value problems.

The total error of the HDPM, which now takes into account both for the modeling and numerical errors, is split into several terms, each accounting for a parameter in the method. The parameters are: the choice of the homogenized material property, the partition into subdomains, the coarse finite element mesh used to solve the homogenized problem and the fine meshes used to solve the subdomain problems. Numerical experiments are carried out on 1-D problems for which the exact solutions are easily calculated. The experiments reveal that the influence of the coarse and fine meshes are very different.

When free of numerical error the HDPM is based on four main results. We rewrite these in the framework of the error in the constitutive law. This leads to a clear mechanical interpretation of the results. Moreover, it allows us to extend the results to nonlinear constitutive laws and to find new properties of the HDPM. Some of the theoretical results are validated for specific cases involving known numerical error.

Finally, explicit a posteriori upper bounds are derived for the total and numerical error. A simple adaptive strategy is presented for choosing the fine mesh and the subdomain partition. The strategy is tested on a 1-D model problem.

1. Introduction

The Homogenized Dirichlet Projection Method (HDPM) [1,2] is an efficient method for solving large heterogeneous elastic problems. Basically, the method consists first in solving a homogenized problem with uniform material properties. Then, by computing an explicit error bound, the distance between the homogenized and exact solution is estimated. On subdomains where this distance is too large, a local analysis is performed using the homogenized solution as boundary conditions on the subdomains. This produces a perturbation of the homogenized solution which is closer to the exact solution. We refer as HDPM solution, the perturbation of the homogenized solution. Thus, the method introduces two types of problems: a regularized one with a uniform material property and a local analysis on subdomains

*Corresponding author.

with the real fine-scale material properties. In practice, these two problems cannot be solved exactly, therefore numerical errors are introduced.

This paper is the first experimental investigation on the interactions between the numerical and modeling errors in the HDPM. Some theoretical results were already presented in [2]. We consider that the homogenized solution is obtained through the classical finite element method using a coarse mesh (H-mesh), and that the local solutions, over the subdomains, are obtained with fine meshes (h-meshes). In this investigation, we neglect the error introduced by the use of iterative solvers and numerical integration of the stiffness matrices. In order to isolate the numerical influence, we split the total error into several terms. Numerical experiments are then carried out on 1-D problems, for which the exact solutions are easily calculated.

The HDPM is based on four results [1,2]:

- an explicit upper bound of the distance between the homogenized solution and the exact fine-scale solution in the energy norm;
- the fact that the HDPM solution is closer in the energy norm to the exact solution than the original homogenized solution;
- a local bound on the difference between the HDPM and the homogenized solutions;
- an upper bound on the difference between the HDPM and exact solution in the energy norm.

These four results hold when the method is free of numerical errors. An important question then is to determine what can be said in the case of numerical errors. This question is also addressed in the present paper.

The exact and homogenized solutions differ because they do not satisfy the same constitutive law. Thus, the quality of the homogenized solution may be estimated by the way it satisfies the exact constitutive law. This fact makes it possible to rewrite all the main results of the HDPM using the error-in-the-constitutive-law concept [3], thereby leading to a clear mechanical interpretation of these results. Moreover, written in this framework, the HDPM may be extended to problems involving nonlinear constitutive laws.

Since the modeling error may be expressed as an error in the constitutive law, we are naturally led to estimate the numerical error in the same way [3]. Computable a posteriori upper bounds are obtained for the total and numerical errors. The estimates are then used in a simple adaptive strategy to adapt the h-mesh size and the subdomain size.

The paper is organized as follows. In Section 2, an outline of the HDPM is given, which takes into account the numerical approximations. A decomposition of the total error is also presented. In Section 3, numerical experiments are carried out on a 1-D problem to analyze the influence of the numerical errors. In Section 4, the main results of the HDPM are restated in the error-in-the-constitutive-law framework. They are then studied in the presence of numerical errors in Section 5. Finally, a posteriori error estimation and a simple adaptive strategy for choosing the mesh for the local problems and the subdomain partition are proposed and tested on 1-D problems in Section 6.

2. HDPM with finite element approximation

2.1. The reference problem

We consider a material body composed of a linearly-elastic material in static equilibrium under the action of given body forces $\boldsymbol{f}_g$ and surface tractions $\boldsymbol{t}_g$. The domain, Ω, occupied by the material body is considered regular: a simply-connected domain with Lipschitz boundary $\partial\Omega$. The boundary $\partial\Omega$ consists of a portion Γ_u where displacements $\boldsymbol{u}_g$ are prescribed and a portion Γ_t where tractions $\boldsymbol{t}_g$ are prescribed,

$$\partial\Omega = \overline{\Gamma_u \cup \Gamma_t}, \qquad \Gamma_u \cap \Gamma_t = \emptyset.$$

The problem to be solved on Ω is to find a triple $(\boldsymbol{u}, \boldsymbol{\epsilon}, \boldsymbol{\sigma})$ such that the kinematic constraints (1), the equilibrium equation (2) and the constitutive law (3) hold:

$$\boldsymbol{\epsilon} = \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^t) \quad \text{on } \Omega \quad \text{and} \quad \boldsymbol{u} = \boldsymbol{u}_g \quad \text{on } \Gamma_u; \tag{1}$$

$$\nabla \cdot \boldsymbol{\sigma} = -\boldsymbol{f}_g \quad \text{on } \Omega \quad \text{and} \quad \boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{t}_g \quad \text{on } \Gamma_t; \tag{2}$$

$$\boldsymbol{\sigma} = \boldsymbol{E}\boldsymbol{\epsilon} \quad \text{on } \Omega. \tag{3}$$

Here $\boldsymbol{u}$, $\boldsymbol{\epsilon}$ and $\boldsymbol{\sigma}$ are the displacement, strain and stress, respectively and $\boldsymbol{n}$ is the outward normal to the boundary. The elasticity tensor $\boldsymbol{E}$ is a function of the position $\boldsymbol{x}$ i.e. $\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{x})$, and describes the microstructure of the material.

We remark that by eliminating $\boldsymbol{\epsilon}$ and $\boldsymbol{\sigma}$ in (1)–(3), and by taking into account the symmetries of $\boldsymbol{E}$, we have the classical elasticity problem for the displacement field,

$$-\nabla \cdot (\boldsymbol{E}\nabla \boldsymbol{u}) = \boldsymbol{f}_g \quad \text{on } \Omega, \qquad \boldsymbol{u} = \boldsymbol{u}_g \text{ on } \Gamma_u, \qquad \boldsymbol{E}\nabla \boldsymbol{u} \cdot \boldsymbol{n} = \boldsymbol{t}_g \quad \text{on } \Gamma_t. \tag{4}$$

A weak formulation of (4) is as follows:

$$\text{Find } \boldsymbol{u} \in V \text{ such that } \quad \mathcal{B}(\boldsymbol{u}, \boldsymbol{v}) = \mathcal{F}(\boldsymbol{v}) \ \ \forall \boldsymbol{v} \in V^o. \tag{5}$$

where

$$V = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v} = \boldsymbol{u}_g \text{ on } \Gamma_u\}, \qquad V^o = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v} = 0 \text{ on } \Gamma_u\}, \tag{6}$$

$$\mathcal{B}(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega (\boldsymbol{E}\nabla \boldsymbol{u}) : \nabla \boldsymbol{v} \, d\boldsymbol{x}, \qquad \mathcal{F}(\boldsymbol{v}) = \int_\Omega \boldsymbol{f}_g \cdot \boldsymbol{v} \, d\boldsymbol{x} + \int_{\Gamma_t} \boldsymbol{t}_g \cdot \boldsymbol{v} \, d\boldsymbol{s}. \tag{7}$$

2.2. The Homogenized Dirichlet Projection Method

Taking into account the unavoidable numerical errors, the HDPM can be summarized by four steps. The presentation which follows is valid for any space dimension but is illustrated for the 1-D problem, Fig. 1.

- Step 1: The so-called fine-scale problem, that is characterized by a highly heterogeneous material, Fig. 1, is replaced by an homogenized problem, Fig. 2, written:
  Find a triple $(\boldsymbol{u}^0, \boldsymbol{\epsilon}^0, \boldsymbol{\sigma}^0)$ such that

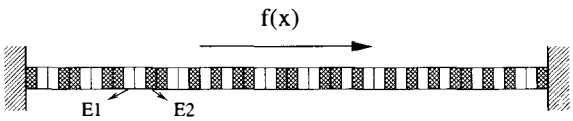
$$\boldsymbol{\epsilon}^0 = \frac{1}{2}(\nabla \boldsymbol{u}^0 + (\nabla \boldsymbol{u}^0)^t) \quad \text{on } \Omega \quad \text{and} \quad \boldsymbol{u}^0 = \boldsymbol{u}_g \quad \text{on } \Gamma_u; \tag{8}$$

$$\nabla \cdot \boldsymbol{\sigma}^0 = -\boldsymbol{f}_g \quad \text{on } \Omega \quad \text{and} \quad \boldsymbol{\sigma}^0 \cdot \boldsymbol{n} = \boldsymbol{t}_g \quad \text{on } \Gamma_t; \tag{9}$$

$$\boldsymbol{\sigma}^0 = \boldsymbol{E}^0 \boldsymbol{\epsilon}^0 \quad \text{on } \Omega. \tag{10}$$

  $\boldsymbol{E}^0$ is the homogenized elasticity tensor, which, for convenience, is assumed to be constant over the domain. We note that a weak form of (8)–(10) analogous to (5) can be obtained with $\mathcal{B}(\boldsymbol{u}, \boldsymbol{v})$ replaced by $\mathcal{B}_0(\boldsymbol{u}^0, \boldsymbol{v}) = \int_\Omega (\boldsymbol{E}^0 \nabla \boldsymbol{u}^0) : \nabla \boldsymbol{v} \, d\boldsymbol{x}$.
- Step 2: In general, the homogenized problem cannot be solved exactly, so numerical approximations are introduced. For example, using the finite element method on a mesh parameterized by a mesh size H (Fig. 3), the finite element solution is denoted by $(\boldsymbol{u}^{0,H}, \boldsymbol{\epsilon}^{0,H}, \boldsymbol{\sigma}^{0,H})$ and satisfies

$$\boldsymbol{\epsilon}^{0,H} = \frac{1}{2}(\nabla \boldsymbol{u}^{0,H} + (\nabla \boldsymbol{u}^{0,H})^t) \text{ on } \Omega \quad \text{and} \quad \boldsymbol{u}^{0,H} = \boldsymbol{u}_g \quad \text{on } \Gamma_u;$$



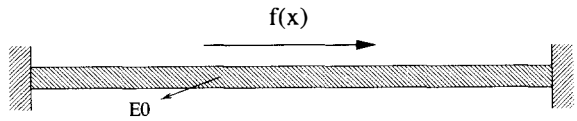


Fig. 1. The reference problem: geometry, loading and material property. The exact solution is denoted by $(\boldsymbol{u}, \boldsymbol{\epsilon}, \boldsymbol{\sigma})$.

Fig. 2. The homogenized problem whose exact solution is denoted by $(\boldsymbol{u}^0, \boldsymbol{\epsilon}^0, \boldsymbol{\sigma}^0)$.

$$\int_\Omega \boldsymbol{\sigma}^{0,H} : \nabla \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} - \int_\Omega \boldsymbol{f}_\mathrm{g} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} - \int_{\Gamma_\mathrm{t}} \boldsymbol{t}_\mathrm{g} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{s} = 0 \quad \forall \boldsymbol{v} \in V^{\mathrm{o},H}; \tag{11}$$

$$\boldsymbol{\sigma}^{0,H} = \boldsymbol{E}^0 \boldsymbol{\epsilon}^{0,H} \quad \text{on } \Omega \tag{12}$$

where $V^{\mathrm{o},H} \subset V^\mathrm{o}$ is an admissible finite element displacement space parameterized by the mesh size H. We assume that $V^{\mathrm{o},H}$ belongs to a family of subspaces constructed so that $\boldsymbol{u}^{0,H} \to \boldsymbol{u}^0 \in V$ as $H \to 0$. Also note that the finite element strain and stress are only well defined on the interior of each element. In other words, there are jumps in the tractions.

- Step 3: Once the homogenized problem is solved, the domain is partitioned into N subdomains Ω_k, $k = 1, 2, \ldots, N$,

$$\bigcup_{k=1}^{N} \overline{\Omega_k} = \overline{\Omega}, \qquad \Omega_i \cap \Omega_j = \emptyset, \quad i \neq j.$$

Fig. 4 shows a uniform partition of the domain in subdomains of size Δ. The subdomains are considered regular, simply connected, with Lipschitz boundary. On each subdomain it is possible to estimate the quality of the homogenized solution [1]. On subdomains where the quality of the homogenized solution is poor compared with the exact solution, a local analysis is performed. The local problem on the k^{th} subdomain, Ω_k, reads:
Find the triple $(\tilde{\boldsymbol{u}}_k^{0,H}, \tilde{\boldsymbol{\epsilon}}_k^{0,H}, \tilde{\boldsymbol{\sigma}}_k^{0,H})$ such that

$$\tilde{\boldsymbol{\epsilon}}_k^{0,H} = \frac{1}{2}(\nabla \tilde{\boldsymbol{u}}_k^{0,H} + (\nabla \tilde{\boldsymbol{u}}_k^{0,H})^t) \quad \text{on } \Omega_k \quad \text{and} \quad \tilde{\boldsymbol{u}}_k^{0,H} = \boldsymbol{u}^{0,H} \quad \text{on } \partial\Omega_k \backslash \partial\Omega, \tag{13}$$

$$\tilde{\boldsymbol{u}}_k^{0,H} = \boldsymbol{u}_\mathrm{g} \quad \text{on } \partial\Omega_k \cap \Gamma_\mathrm{u}; \tag{14}$$

$$\nabla \cdot \tilde{\boldsymbol{\sigma}}_k^{0,H} = -\boldsymbol{f}_\mathrm{g} \quad \text{on } \Omega_k \quad \text{and} \quad \tilde{\boldsymbol{\sigma}}_k^{0,H} \cdot \boldsymbol{n} = \boldsymbol{t}_\mathrm{g} \quad \text{on } \partial\Omega_k \cap \Gamma_\mathrm{t}; \tag{15}$$

$$\tilde{\boldsymbol{\sigma}}_k^{0,H} = \boldsymbol{E}\, \tilde{\boldsymbol{\epsilon}}_k^{0,H} \quad \text{on } \Omega_k \tag{16}$$

where $\partial\Omega_k$ denotes the boundary of Ω_k. The homogenized displacements, $\boldsymbol{u}^{0,H}$, are imposed on the boundary of the subdomain, except on the boundary, $\partial\Omega$, where the actual boundary conditions are satisfied. Note also that the local analysis is solved with the actual microstructure described by $\boldsymbol{E}$.

For the sake of simplicity, we assume that the local analysis is performed on every subdomain, but this is seldom the case in application of the HDPM. The final solution $(\tilde{\boldsymbol{u}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H}, \tilde{\boldsymbol{\sigma}}^{0,H})$ over the domain is then constructed in the following manner

$$\tilde{\boldsymbol{u}}^{0,H} = \boldsymbol{u}^{0,H} + \sum_{k=1}^{N} \mathcal{E}_k (\tilde{\boldsymbol{u}}_k^{0,H} - \boldsymbol{u}^{0,H});$$

$$\tilde{\boldsymbol{\epsilon}}^{0,H} = \boldsymbol{\epsilon}^{0,H} + \sum_{k=1}^{N} \mathcal{E}_k (\tilde{\boldsymbol{\epsilon}}_k^{0,H} - \boldsymbol{\epsilon}^{0,H});$$

$$\tilde{\boldsymbol{\sigma}}^{0,H} = \boldsymbol{\sigma}^{0,H} + \sum_{k=1}^{N} \mathcal{E}_k (\tilde{\boldsymbol{\sigma}}_k^{0,H} - \boldsymbol{\sigma}^{0,H}).$$

where $\mathcal{E}_k$ is a scalar function defined on Ω with a value of unity on Ω_k and zero elsewhere.

Let us define Γ_int as the union of the boundaries of all the subdomain modulo the boundary of the domain

$$\Gamma_\mathrm{int} = \left(\bigcup_{k=1}^{N} \partial\Omega_k \right) \backslash \partial\Omega.$$

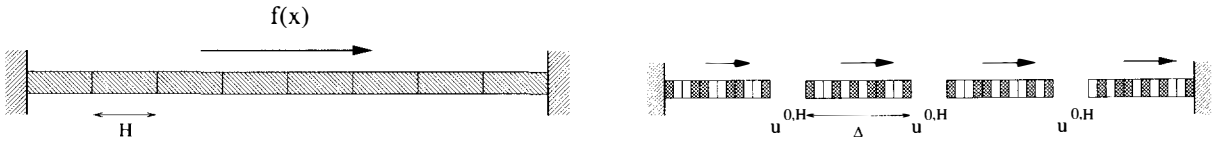


Fig. 3. The finite element problem associated with the homogenized problem. The mesh size is H and the finite element solution is $(\boldsymbol{u}^{0,H}, \boldsymbol{\epsilon}^{0,H}, \boldsymbol{\sigma}^{0,H})$.

Fig. 4. The subdomain problems. The subdomain size is Δ. The finite element displacements $\boldsymbol{u}^{0,H}$ are applied at the boundaries of the subdomains, in the interior of the domain. The exact solution of the subdomain problems is denoted by $(\tilde{\boldsymbol{u}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H}, \tilde{\boldsymbol{\sigma}}^{0,H})$.

By construction, the displacement field $\tilde{\boldsymbol{u}}^{0,H}$ is continuous across Γ_{int}. The normal stresses $\tilde{\boldsymbol{\sigma}}^{0,H}\boldsymbol{n}$ are not in general continuous across Γ_{int}. Let Ω_1 and Ω_2 be two subdomains having a common boundary. The traction jump $[\tilde{\boldsymbol{\sigma}}^{0,H} \cdot \boldsymbol{n}]$ on this boundary is defined by

$$[\tilde{\boldsymbol{\sigma}}^{0,H} \cdot \boldsymbol{n}] = \tilde{\boldsymbol{\sigma}}_1^{0,H} \cdot \boldsymbol{n}_1 + \tilde{\boldsymbol{\sigma}}_2^{0,H} \cdot \boldsymbol{n}_2$$

where $\boldsymbol{n}_i$ is the outward normal to the boundary for the subdomain i and $\tilde{\boldsymbol{\sigma}}_i^{0,H}$ is the stress state in the subdomain i on the boundary $(i = 1,2)$.

Finally, if we assume that no numerical error was introduced in Step 2, the solution of the Step 3 just described will be denoted $(\tilde{\boldsymbol{u}}_k^0, \tilde{\boldsymbol{\epsilon}}_k^0, \tilde{\boldsymbol{\sigma}}_k^0)$ for each subdomain and $(\tilde{\boldsymbol{u}}^0, \tilde{\boldsymbol{\epsilon}}^0, \tilde{\boldsymbol{\sigma}}^0)$ on the whole domain. The solution $(\tilde{\boldsymbol{u}}_k^0, \tilde{\boldsymbol{\epsilon}}_k^0, \tilde{\boldsymbol{\sigma}}_k^0)$ satisfies problem (13)–(16) with all the H superscripts removed.

- Step 4: In fact, as in the case of the homogenized problem, the subdomain problems cannot be solved exactly, introducing further error components. On each subdomain, Ω_k, the exact solution, $(\tilde{\boldsymbol{u}}_k^{0,H}, \tilde{\boldsymbol{\epsilon}}_k^{0,H}, \tilde{\boldsymbol{\sigma}}_k^{0,H})$, of the problem (13)–(16) is approximated by a finite element solution parameterized by a mesh size h (Fig. 5). This approximate solution is denoted by $(\tilde{\boldsymbol{u}}_k^{0,H,h}, \tilde{\boldsymbol{\epsilon}}_k^{0,H,h}, \tilde{\boldsymbol{\sigma}}_k^{0,H,h})$ and satisfies

$$\tilde{\boldsymbol{\epsilon}}_k^{0,H,h} = \frac{1}{2}(\nabla \tilde{\boldsymbol{u}}_k^{0,H,h} + (\nabla \tilde{\boldsymbol{u}}_k^{0,H,h})^t) \quad \text{on } \Omega_k, \qquad \tilde{\boldsymbol{u}}_k^{0,H,h} = \boldsymbol{u}^{0,H} \quad \text{on } \partial\Omega_k \backslash \partial\Omega; \tag{17}$$

$$\tilde{\boldsymbol{u}}_k^{0,H,h} = \boldsymbol{u}_{\text{g}} \quad \text{on } \partial\Omega_k \cap \Gamma_{\text{u}}; \tag{18}$$

$$\int_{\Omega_k} \tilde{\boldsymbol{\sigma}}_k^{0,H,h} : \nabla \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} - \int_{\Omega_k} \boldsymbol{f}_{\text{g}} \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} - \int_{\Gamma_{\text{t}} \cap \partial\Omega_k} \boldsymbol{t}_{\text{g}} \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{s} = 0 \quad \forall \boldsymbol{v} \in V_k^{\text{o},h}; \tag{19}$$

$$\tilde{\boldsymbol{\sigma}}_k^{0,H,h} = \boldsymbol{E}\, \tilde{\boldsymbol{\epsilon}}_k^{0,H,h} \quad \text{on } \Omega_k \tag{20}$$

where

$$V_k^{\text{o},h} = \{\boldsymbol{v} \in V_k^h : \boldsymbol{v} = 0 \text{ on } \partial\Omega_k \backslash \Gamma_{\text{t}}\}.$$

V_k^h is a finite element displacement space parameterized by the mesh size h on subdomain Ω_k.

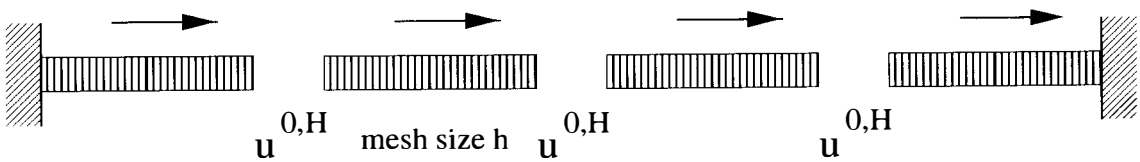


Fig. 5. The finite element problems associated to the subdomain problems. The mesh size is h and the finite element solution is denoted by $(\tilde{\boldsymbol{u}}^{0,H,h}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}, \tilde{\boldsymbol{\sigma}}^{0,H,h})$.

The final solution of the HDPM is denoted by $(\tilde{\boldsymbol{u}}^{0,H,h}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}, \tilde{\boldsymbol{\sigma}}^{0,H,h})$ and is defined in the following manner:

$$\tilde{\boldsymbol{u}}^{0,H,h} = \boldsymbol{u}^{0,H} + \sum_{k=1}^{N} \mathcal{E}_k(\tilde{\boldsymbol{u}}_k^{0,H,h} - \boldsymbol{u}^{0,H});$$

$$\tilde{\boldsymbol{\epsilon}}^{0,H,h} = \boldsymbol{\epsilon}^{0,H} + \sum_{k=1}^{N} \mathcal{E}_k(\tilde{\boldsymbol{\epsilon}}_k^{0,H,h} - \boldsymbol{\epsilon}^{0,H});$$

$$\tilde{\boldsymbol{\sigma}}^{0,H,h} = \boldsymbol{\sigma}^{0,H} + \sum_{k=1}^{N} \mathcal{E}_k(\tilde{\boldsymbol{\sigma}}_k^{0,H,h} - \boldsymbol{\sigma}^{0,H}).$$

As before, we assume that $V_k^h \subset V_k$ are members of appropriate families of subspace constructed so that $\tilde{\boldsymbol{u}}_k^{0,H,h} \to \tilde{\boldsymbol{u}}_k^{0,H}$ in V_k as $h \to 0$, where

$$V_k = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega_k) : \boldsymbol{v} = \boldsymbol{u}^{0,H} \text{ on } \partial\Omega_k \backslash \partial\Omega \text{ and } \boldsymbol{v} = \boldsymbol{u}_g \text{ on } \partial\Omega_k \cap \Gamma_u\}.$$

2.3. *Decomposition of the error*

Let us summarize the various solutions and their finite element approximations:

- $(\boldsymbol{u}, \boldsymbol{\epsilon} = \boldsymbol{\epsilon}(\boldsymbol{u}), \boldsymbol{\sigma} = \boldsymbol{E}\boldsymbol{\epsilon})$ = the exact solution, also called fine-scale solution;
- $(\boldsymbol{u}^0, \boldsymbol{\epsilon}^0 = \boldsymbol{\epsilon}(\boldsymbol{u}^0), \boldsymbol{\sigma}^0 = \boldsymbol{E}^0\boldsymbol{\epsilon}^0)$ = the homogenized solution;
- $(\boldsymbol{u}^{0,H}, \boldsymbol{\epsilon}^{0,H} = \boldsymbol{\epsilon}(\boldsymbol{u}^{0,H}), \boldsymbol{\sigma}^{0,H} = \boldsymbol{E}^0\boldsymbol{\epsilon}^{0,H})$ = the coarse-mesh finite element approximation of the homogenized solution;
- $(\tilde{\boldsymbol{u}}^0, \tilde{\boldsymbol{\epsilon}}^0 = \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}^0), \tilde{\boldsymbol{\sigma}}^0 = \boldsymbol{E}\tilde{\boldsymbol{\epsilon}}^0)$ = the perturbed solution defined on N subdomains on which the exact $\boldsymbol{u}^0$ is prescribed as boundary data, on the interior;
- $(\tilde{\boldsymbol{u}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H} = \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}^{0,H}), \tilde{\boldsymbol{\sigma}}^{0,H} = \boldsymbol{E}\tilde{\boldsymbol{\epsilon}}^{0,H})$ = the perturbed solution defined on the N subdomains on which the approximate homogenized solution $\boldsymbol{u}^{0,H}$ is prescribed as boundary data, on the interior;
- $(\tilde{\boldsymbol{u}}^{0,H,h}, \tilde{\boldsymbol{\epsilon}}^{0,H,h} = \boldsymbol{\epsilon}(\tilde{\boldsymbol{u}}^{0,H,h}), \tilde{\boldsymbol{\sigma}}^{0,H,h} = \boldsymbol{E}\tilde{\boldsymbol{\epsilon}}^{0,H,h})$ = the fine-mesh approximation of $(\tilde{\boldsymbol{u}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H}, \tilde{\boldsymbol{\sigma}}^{0,H})$.

$(\tilde{\boldsymbol{u}}^{0,H,h}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}, \tilde{\boldsymbol{\sigma}}^{0,H,h})$ is referred as the HDPM solution. The error in the energy norm associated to this solution reads:

$$e = \|\boldsymbol{u} - \tilde{\boldsymbol{u}}^{0,H,h}\|_{E(\Omega)}, \qquad \text{where } \|\boldsymbol{v}\|_{E(\Omega)}^2 = \int_\Omega \boldsymbol{E}\nabla\boldsymbol{v} : \nabla\boldsymbol{v}\, \mathrm{d}\boldsymbol{x}. \tag{21}$$

This error depends on four factors. Two are modeling-related: the homogenized material tensor $\boldsymbol{E}^0$ and the partition in subdomains; and two are numerically-related: the H-mesh to solve the homogenized problem and the h-meshes to solve the subdomain problems.

We first separate the error into a modeling and a numerical part:

PROPERTY 1. If $\tilde{\boldsymbol{u}}^{0,H,h} = \boldsymbol{u}^{0,H}$ on Γ_{int} then

$$e^2 = \underbrace{\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)}^2}_{\text{modeling}} + \underbrace{\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H,h}\|_{E(\Omega)}^2}_{\text{numerical}} + \underbrace{2\int_{\Gamma_{\text{int}}} [\tilde{\boldsymbol{\sigma}}^0 \cdot \boldsymbol{n}] \cdot (\boldsymbol{u}^{0,H} - \boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{s}}_{\text{coupling}}. \tag{22}$$

PROOF. Owing to the definition of the norm $\|\cdot\|_{E(\Omega)}$ in (21), we have

$$e^2 = \|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)}^2 + \|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H,h}\|_{E(\Omega)}^2 + 2\int_\Omega \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^0)\, \mathrm{d}\boldsymbol{x}.$$

The third term of the right hand side can be transformed as follows:

$$\int_\Omega \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^0)\, \mathrm{d}\boldsymbol{x} = \int_\Omega (\tilde{\boldsymbol{\sigma}}^0 - \boldsymbol{\sigma}) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^0)\, \mathrm{d}\boldsymbol{x}$$

$$= -\underbrace{\int_{\Omega} (\nabla \cdot \tilde{\boldsymbol{\sigma}}^0 - \nabla \cdot \boldsymbol{\sigma}).(\tilde{\boldsymbol{u}}^{0,H,h} - \tilde{\boldsymbol{u}}^0)\, \mathrm{d}\boldsymbol{x}}_{0}$$

$$+ \int_{\Gamma_{\text{int}}} [(\tilde{\boldsymbol{\sigma}}^0 - \boldsymbol{\sigma}) \cdot \boldsymbol{n}] \cdot (\tilde{\boldsymbol{u}}^{0,H,h} - \tilde{\boldsymbol{u}}^0)\, \mathrm{d}\boldsymbol{s}$$

$$= \int_{\Gamma_{\text{int}}} [\tilde{\boldsymbol{\sigma}}^0 \cdot \boldsymbol{n}] \cdot (\boldsymbol{u}^{0,H} - \boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{s}.$$

We have used the fact that $\boldsymbol{u}^0 = \tilde{\boldsymbol{u}}^0$ on Γ_{int}. □

The coupling term can be positive or negative. It represents the work done by the jump in traction $[\tilde{\boldsymbol{\sigma}}^0 \cdot \boldsymbol{n}]$ moving through the difference in the numerical and exact homogenized displacement. Thus, it depends both on the modeling and numerical errors. On the contrary, if we isolate the numerical error coming from the local analysis, we get the following decomposition:

PROPERTY 2. If $\tilde{\boldsymbol{u}}^{0,H,h} = \tilde{\boldsymbol{u}}^{0,H}$ on Γ_{int}, then

$$e^2 = \underbrace{\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)}}_{\text{mod.+num. } (H\text{-mesh})} + \underbrace{\|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)}}_{\text{num. } (h\text{-mesh})}. \tag{23}$$

PROOF. We have

$$e^2 = \|\boldsymbol{u} - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)} + \|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)} + 2\int_{\Omega} \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^{0,H} - \nabla\boldsymbol{u}) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x}$$

and

$$\int_{\Omega} \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^{0,H} - \nabla\boldsymbol{u}) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x} = \int_{\Omega} (\tilde{\boldsymbol{\sigma}}^{0,H} - \boldsymbol{\sigma}) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x}$$

$$= -\underbrace{\int_{\Omega} (\nabla \cdot \tilde{\boldsymbol{\sigma}}^{0,H} - \nabla \cdot \boldsymbol{\sigma}) \cdot (\tilde{\boldsymbol{u}}^{0,H,h} - \tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x}}_{0}$$

$$+ \int_{\Gamma_{\text{int}}} [(\tilde{\boldsymbol{\sigma}}^{0,H} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}] \cdot (\tilde{\boldsymbol{u}}^{0,H,h} - \tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{s}$$

$$= 0. \quad \square$$

Let us now decompose the modeling and numerical errors appearing in Property 1. The modeling error can be decomposed into the error due to the homogenization process minus that gained by doing the local analysis (Property 3). The numerical error can be decomposed into the error due to the H-mesh approximation and the error due to the h-mesh approximation (Property 4).

PROPERTY 3.

$$\underbrace{\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|^2_{E(\Omega)}}_{\text{modeling error}} = \underbrace{\|\boldsymbol{u} - \boldsymbol{u}^0\|^2_{E(\Omega)}}_{\text{homogenization}} - \underbrace{\|\boldsymbol{u}^0 - \tilde{\boldsymbol{u}}^0\|^2_{E(\Omega)}}_{\text{sudom. analysis}}. \tag{24}$$

PROOF. We have

$$\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|^2_{E(\Omega)} = \|\boldsymbol{u} - \boldsymbol{u}^0\|^2_{E(\Omega)} - \|\boldsymbol{u}^0 - \tilde{\boldsymbol{u}}^0\|^2_{E(\Omega)} + 2\int_{\Omega} \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}) : (\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{x}$$

and

$$\int_{\Omega} \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}) : (\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{x} = \int_{\Omega} (\tilde{\boldsymbol{\sigma}}^0 - \boldsymbol{\sigma}) : (\nabla\tilde{\boldsymbol{u}}^0 - \nabla\boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{x}$$

$$= \int_{\Gamma_{\text{int}}} [(\tilde{\boldsymbol{\sigma}}^0 - \boldsymbol{\sigma}) \cdot \boldsymbol{n}] \cdot (\tilde{\boldsymbol{u}}^0 - \boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{s}$$
$$= 0$$

owing to the fact that $\boldsymbol{u}^0 = \tilde{\boldsymbol{u}}^0$ on Γ_{int}. □

PROPERTY 4. If $\tilde{\boldsymbol{u}}^{0,H,h} = \tilde{\boldsymbol{u}}^{0,H}$ on Γ_{int} then

$$\underbrace{\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)}}_{\text{numerical}} = \underbrace{\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)}}_{\text{num. } (H\text{-mesh})} + \underbrace{\|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)}}_{\text{num. } (h\text{-mesh})}. \tag{25}$$

PROOF. We have

$$\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)} = \|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)} + \|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)}$$
$$+ 2\int_\Omega \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^{0,H} - \nabla\tilde{\boldsymbol{u}}^0) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x}$$

and

$$\int_\Omega \boldsymbol{E}(\nabla\tilde{\boldsymbol{u}}^{0,H} - \nabla\tilde{\boldsymbol{u}}^0) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x} = \int_\Omega (\tilde{\boldsymbol{\sigma}}^{0,H} - \tilde{\boldsymbol{\sigma}}^0) : (\nabla\tilde{\boldsymbol{u}}^{0,H,h} - \nabla\tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{x}$$
$$= \int_{\Gamma_{\text{int}}} [(\tilde{\boldsymbol{\sigma}}^{0,H} - \tilde{\boldsymbol{\sigma}}^0) \cdot \boldsymbol{n}] \cdot (\tilde{\boldsymbol{u}}^{0,H,h} - \tilde{\boldsymbol{u}}^{0,H})\, \mathrm{d}\boldsymbol{s}$$
$$= 0.$$

Gathering Properties 1, 3 and 4, we finally get the following decomposition for the error:

PROPERTY 5. If $\tilde{\boldsymbol{u}}^{0,H,h} = \tilde{\boldsymbol{u}}^{0,H} = \boldsymbol{u}^{0,H}$ on Γ_{int}, then

$$e^2 = \|\boldsymbol{u} - \boldsymbol{u}^0\|^2_{E(\Omega)} - \|\boldsymbol{u}^0 - \tilde{\boldsymbol{u}}^0\|^2_{E(\Omega)} + \|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)} + \|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)}$$
$$+ 2\int_{\Gamma_{\text{int}}} [\tilde{\boldsymbol{\sigma}}^0 \cdot \boldsymbol{n}] \cdot (\boldsymbol{u}^{0,H} - \boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{s}. \quad \square \tag{26}$$

Let us establish precisely the meaning of each of these terms:

- The first term, $\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)}$, is the error introduced by replacing $\boldsymbol{E}$ by $\boldsymbol{E}^0$ in the original problem. It depends on $\boldsymbol{E}^0$ and it is zero if $\boldsymbol{E}^0$ and $\boldsymbol{E}$ coincide.
- The second term, $\|\boldsymbol{u}^0 - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)}$, is what we gain (minus sign in (26)) in solving the subdomain problems. It depends on $\boldsymbol{E}^0$ and the subdomain partition, symbolically denoted by Δ. As the subdomain size tends to zero, $\Delta \to 0$ symbolically, this term tends to zero since $\tilde{\boldsymbol{u}}^0 \to \boldsymbol{u}^0$ and it is zero if $\boldsymbol{E}^0$ and $\boldsymbol{E}$ coincide. Finally note that the two first terms in the right hand side of (26) form a quantity that is always greater or equal to zero, by Property 3.
- The third term, $\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H}\|_{E(\Omega)}$, takes into account the numerical errors introduced in solving the homogenized problems. It depends on $\boldsymbol{E}^0$, Δ and the H-mesh size. It tends to zero as $H \to 0$ since $\boldsymbol{u}^{0,H} \to \boldsymbol{u}^0$. But is *not* zero in general if $\boldsymbol{E}^0$ and $\boldsymbol{E}$ coincide. We have

  $$\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)} = \int_{\Gamma_{\text{int}}} [(\tilde{\boldsymbol{\sigma}}^{0,H} - \tilde{\boldsymbol{\sigma}}^0) \cdot \boldsymbol{n}] \cdot (\boldsymbol{u}^{0,H} - \boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{s}.$$

  Thus, the third term depends on the quality of the finite element solution $\boldsymbol{u}^{0,H}$ only on the boundary of the subdomains, Γ_{int}.
- The fourth term, $\|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|_{E(\Omega)}$, takes into account the numerical errors introduced by solving approximately the subdomain problems. It tends to zero as h tends to zero since $\tilde{\boldsymbol{u}}^{0,H,h} \to \tilde{\boldsymbol{u}}^{0,H}$ and it is not zero, in general, if $\boldsymbol{E}^0$ and $\boldsymbol{E}$ coincide.

- The last term in (26) is the coupling term whose sign can be positive or negative. It tends to zero as $H \to 0$ and it is zero if $\boldsymbol{E}^0$ and $\boldsymbol{E}$ coincide. Thus, this term exhibits a coupling between the numerical and modeling errors.

3. Influence of the numerical errors

We study here the influence of the h- and H-mesh on the error in connection with a model problem. The influence of the subdomain size is also studied since it has many features similar to the influence of the h-mesh. The study is carried out on the 1-D problem described in Fig. 1. The problem to be solved on $(0,1)$ is to find a triple (u, ϵ, σ) such that

$$\epsilon(x) = \frac{\mathrm{d}u(x)}{\mathrm{d}x}, \quad x \in (0,1) \quad \text{and} \quad u(0) = 0, \quad u(1) = 0; \tag{27}$$

$$\frac{\mathrm{d}\sigma(x)}{\mathrm{d}x} = -f(x), \quad f(x) = 1, \quad x \in (0,1); \tag{28}$$

$$\sigma(x) = E\epsilon(x), \quad x \in (0,1). \tag{29}$$

The Young modulus $E = E(x)$ corresponds to a two-phase material with equal volumetric distribution of each phase. More precisely, we consider that the rod is composed of 'particles' of size $d = 1/P$, P being the total number of particles. These particles have a Young modulus of E_1 or E_2. The partition of the phases on the rod is generated randomly.

3.1. Influence of the h-mesh

The h-mesh size influences only the fourth term in the decomposition of the error (26). We introduce the notations

$$\epsilon_{\mathrm{num},h} = \frac{\|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}}, \qquad \epsilon = \frac{\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^{0,H,h}\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}}$$

and we first consider the case of one subdomain. So, $\epsilon = \epsilon_{\mathrm{num},h} = \epsilon_{\mathrm{num},h}(h)$. We thus study the behavior of the finite element method for a heterogeneous material. Piecewise linear finite elements are considered.

3.1.1. Non-matching mesh

Fig. 6 shows the behavior of $\epsilon_{\mathrm{num},h}$ for uniform meshes non matching the particles boundaries (the number of particles is a power of 10 and the number of elements a power of 2). We observe that the error is almost constant during a range of mesh sizes and then starts to decrease with a rate of convergence $O(\sqrt{h})$. The ratio of the size of the element and the size of the particles is about three or four when the error starts to decrease.

The almost constant value of the error before the decrease does not depend on the size of particles nor on the size of the elements. It only depends on τ, the 'mismatch ratio' E_1/E_2. Fig. 7 shows two

Table 1
Evolution of the error (%) with the number of elements for several values of the mismatch τ. The number of particles is 10^4

		Number of elements				
τ	$\frac{\lvert 1-\tau\rvert}{(1+\tau)} * 10^2$	1	4	16	64	256
100	98.0	100	98.2	98.0	98.0	98.0
10	81.8	100	83.1	81.9	81.7	81.4
5	66.7	100	69.1	66.8	66.6	66.2
2	33.3	100	41.2	33.8	33.3	33.0

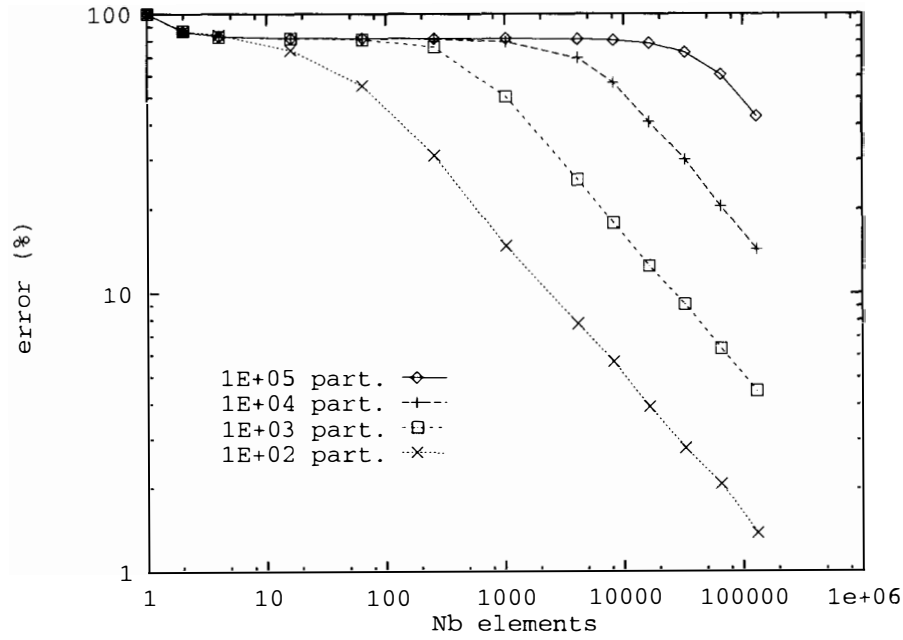


Fig. 6. Evolution of the numerical error $\epsilon_{\text{num},h}$ (%) with the number of elements in the h-mesh for several numbers of particles. The mismatch is $\tau = 10$.

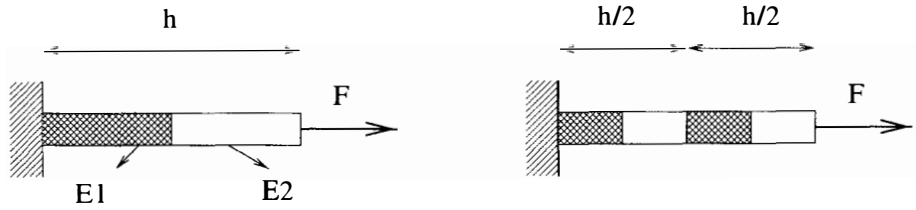


Fig. 7. Two finite element problems having the same numerical error.

finite element problems for which the numerical errors are identically the same. The relative error is $C(\tau) = |1 - \tau|/(1 + \tau)$. Although our 1-D problem is different from the one described in Fig. 7, the value $C(\tau)$ is very close to the actual value for our problem (Table 1). Note that when $\tau \to +\infty$, the error stays at a level of 100% before a significant decrease.

The error exhibits three main trends (Fig. 8). First, $O(h)$ convergence is observed until an error $|1 - \tau|/(1 + \tau)$ is reached (when $\tau = 1$ we observe an order of convergence of unity all the way as $h \to 0$). Then, the error decreases very slowly. Finally, when h is around three times the size of the particle we observe a rate of convergence of $O(\sqrt{h})$. This rate is in agreement with the classical convergence results of the finite element method.

The following empirical formula gives an idea of the number of elements needed per particle, for this problem, in order to achieve a numerical error ϵ_0

$$\frac{d}{h} \simeq \frac{1}{r_{\text{crit}}} \left(\frac{|1 - \tau|}{(1 + \tau)} \right)^2 \frac{1}{\epsilon_0^2}.$$

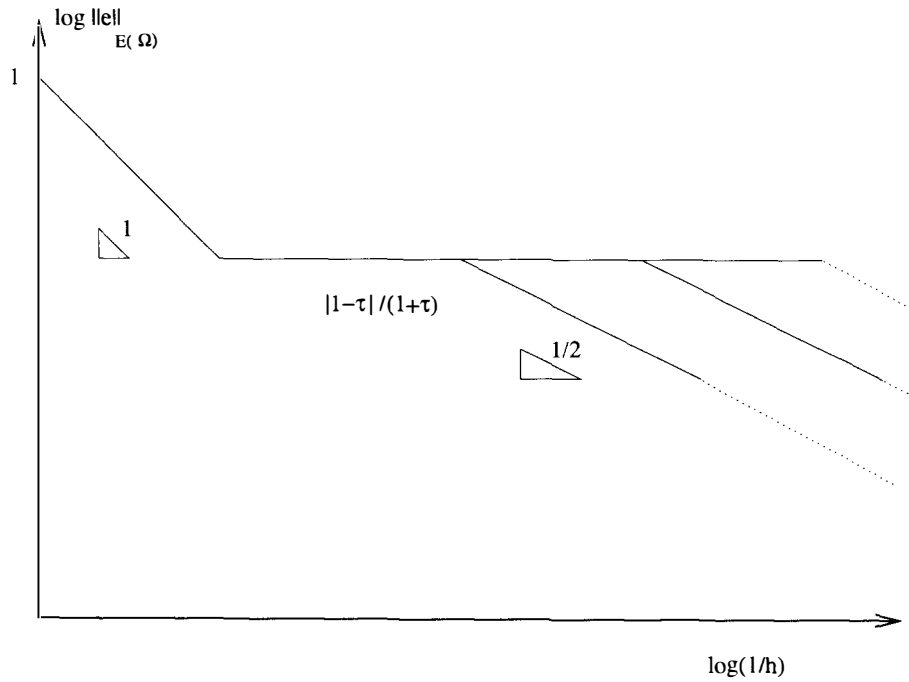


Fig. 8. The three main behavior phases of the numerical error $\epsilon_{\text{num},h}$ with non-matching meshes.

d is the size of the particle and r_{crit} is the critical ratio h/d needed to obtain a significant decrease of the error ($r_{\text{crit}} \sim 3$). With $\tau = 10$, about 5 elements per particle are needed for 20% error, about 20 for 10% and about 85 for 5%. Due to the poor convergence rates the numerical effort needed is very large.

Concerning the influence of the subdomain size Δ, we observe that $\epsilon_{\text{num},h}$ increases slightly for a given h-mesh when Δ is decreased, if $\boldsymbol{E}^0 = \langle \boldsymbol{E}^{-1} \rangle^{-1}$. The opposite behavior is observed if $\boldsymbol{E}^0 = \langle \boldsymbol{E} \rangle$. However, in general, whatever the value of Δ and the choice of $\boldsymbol{E}^0$, the same general behavior as depicted in Fig. 9 may be observed. The only difference is that the limiting value $|1 - \tau|/(1 + \tau)$ is no longer valid.

3.1.2. Matching mesh

We now consider the case where the mesh does take into account the boundary of the particles, i.e. a node is placed at each boundary between two particles. Fig. 9 shows the evolution of the error with the number of elements compared with the previous case where the mesh was not taking into account the boundaries of the particles. The error is now dramatically smaller and does not depend on the number of particles. The error is simply given by the mesh size h.

3.2. Influence of the H-mesh

We denote this influence by δ_H. It gathers the third and last term of the decomposition (26)

$$\delta_H = \epsilon^2_{\text{num},H} + \gamma_H \tag{30}$$

where

$$\epsilon_{\text{num},H} = \frac{\|\tilde{\boldsymbol{u}}^0 - \tilde{\boldsymbol{u}}^{0,H}\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}}, \qquad \gamma_H = \frac{2}{\|\boldsymbol{u}\|^2_{E(\Omega)}} \int_{\Gamma_{\text{int}}} [\tilde{\boldsymbol{\sigma}}^0 \cdot \boldsymbol{n}] \cdot (\boldsymbol{u}^{0,H} - \boldsymbol{u}^0)\, \mathrm{d}\boldsymbol{s}. \tag{31}$$

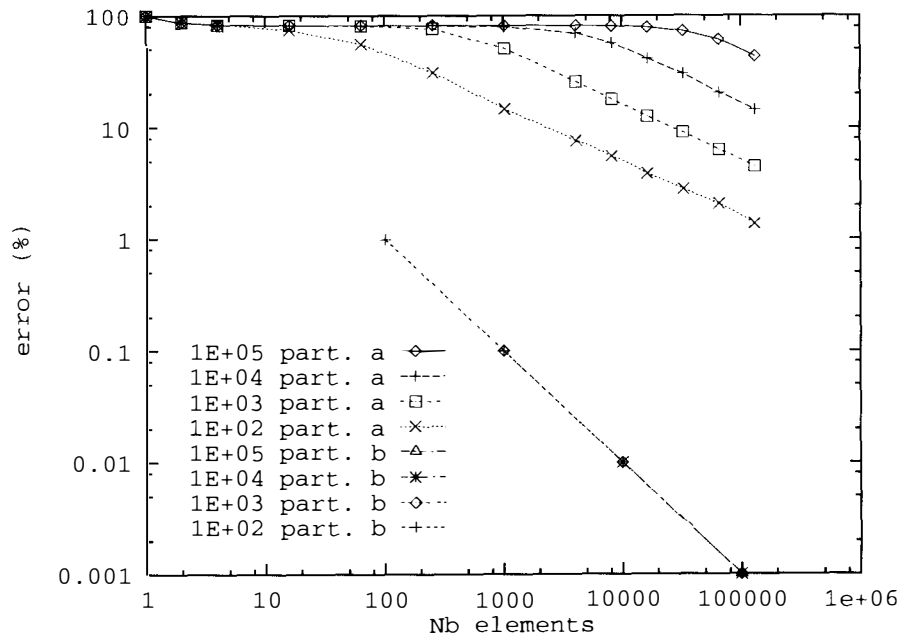


Fig. 9. Evolution of the numerical error $\epsilon_{\text{num},h}$ (%) with the number of elements in the h-mesh for several numbers of particles. The mismatch is $\tau = 10$. The a and b cases correspond to meshes not matching and matching the boundary of the particles, respectively.

If the assumptions of the Property 5 holds, one can easily show that

$$\delta_H = \frac{1}{\|\boldsymbol{u}\|^2_{E(\Omega)}} \int_{\Gamma_{\text{int}}} [(\tilde{\boldsymbol{\sigma}}^{0,H} + \tilde{\boldsymbol{\sigma}}^{0}) \cdot \boldsymbol{n}] \cdot (\boldsymbol{u}^{0,H} - \boldsymbol{u}^{0})\, \mathrm{d}\boldsymbol{s}. \tag{32}$$

For the 1-D problem under consideration, the finite element solution $\boldsymbol{u}^{0,H}$ is such that $\boldsymbol{u}^{0,H} = \boldsymbol{u}^{0}$ at the nodes. So δ_H is zero since the subdomains boundaries are chosen so as to coincide with nodes of the H-mesh. However, the classical numerical error in the energy norm $\|\boldsymbol{u}^{0} - \boldsymbol{u}^{0,H}\|_{E^0(\Omega)}$ associated to $\boldsymbol{u}^{0,H}$ is not zero. This remark highlights the fact that δ_H depends only on the quality of $\boldsymbol{u}^{0,H}$ on the boundary of the subdomains and not over the whole domain.

In order to purposely avoid exact finite element values at node, we modify the problem by introducing a non constant section $S(x)$. The problem is now

$$\frac{\mathrm{d}}{\mathrm{d}x}(S(x)E(x)\frac{\mathrm{d}u(x)}{\mathrm{d}x}) = -f(x), \qquad u(0) = 0, \qquad u(1) = 0 \quad \text{with } S(x) = e^{x} \text{ and } f(x) = e^{-x}$$

and the homogenized problem is

$$\frac{\mathrm{d}}{\mathrm{d}x}(S(x)E^{\bullet}\frac{\mathrm{d}u^{0}(x)}{\mathrm{d}x}) = -f(x), \qquad u^{0}(0) = 0, \qquad u^{0}(1) = 0.$$

Tables 2 and 3 show the behavior of $\epsilon_{\text{num},H}$ and γ_H as the H-mesh is refined for two different choices of $\boldsymbol{E}^0$.

We observe that:

- The sign of γ_H can be positive or negative depending on the choice of $\boldsymbol{E}^0$;
- $\epsilon_{\text{num},H}$ is very stable with respect to the number of subdomains. γ_H is very stable in the case $\boldsymbol{E}^0 = \langle \boldsymbol{E} \rangle$ but not in the case $\boldsymbol{E}^0 = \langle \boldsymbol{E}^{-1} \rangle^{-1}$;

Table 2
Influence of the number of elements in the H-mesh on $\epsilon_{\text{num},H}$ and γ_H for a variable number of subdomains (N). $\boldsymbol{E}^0 = \langle \boldsymbol{E}^{-1} \rangle^{-1}$ and $\tau = 10$

		Number of elements in the H-mesh (1000 particles)						
	N	4	8	16	32	64	128	256
$\epsilon_{\text{num},H}$	4	1.00(−2)	2.53(−3)	6.34(−4)	1.59(−4)	3.97(−5)	9.91(−6)	2.48(−6)
	16			6.66(−4)	1.67(−4)	4.16(−5)	1.04(−5)	2.60(−6)
	64					4.28(−5)	1.07(−5)	2.67(−6)
γ_H	4	−2.70(−4)	−6.79(−5)	−1.70(−5)	−4.25(−6)	−1.06(−6)	−2.66(−7)	−6.64(−8)
	16			−4.19(−5)	−1.05(−5)	−2.62(−6)	−6.55(−7)	−1.64(−7)
	64					−6.69(−6)	−1.67(−6)	−4.18(−7)

Table 3
Influence of the number of elements in the H-mesh on $\epsilon_{\text{num},H}$ and γ_H for a variable number of subdomains (N). $\boldsymbol{E}^0 = \langle \boldsymbol{E} \rangle$ and $\tau = 10$

		Number of elements in the H-mesh (1000 particles)						
	N	4	8	16	32	64	128	256
$\epsilon_{\text{num},H}$	4	3.33(−3)	8.37(−4)	2.10(−4)	5.24(−5)	1.31(−5)	3.28(−6)	8.19(−7)
	16			2.20(−4)	5.50(−5)	1.38(−5)	3.44(−6)	8.60(−7)
	64					1.41(−5)	3.53(−6)	8.84(−7)
γ_H	4	4.26(−3)	1.07(−3)	2.68(−4)	6.70(−5)	1.67(−5)	4.19(−6)	1.05(−6)
	16			2.88(−4)	7.19(−5)	1.80(−5)	4.49(−6)	1.12(−6)
	64					1.77(−5)	4.42(−6)	1.11(−6)

- $\epsilon_{\text{num},H}$ and γ_H are of the order of $O(H^2)$ as $H \to 0$. This comes from the fact that we have super-convergence at the nodes: $(\boldsymbol{u}^{0,H} - \boldsymbol{u}^0) \sim O(H^2)$.
- Since $\epsilon_{\text{num},H}$ and γ_H converge at the same rate, the influence of $\epsilon_{\text{num},H}$ is quickly negligible in front of the influence of γ_H in δ_H, see (30).

Finally, to compare the influence of the H-mesh and of the h-mesh on the error, we must compare δ_H to $\epsilon^2_{\text{num},h}$ and observe that:

- $\epsilon_{\text{num},h}$ is always positive; δ_H is not necessarily positive or negative;
- $\epsilon_{\text{num},h}$ depends on the quality of $\tilde{\boldsymbol{u}}^{0,H,h}$ on each subdomain; δ_H depends on the quality of $\tilde{\boldsymbol{u}}^{0,H}$ only on the boundary of the subdomains;
- The rate of convergence of δ_H as $H \to 0$ depends on the regularity of the homogenized solution $\boldsymbol{u}^0$. On the contrary, the rate of convergence of $\epsilon_{\text{num},h}$ as $h \to 0$ depends on the regularity of the local solution $\tilde{\boldsymbol{u}}^{0,H}$ and the type of mesh (matching or non matching mesh).

3.3. *Influence of the subdomain size*

Let us introduce the notations

$$\epsilon_{\text{mod,o}} = \frac{\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}}, \qquad \epsilon_{\text{mod},\Delta} = \frac{\|\boldsymbol{u}^0 - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}}. \tag{33}$$

Fig. 10 shows the behavior of $\epsilon_{\text{mod},\Delta}$ as the size of the subdomains, Δ, tends to zero. One can observe, first, a very slow decrease of $\epsilon_{\text{mod},\Delta}$. Then, when the size of the subdomains is around three or four times the size of the particles, $\epsilon_{\text{mod},\Delta}$ decreases as the square root of the size of the subdomains. Fig. 10 is very similar to Fig. 6. The critical subdomain size to obtain a significant decrease of $\epsilon_{\text{mod},\Delta}$ is the same as the critical h size to obtain a significant decrease of the numerical error $\epsilon_{\text{num},h}$: this critical length is three to four times the size of the particles. Furthermore, the asymptotic rates of convergence are the same.

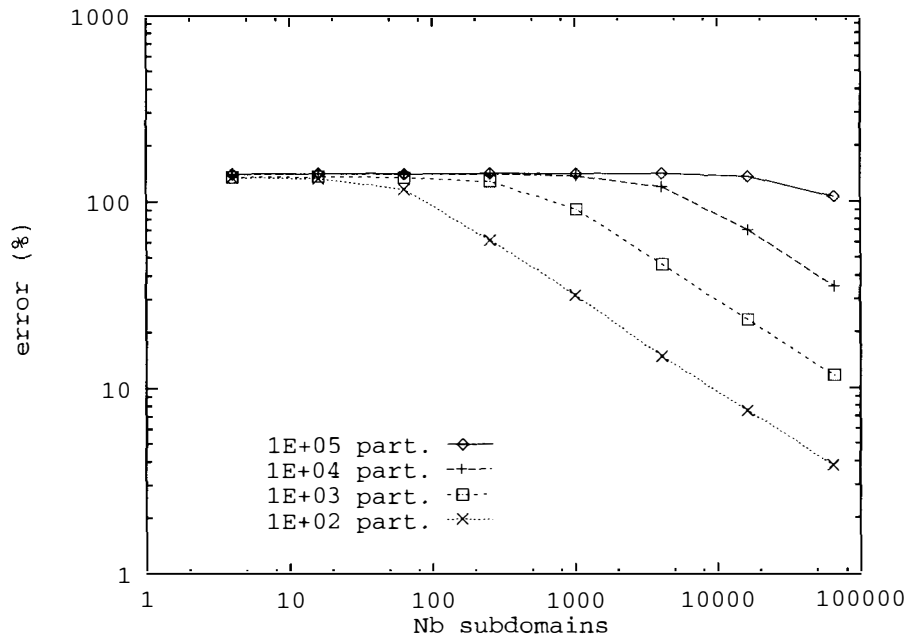


Fig. 10. Evolution of $\epsilon_{\text{mod},\Delta}$ (%) with the number of subdomains, for several numbers of particles. The mismatch is $\tau = 10$ and $\boldsymbol{E}^{\bullet} = \langle \boldsymbol{E}^{-1} \rangle^{-1}$.

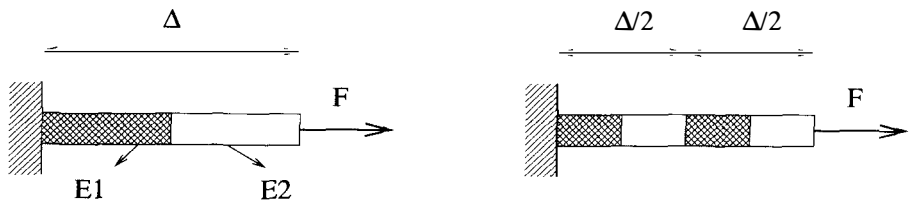


Fig. 11. These two cases have the same $\epsilon_{\text{mod},\Delta}$ error.

There is however, a conceptual difference between $\epsilon_{\text{mod},\Delta}$ and $\epsilon_{\text{num},h}$: $\epsilon_{\text{mod},\Delta}$ is what we gain by solving exactly the fine scale problems and $\epsilon_{\text{num},h}$ is the numerical error introduced by solving the local problem. Thus, the size of interest for the subdomain is *at least* three or four times the size of the particles and the size of interest for h is *at most* three or four times the size of the particles. The critical length, three to four times the size of particles, separates the long length range (homogenized problem) and the short length range (local analysis).

Fig. 11 shows two cases for which $\epsilon_{\text{mod},\Delta}$ is the same: $\epsilon_{\text{mod},\Delta} = 0.5|1 - \tau|/\sqrt{\tau}$. With $\tau = 10$, we have an error of 142.3%. This value is very close to the constant error value in Fig. 10. The relationship between Figs. 11 and 7 is noteworthy.

4. The error-in-the-constitutive-law framework

The homogenized solution and the exact solution both fulfill the kinematic constraints (1) and the equilibrium equation (2). They differ because they do not satisfy the same constitutive law. A way to

measure the distance between the two solutions is to compute the distance in the energy norm, between the exact and homogenized displacement: $e = \|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)}$. Another way is to measure the way the couple $(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)$ satisfies the fine-scale constitutive law.

The error on the constitutive law concept is very general and relies on a strong mechanical interpretation. It has been applied to many types of problem. Let us cite the design of error estimator for finite element computation [3] and the adjustment of finite element models using vibration tests [4]. For our purposes, this concept can be summarized as follows. Suppose we have at hand a stress–strain couple $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}})$ satisfying the kinematic constraints (1) and the equilibrium equation (2). This couple will be the exact solution of the problem if and only if it satisfies the actual constitutive law:

$$\hat{\boldsymbol{\sigma}} = \boldsymbol{E}\hat{\boldsymbol{\epsilon}} \quad \text{on } \Omega. \tag{34}$$

Thus, the quality of the couple $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}})$ can be measured by the way it satisfies (34). Let us introduce the quantity η defined by

$$\eta(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}) = \varphi^*(\hat{\boldsymbol{\sigma}}) + \varphi(\hat{\boldsymbol{\epsilon}}) - \hat{\boldsymbol{\sigma}} : \hat{\boldsymbol{\epsilon}}, \quad \text{where} \quad \varphi^*(\hat{\boldsymbol{\sigma}}) = \frac{1}{2}\hat{\boldsymbol{\sigma}} : \boldsymbol{E}^{-1}\hat{\boldsymbol{\sigma}} \quad \text{and} \quad \varphi(\hat{\boldsymbol{\epsilon}}) = \frac{1}{2}\hat{\boldsymbol{\epsilon}} : \boldsymbol{E}\hat{\boldsymbol{\epsilon}}.$$

$\varphi(\hat{\boldsymbol{\epsilon}})$ is the strain energy and $\varphi^*(\hat{\boldsymbol{\sigma}})$ is the complementary energy, its dual via the classical Legendre–Fenchel transformation,

$$\varphi^*(\hat{\boldsymbol{\sigma}}) = \sup_{\boldsymbol{\epsilon}'} (\hat{\boldsymbol{\sigma}} : \boldsymbol{\epsilon}' - \varphi(\boldsymbol{\epsilon}')).$$

It follows that η has the following two classical properties:

$$\eta(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}) \geqslant 0 \quad \forall (\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}); \tag{35}$$

$$\eta(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}) = 0 \Leftrightarrow \hat{\boldsymbol{\sigma}} = \boldsymbol{E}\hat{\boldsymbol{\epsilon}}. \tag{36}$$

Thus, a measure of the absolute error associated with the couple $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}})$ may be defined by:

$$\Upsilon(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}) = \left\{ 2 \int_\Omega \eta(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}) \, \mathrm{d}\boldsymbol{x} \right\}^{1/2} \tag{37}$$

We shall refer to this error-in-the-constitutive-law approach as the ECL framework.

It is interesting to note that

$$\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)} = \Upsilon(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) \quad \text{and} \quad \|\boldsymbol{u}\|_{E(\Omega)} = \Upsilon(\boldsymbol{\sigma}, 0). \tag{38}$$

In other words, the distance between $\boldsymbol{u}$ and $\boldsymbol{u}^0$ in the energy norm, is equivalent to the way the couple $(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)$ satisfies the fine-scale constitutive law as measured by the functional $\Upsilon(\cdot, \cdot)$.

We shall now rewrite and demonstrate the main results of the HDPM in the ECL framework. This will lead to a mechanical interpretation of the results. We will also establish a new result.

- The first result of the HDPM is that it is possible to compute an explicit a posteriori upper bound for the homogenization error [1]. In particular, the exact solution need not be known to compute the bound:

$$\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)} \leqslant \zeta \quad \text{where} \quad \zeta = \|\mathcal{I}_0 \nabla \boldsymbol{u}^0\|_{E(\Omega)}, \qquad \mathcal{I}_0 = I - \boldsymbol{E}^{-1}\boldsymbol{E}^0. \tag{39}$$

After some manipulations, this result can be proved to be equivalent to

$$\Upsilon(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) \leqslant \Upsilon(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0). \tag{40}$$

In other words, an upper bound is obtained by replacing in the error in the constitutive law the exact stress field by the homogenized field. $\Upsilon(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)$ measures the way the homogenized solution satisfies the actual behavior. The following Lemma is the key to demonstrate all the main results of the HDPM using the ECL framework:

LEMMA 1. Let $(\boldsymbol{\sigma}_1, \boldsymbol{\epsilon}_1)$ and $(\boldsymbol{\sigma}_2, \boldsymbol{\epsilon}_2)$ be two stress–strain couples satisfying the following orthogonality condition:

$$\int_{\Omega} (\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2) : (\boldsymbol{\epsilon}_1 - \boldsymbol{\epsilon}_2)\, \mathrm{d}\boldsymbol{x} = 0. \tag{41}$$

Then, we have

$$\Upsilon^2(\boldsymbol{\sigma}_1, \boldsymbol{\epsilon}_1) + \Upsilon^2(\boldsymbol{\sigma}_2, \boldsymbol{\epsilon}_2) = \Upsilon^2(\boldsymbol{\sigma}_1, \boldsymbol{\epsilon}_2) + \Upsilon^2(\boldsymbol{\sigma}_2, \boldsymbol{\epsilon}_1). \quad \square$$

The pairs $(\boldsymbol{\sigma}, \boldsymbol{\epsilon})$ and $(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)$ satisfy the orthogonality condition. Moreover, $\Upsilon^2(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) = 0$. So, by the Lemma 1, we have

$$\Upsilon^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0) = \Upsilon^2(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) + \Upsilon^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}) \tag{42}$$

which proves the upper-bound. $\square$

- The second principle of the HDPM is that the modeling error is reduced by carrying out the subdomain analysis [1]. Suppose that the subdomain analysis is performed on the subdomain Ω_k. We have

$$\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|_{E(\Omega_k)} \leqslant \|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega_k)}.$$

This result may be rewritten

$$\Upsilon_k(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0) \leqslant \Upsilon_k(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) \tag{43}$$

where

$$\Upsilon_k(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}}) = \left\{ 2 \int_{\Omega_k} \eta(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}})\, \mathrm{d}\boldsymbol{x} \right\}^{1/2}.$$

PROOF OF (43). The pairs $(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)$ and $(\tilde{\boldsymbol{\sigma}}^0, \tilde{\boldsymbol{\epsilon}}^0)$ satisfy the orthogonality condition on Ω_k. Then by Lemma 1, since $\Upsilon_k(\tilde{\boldsymbol{\sigma}}^0, \tilde{\boldsymbol{\epsilon}}^0) = 0$, we have

$$\Upsilon_k^2(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) = \Upsilon_k^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0) + \Upsilon_k^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0), \tag{44}$$

proving the result. $\square$

- The ζ quantity defined in the first result is not, in general, an upper bound locally, i.e. the following inequality is not satisfied in general over a given subdomain, Ω_k:

$$\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega_k)} \leqslant \zeta_k \quad \text{where} \quad \zeta_k = \|\mathcal{I}_0 \nabla \boldsymbol{u}^0\|_{E(\Omega_k)}. \tag{45}$$

However, the following inequality, which is the third result of the HDPM [1], holds:

$$\|\tilde{\boldsymbol{u}}^0 - \boldsymbol{u}^0\|_{E(\Omega_k)} \leqslant \zeta_k. \tag{46}$$

In other words, it is possible to determine locally where the local solution process will produce a significant change in the homogenized solution.

In terms of the ECL, we have

$$\Upsilon_k(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0) \leqslant \Upsilon_k(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0). \tag{47}$$

PROOF OF (47). Applying the Lemma 1 with the pairs $(\tilde{\boldsymbol{\sigma}}^0, \tilde{\boldsymbol{\epsilon}}^0)$ and $(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)$, we have

$$\Upsilon_k^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0) = \Upsilon_k^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0) + \Upsilon_k^2(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0). \quad \square$$

- The fourth principle is that an upper bound also exists for the modeling error obtained after the subdomain analysis [2]:

$$\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)} \leqslant \psi \quad \text{where} \quad \psi^2 = 2(\mathcal{J}(\tilde{\boldsymbol{u}}^0) - \mathcal{J}(\boldsymbol{u}^0)) + \zeta^2. \tag{48}$$

$\mathcal{J}(\cdot)$ is the potential energy associated to the displacement field $\cdot$:

$$\mathcal{J}(\cdot) = \int_{\Omega} \varphi(\boldsymbol{\epsilon}(\cdot))\, \mathrm{d}\boldsymbol{x} - \int_{\Omega} \boldsymbol{f}_{\mathrm{g}} \cdot (\cdot)\, \mathrm{d}\boldsymbol{x} - \int_{\Gamma_{\mathrm{t}}} \boldsymbol{t}_{\mathrm{g}} \cdot (\cdot)\, \mathrm{d}\boldsymbol{s}.$$

The result (48) may be rewritten

$$\Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0) \leqslant \Upsilon(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0). \tag{49}$$

In other words, as for the first result, an upper bound is obtained by replacing the exact stress by the homogenized one.

PROOF OF (49). Applying the Lemma 1 with the pairs $(\boldsymbol{\sigma}, \boldsymbol{\epsilon})$ and $(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0)$, we have

$$\Upsilon^2(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0) = \Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0) + \Upsilon^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}). \quad \square \tag{50}$$

The ECL framework has allowed us to find a new mechanical interpretation of the HDPM results: the bounds are obtained by using the homogenized stress field in the error in the constitutive law expression. In the first and last results, it is the global exact stress field $\boldsymbol{\sigma}$ that is replaced by the homogenized one (see (40) and (49)) and in the third result it is the local stress field $\tilde{\boldsymbol{\sigma}}^0$ that is replaced by the homogenized stress (see (47)).

We also note that all the HDPM results now extend to the nonlinear case. In the nonlinear case, all the proofs remain valid, only the expression of the potential φ and φ^* change. Let us inquire the meaning of the Y functional in the nonlinear case. Assuming that $\hat{\boldsymbol{\sigma}}$ satisfies the equilibrium equation (2) and $\hat{\boldsymbol{u}}$ satisfies the kinematic constraints (1), we can write

$$\frac{1}{2}\Upsilon^2(\hat{\boldsymbol{\sigma}}, \boldsymbol{\epsilon}(\hat{\boldsymbol{u}})) = \mathcal{J}(\hat{\boldsymbol{u}}) - \Pi(\hat{\boldsymbol{\sigma}}) \tag{51}$$

where $\mathcal{J}(\hat{\boldsymbol{u}})$ is the potential energy and $\Pi(\hat{\boldsymbol{\sigma}})$ the complementary potential energy:

$$\mathcal{J}(\hat{\boldsymbol{u}}) = \int_\Omega \varphi(\boldsymbol{\epsilon}(\hat{\boldsymbol{u}}))\, \mathrm{d}\boldsymbol{x} - \int_\Omega \boldsymbol{f}_\mathrm{g} \cdot (\hat{\boldsymbol{u}})\, \mathrm{d}\boldsymbol{x} - \int_{\Gamma_\mathrm{t}} \boldsymbol{t}_\mathrm{g} \cdot (\hat{\boldsymbol{u}})\, \mathrm{d}\boldsymbol{s}, \tag{52}$$

$$\Pi(\hat{\boldsymbol{\sigma}}) = -\int_\Omega \varphi^*(\hat{\boldsymbol{\sigma}})\, \mathrm{d}\boldsymbol{x} + \int_{\Gamma_\mathrm{u}} (\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{n}) \cdot \boldsymbol{u}_\mathrm{g}\, \mathrm{d}\boldsymbol{s}. \tag{53}$$

Thus Y is linked to the difference between the potential energy and the complementary potential energy. It is always positive and zero if and only if the couple $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\epsilon}} = \boldsymbol{\epsilon}(\hat{\boldsymbol{u}}))$ satisfies the behavior described by the relation:

$$\varphi^*(\hat{\boldsymbol{\sigma}}) + \varphi(\hat{\boldsymbol{\epsilon}}) - \hat{\boldsymbol{\sigma}} : \hat{\boldsymbol{\epsilon}} = 0. \tag{54}$$

The main results of the HDPM are stated in terms of explicit upper-bounds. The quality of these bounds, is defined by the following effectivity indices, all greater than or equal to one:

$$\theta_0 = \frac{\zeta}{\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)}}, \qquad \widetilde{\theta}_0 = \frac{\psi}{\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)}}, \qquad \theta_{\mathrm{s},k} = \frac{\zeta_k}{\|\tilde{\boldsymbol{u}}^0 - \boldsymbol{u}^0\|_{E(\Omega_k)}},$$

or equivalently, in terms of the error in the constitutive law:

$$\theta_0 = \frac{\Upsilon(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)}{\Upsilon(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)}, \qquad \widetilde{\theta}_0 = \frac{\Upsilon(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0)}{\Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0)}, \qquad \theta_{\mathrm{s},k} = \frac{\Upsilon_k(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)}{\Upsilon_k(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)}.$$

We may also define the global sensitivity effectivity index, θ_s:

$$\theta_\mathrm{s} = \frac{\zeta}{\|\tilde{\boldsymbol{u}}^0 - \boldsymbol{u}^0\|_{E(\Omega)}} = \frac{\Upsilon(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)}{\Upsilon(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)}.$$

Of course, θ_s is bounded from below by the best effectivity over the subdomain and from above by the worst effectivity index:

$$\min_{k=1,\ldots,N} \theta_{\mathrm{s},k} \leqslant \theta_\mathrm{s} \leqslant \max_{k=1,\ldots,N} \theta_{\mathrm{s},k}.$$

The following property establishes that the effectivity index on the ψ estimate, $\widetilde{\theta}_0$, is always greater than the effectivity index on the ζ estimate, θ_0. The global sensitivity index, θ_s, is also always greater than θ_0. Finally, θ_0, $\widetilde{\theta}_0$ and θ_s are linked in a relation:

PROPERTY 6.

- $\theta_0 \leqslant \widetilde{\theta}_0$ and $\theta_0 = 1 \Leftrightarrow \widetilde{\theta}_0 = 1$;
- $\theta_0 \leqslant \theta_s$;
- $(\theta_0^2 - \theta_0^2)(\theta_s^2 - \theta_0^2) = \theta_0^2(\theta_0^2 - 1)$.

PROOF. Applying the Lemma 1 to the pairs $(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)$ and $(\tilde{\boldsymbol{\sigma}}^0, \tilde{\boldsymbol{\epsilon}}^0)$, we obtain since $\Upsilon(\tilde{\boldsymbol{\sigma}}^0, \tilde{\boldsymbol{\epsilon}}^0) = 0$

$$\Upsilon^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0) = \Upsilon^2(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0). \tag{55}$$

Summing relation (44) over all the subdomains yields

$$\Upsilon^2(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) = \Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0). \tag{56}$$

Dividing relation (55) by relation (56), we obtain

$$\theta_0^2 = \frac{\Upsilon^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)}{\Upsilon^2(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)} = \frac{\Upsilon^2(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)}{\Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)} \leqslant \frac{\Upsilon^2(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0)}{\Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0)} = \widetilde{\theta}_0^2 \tag{57}$$

and clearly $\widetilde{\theta}_0 = 1 \Leftrightarrow \widetilde{\theta}_0 = 1$. The second relation is obvious since

$$\Upsilon(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0) \leqslant \Upsilon(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)$$

and the third one is obtained by eliminating the quantity $\Upsilon(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)$ between the two relations:

$$\theta_s = \frac{\Upsilon(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)}{\Upsilon(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)} \quad \text{and} \quad \widetilde{\theta}_0^2 = \frac{\Upsilon^2(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0) - \Upsilon^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)}{\Upsilon^2(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0) - \Upsilon^2(\tilde{\boldsymbol{\sigma}}^0, \boldsymbol{\epsilon}^0)}. \quad \square$$

Table 4 shows the influence of the number of subdomains on the modeling error

$$\epsilon = \frac{\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^0\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}} = \frac{\Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^0)}{\Upsilon(\boldsymbol{\sigma}, 0)}$$

and on the effectivity index $\widetilde{\theta}_0$. We see that $\widetilde{\theta}_0$ tends monotonically to θ_0 as the subdomain size is reduced.

Table 4
Evolution of the modeling error ϵ and the effectivity index $\widetilde{\theta}_0$ with respect to the number of subdomains, N, for three different values of τ. $\boldsymbol{E}^0 = \langle \boldsymbol{E}^{-1} \rangle^{-1}$. When $N = +\infty$, $\epsilon = \frac{\|\boldsymbol{u} - \boldsymbol{u}^0\|_{E(\Omega)}}{\|\boldsymbol{u}\|_{E(\Omega)}} = \frac{\Upsilon(\boldsymbol{\sigma}, \boldsymbol{\epsilon}^0)}{\Upsilon(\boldsymbol{\sigma}, 0)}$ and $\widetilde{\theta}_0 = \theta_0$

	$\tau = 10$		$\tau = 50$		$\tau = 100$	
N	ϵ	$\widetilde{\theta}_0$	ϵ	$\widetilde{\theta}_0$	ϵ	$\widetilde{\theta}_0$
1	0.	$+\infty$	0.	$+\infty$	0.	$+\infty$
2	2.702(−4)	44.70039184	3.114(−4)	45.66017533	3.169(−4)	45.79441773
4	1.739(−2)	1.21748697	2.046(−2)	1.21758072	2.088(−2)	1.21759302
16	3.505(−2)	1.05768583	4.138(−2)	1.05734892	4.225(−2)	1.05730230
64	7.672(−2)	1.01231204	9.128(−2)	1.01205426	9.329(−2)	1.01201730
256	1.499(−1)	1.00323777	1.809(−1)	1.00308147	1.854(−1)	1.00305648
1024	3.232(−1)	1.00069796	4.959(−1)	1.00041075	6.074(−1)	1.00028513
4096	7.648(−1)	1.00012465	1.692	1.00003527	2.356	1.00001896
$+\infty$	1.421	1.00003609	3.460	1.00000844	4.944	1.00000431

5. Numerical results for the HDPM

The results of the HDPM method rely on the assumption that the homogenized and the local solutions are free of numerical errors. In practice, this is not the case. Thus, we seek for the meaning of the results, mainly the two first results, when numerical errors occur. We keep the error-in-the-constitutive-law framework allowing us to deal with both linear or nonlinear constitutive laws.

It is interesting to note that the inequality expressed in (40) still holds if the exact fields are replaced by the approximate fields. Defining $(\boldsymbol{\sigma}^H, \boldsymbol{\epsilon}^H)$ as the finite element solution obtained with the real material and the H-mesh, we have

PROPERTY 7.

$$\Upsilon(\boldsymbol{\sigma}^H, \boldsymbol{\epsilon}^{0,H}) \leqslant \Upsilon(\boldsymbol{\sigma}^{0,H}, \boldsymbol{\epsilon}^{0,H}).$$

PROOF. The proof is obtained by the Lemma 1 using the pairs $(\boldsymbol{\sigma}^H, \boldsymbol{\epsilon}^H)$ and $(\boldsymbol{\sigma}^{0,H}, \boldsymbol{\epsilon}^{0,H})$. The orthogonality condition (41) is indeed satisfied at the finite element level. □

Defining the numerical effectivity index, θ_0^H, as

$$\theta_0^H = \frac{\Upsilon(\boldsymbol{\sigma}^{0,H}, \boldsymbol{\epsilon}^{0,H})}{\Upsilon(\boldsymbol{\sigma}^H, \boldsymbol{\epsilon}^H)},$$

we thus have $\theta_0^H \geqslant 1$ like $\theta_0 \geqslant 1$. Numerical experiments carried out in [1] for 1-D and 3-D problems confirm the property, although in this study an iterative solver and approximate spatial integration were used in the 3-D case. It was also noticed in [1] that θ_0^H is very stable with respect to the mesh size H. Finally, note that, as the ζ estimate ($\Upsilon(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0)$), $\Upsilon(\boldsymbol{\sigma}^{0,H}, \boldsymbol{\epsilon}^{0,H})$ is zero if E and E^0 coincide, even though numerical errors occur.

By performing the local analysis, we know from the second result of the HDPM that we obtain a perturbation of the homogenized solution, $\tilde{\boldsymbol{u}}^0$, that is closer to the exact solution than $\boldsymbol{u}^0$. This might not hold for the numerical approximation of $\tilde{\boldsymbol{u}}^0$ because it is not an exact solution of the subdomain problems. Fortunately, we have the following result:

PROPERTY 8. Let $\boldsymbol{u}_k^{0,H}$ be the restriction of $\boldsymbol{u}^{0,H}$ to Ω_k and $\boldsymbol{\epsilon}_k^{0,H} = \boldsymbol{\epsilon}(\boldsymbol{u}_k^{0,H})$. If $(\boldsymbol{u}_k^{0,H} - \tilde{\boldsymbol{u}}_k^{0,H,h}) \in V_k^h$ then

$$\Upsilon_k(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}_k^{0,H,h}) \leqslant \Upsilon_k(\boldsymbol{\sigma}, \boldsymbol{\epsilon}_k^{0,H}).$$

PROOF. We apply Lemma 1 with the pairs $(\tilde{\boldsymbol{\sigma}}_k^{0,H,h}, \tilde{\boldsymbol{\epsilon}}_k^{0,H,h})$ and $(\boldsymbol{\sigma}, \boldsymbol{\epsilon}_k^{0,H})$. Since $(\boldsymbol{u}_k^{0,H} - \tilde{\boldsymbol{u}}_k^{0,H,h}) \in V_k^h$ the orthogonality condition holds. □

Finally, we stress that two quantities equivalent when free of numerical errors may become quite different in the case of numerical error. For instance, as seen in the previous section, the following two expressions of ψ are equivalent:

$$\psi = \left(2(\mathcal{J}(\tilde{\boldsymbol{u}}^0) - \mathcal{J}(\boldsymbol{u}^0)) + \|\mathcal{I}_0 \nabla \boldsymbol{u}^0\|^2_{E(\Omega)}\right)^{1/2},$$

$$\psi = \Upsilon(\boldsymbol{\sigma}^0, \tilde{\boldsymbol{\epsilon}}^0).$$

If we inject the approximate solution, we get

$$\psi^{\text{num1}} = \left(2(\mathcal{J}(\tilde{\boldsymbol{u}}^{0,H,h}) - \mathcal{J}(\boldsymbol{u}^{0,H})) + \|\mathcal{I}_0 \nabla \boldsymbol{u}^{0,H}\|^2_{E(\Omega)}\right)^{1/2},$$

$$\psi^{\text{num2}} = \Upsilon(\boldsymbol{\sigma}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}).$$

ψ^{num1} and ψ^{num2} no longer coincide since the first one may be the square root of a negative value, whereas the second one is always well defined.

6. A posteriori error estimation and adaptive strategy

From (23), we recall that

$$\|\boldsymbol{u} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)} = \|\boldsymbol{u} - \tilde{\boldsymbol{u}}^{0,H}\|^2_{E(\Omega)} + \|\tilde{\boldsymbol{u}}^{0,H} - \tilde{\boldsymbol{u}}^{0,H,h}\|^2_{E(\Omega)}$$

which can be rewritten

$$\Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) = \Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^{0,H}) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}). \tag{58}$$

The Properties 9, 10 and 11 below give upper-bounds for the three terms in (58). The spaces $\mathcal{U}$ and $\mathcal{S}$ describe the regularity imposed to the displacement and stress field, respectively.

PROPERTY 9.

$$\Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) \leqslant \Upsilon(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) \quad \forall\, \hat{\boldsymbol{\sigma}} \in \mathcal{S}_{\text{ad}}$$

where

$$\mathcal{S}_{\text{ad}} = \left\{ \boldsymbol{\sigma} \in \mathbb{L}^2_{\text{sym}}(\Omega) : \int_\Omega \boldsymbol{\sigma} : \nabla \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} - \mathcal{F}(\boldsymbol{v}) = 0 \quad \forall \boldsymbol{v} \in V \right\}, \tag{59}$$

$$\mathbb{L}^2_{\text{sym}} = \left\{ \boldsymbol{\tau} = \{\tau_{ij}\} \in (L^2(\Omega))^{n\times n}, \boldsymbol{\tau} = \boldsymbol{\tau}^t \right\}. \tag{60}$$

PROOF. With the pairs $(\boldsymbol{\sigma}, \boldsymbol{\epsilon})$ and $(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^{0,H,h})$, Lemma 1 gives

$$\Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) = \Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) + \Upsilon^2(\hat{\boldsymbol{\sigma}}, \boldsymbol{\epsilon}). \quad \square \tag{61}$$

PROPERTY 10.

$$\Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^{0,H}) \leqslant \Upsilon(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^{0,H}) \quad \forall\, \hat{\boldsymbol{\sigma}} \in \mathcal{S}_{\text{ad}}.$$

PROOF. With the pairs $(\boldsymbol{\sigma}, \boldsymbol{\epsilon})$ and $(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^{0,H})$, we get from Lemma 1

$$\Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\epsilon}}^{0,H}) = \Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\epsilon}}^{0,H}) + \Upsilon^2(\hat{\boldsymbol{\sigma}}, \boldsymbol{\epsilon}). \quad \square \tag{62}$$

PROPERTY 11.

$$\Upsilon(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) \leqslant \Upsilon(\hat{\tilde{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) \quad \forall\, \hat{\tilde{\boldsymbol{\sigma}}} \in \tilde{\mathcal{S}}_{\text{ad}}$$

where

$$\tilde{\mathcal{S}}_{\text{ad}} = \prod_{k=1}^{N} \{ \boldsymbol{\tau} \in \mathbb{L}^2_{\text{sym}}(\Omega_k) : \int_{\Omega_k} \boldsymbol{\tau} : \nabla \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} = \int_{\Omega_k} \boldsymbol{f}_{\text{g}} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x} + \int_{\Gamma_{\text{t}} \cap \partial\Omega_k} \boldsymbol{t}_{\text{g}} \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{s},$$
$$\forall \boldsymbol{v} \in \boldsymbol{H}^1(\Omega_k), \boldsymbol{v} = 0 \text{ on } \Gamma_{\text{int}} \cup \Gamma_{\text{u}} \} \tag{63}$$

PROOF. With the pairs $(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H})$ and $(\hat{\tilde{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\epsilon}}^{0,H,h})$, we have by Lemma 1

$$\Upsilon^2(\hat{\tilde{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) = \Upsilon^2(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\epsilon}}^{0,H,h}) + \Upsilon^2(\hat{\tilde{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\epsilon}}^{0,H}). \quad \square \tag{64}$$

Note that this last Property is also valid on each subdomain. Finally, it is worth mentioning that for linear elasticity, the relations (61), (62) and (64) may also be obtained using the hyper-circle theorem [5].

Let us define the three following effectivity indices, all greater than one:

$$\theta = \frac{\hat{e}}{e}, \qquad \theta_{\text{mod}} = \frac{\hat{e}_{\text{mod}}}{e_{\text{mod}}}, \qquad \theta_{\text{num}} = \frac{\hat{e}_{\text{num}}}{e_{\text{num}}}$$

where

$$\hat{e} = \Upsilon(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h}), \qquad \hat{e}_{\text{mod}} = \Upsilon(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H}), \qquad \hat{e}_{\text{num}} = \Upsilon(\hat{\tilde{\boldsymbol{\sigma}}}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h}), \tag{65}$$

$$e = \Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h}), \qquad e_{\text{mod}} = \Upsilon(\boldsymbol{\sigma}, \tilde{\boldsymbol{\varepsilon}}^{0,H}), \qquad e_{\text{num}} = \Upsilon(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h}). \tag{66}$$

We have the following property:

PROPERTY 12. If $\hat{e}$ and $\hat{e}_{\text{mod}}$ are computed with the same stress field $\hat{\boldsymbol{\sigma}} \in \mathcal{S}_{\text{ad}}$, we have: $\theta \leqslant \theta_{\text{mod}}$ and $\theta = 1. \Leftrightarrow \theta_{\text{mod}} = 1.$

PROOF. With the pairs $(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h})$ and $(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\varepsilon}}^{0,H})$, Lemma 1 gives

$$\Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h}) = \Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H}) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h}). \tag{67}$$

Dividing (67) by (58), we get

$$\theta^2 = \frac{\Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h})}{\Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h})} = \frac{\Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H}) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h})}{\Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\varepsilon}}^{0,H}) + \Upsilon^2(\tilde{\boldsymbol{\sigma}}^{0,H}, \tilde{\boldsymbol{\varepsilon}}^{0,H,h})} \leqslant \frac{\Upsilon^2(\hat{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\varepsilon}}^{0,H})}{\Upsilon^2(\boldsymbol{\sigma}, \tilde{\boldsymbol{\varepsilon}}^{0,H})} = \theta^2_{\text{mod}} \tag{68}$$

and clearly $\theta_{\text{mod}} = 1. \Leftrightarrow \theta = 1.$ □

From the finite element stresses $\boldsymbol{\sigma}^{0,H}$, it is possible to build explicitly a stress field $\hat{\boldsymbol{\sigma}}^{0,H}$ belonging to $\mathcal{S}_{\text{ad}}$ [3]. Similarly, it is possible to build from the finite element stresses $\tilde{\boldsymbol{\sigma}}^{0,H,h}$ stresses $\hat{\tilde{\boldsymbol{\sigma}}}^{0,H,h}$ belonging to $\tilde{\mathcal{S}}_{\text{ad}}$. Thus, the upper bound $\hat{e}$ and $\hat{e}_{\text{num}}$ can be practically computed since $\tilde{\boldsymbol{\varepsilon}}^{0,H,h}$ is known. On the contrary, $\hat{e}_{\text{mod}}$ cannot be computed since $\tilde{\boldsymbol{\varepsilon}}^{0,H}$ is unknown. However, when $\hat{e}^2 \geqslant \hat{e}^2_{\text{num}}$, we can evaluate e_{mod} by $\tilde{e}_{\text{mod}}$:

$$\tilde{e}^2_{\text{mod}} = \hat{e}^2 - \hat{e}^2_{\text{num}}$$

and define the effectivity index

$$\tilde{\theta}_{\text{mod}} = \tilde{e}_{\text{mod}}/e_{\text{mod}}.$$

This latter effectivity is not necessarily greater than one.

Table 5 gives the results for the estimated errors and the effectivity indices for the model problem. The relative errors $\hat{\epsilon}$, $\hat{\epsilon}_{\text{num}}$ and $\tilde{\epsilon}_{\text{mod}}$ are defined by

$$\hat{\epsilon} = \hat{e}/\|\boldsymbol{u}\|_{E(\Omega)}, \qquad \hat{\epsilon}_{\text{num}} = \hat{e}_{\text{num}}/\|\boldsymbol{u}\|_{E(\Omega)}, \qquad \tilde{\epsilon}_{\text{mod}} = \tilde{e}_{\text{mod}}/\|\boldsymbol{u}\|_{E(\Omega)}.$$

One can see that the effectivity indices, θ and θ_{num}, are close to one, especially for θ. For the last two meshes, the modeling error can be evaluated by the difference between the estimated total error and the numerical errors and the modeling effectivity index is very good.

6.1. *A simple adaptive strategy*

An effective adaptive strategy for the HDPM should be able to (1) select the best suited homogenized material property $\boldsymbol{E}^0$, (2) partition Ω into subdomains, (3) produce the H-mesh and (4) the h-mesh in

Table 5
Estimated errors (%) and effectivity indices with a growing number of elements in the h-mesh. There are 1000 particles, 256 subdomains and the mismatch is $\tau = 10$

Nb elts	$\hat{\epsilon}$	θ	$\hat{\epsilon}_{\text{num}}$	θ_{num}	$\tilde{\epsilon}_{\text{mod}}$	$\tilde{\theta}_{\text{mod}}$
256	137.3	1.0001	213.3	1.66		
1024	92.07	1.0001	111.5	1.41		
4096	57.39	1.0004	34.65	1.07	45.76	0.970
16384	49.75	1.0005	15.57	1.02	47.27	0.999

Table 6
An example of simple adaptive strategy. 1000 particles, $\boldsymbol{E}^0 = \langle \boldsymbol{E}^{-1} \rangle^{-1}$ and $\tau = 10$.

N	$\hat{\epsilon}$	$\tilde{\epsilon}_{\text{mod}}$	$\hat{\epsilon}_{\text{num}}$
256	49.77	47.27	15.57
128	32.21	29.16	13.66
64	21.17	16.94	12.70

order to minimize the cost of the computation for a prescribed total accuracy. The problem of choosing $\boldsymbol{E}^0$ and the partition is already a difficult task by itself. Therefore, we make the following assumptions:

- The 'best' homogenized material property is known;
- No numerical errors are introduced by the H-mesh;
- We consider a uniform subdomain partition and uniform h-meshes;
- The subdomain analysis is carried out on each subdomain.

We try to adapt the subdomain size and the h-mesh size to reach a prescribed total accuracy $\epsilon_0\%$. Instead of minimizing the cost of the computation, we impose a given sharing of the modeling and numerical errors. This sharing is described by the parameter α $(0 < \alpha < 1)$: $\epsilon_{\text{mod}} = (\alpha)^{1/2}\epsilon_0$ and $\epsilon_{\text{num}} = (1 - \alpha)^{1/2}\epsilon_0$.

We propose the following simple strategy:

- Step 1: Start with a subdomain size which is at least three or four times the size of a particle;
- Step 2: With this size of subdomain, use standard h-adaptive finite element method to optimize the mesh to reach the prescribed numerical accuracy;
- Step 3: Iteratively reduce the number of subdomains, keeping the h-mesh fixed, until the prescribed modeling accuracy is reached.

As an example, we consider our 1-D problem with 1000 particles and $\tau = 10$. The homogenized modulus $\boldsymbol{E}^0$ is taken as $\boldsymbol{E}^0 = \langle \boldsymbol{E}^{-1} \rangle^{-1}$ and no errors are introduced by the H-mesh since the exact displacements values are obtained at the nodes (cf Section 3.2). We target a 20% error divided into 14.14% for modeling error and 14.14% for the numerical error ($\alpha = 0.5$). We start with 256 subdomains and we optimize the mesh. Note that for our 1-D model problem, if we work with meshes matching the particles boundaries the numerical error is quickly very small. In order to get substantial numerical error, we choose to work with non matching meshes (our 1-D problem is academic and it is highly advised for 2- and 3-D problem to use meshes matching the boundaries of the heterogeneities). With 16,384 elements, we get $\hat{\epsilon}_{\text{num}} = 15.57\%$. The modeling error is $\tilde{\epsilon}_{\text{mod}} = 47.27\%$. The number of subdomains, N, is decreased to 128 and then to 64. The final error is 21.17%. Table 6 gives a summary of the adaptive process.

7. Conclusions

In the Homogenized Dirichlet Projection Method, numerical errors occur when solving the homogenized problem (H-mesh) and when performing the local analysis (h-mesh). The influences of the h- and H-mesh on the error are completely different. The influence of the H-mesh is expressed in terms of the difference between the numerical and exact homogenized displacement on the boundary of the subdomains. As $H \to 0$, the total error decreases or increases depending on the choice made for the homogenized material property. Conversely, the influence of the h-mesh is expressed as the distance in the energy norm between the exact and numerical solution in displacement of the local subdomain problems. If the nodes of the h-mesh and the particle boundaries are not matching, the rate of convergence of the numerical error is very poor, $O(\sqrt{h})$ for piecewise linear elements. Moreover, this rate of convergence is achieved only if the mesh size is *smaller* than a critical length being three to four times the size of the particles. On the contrary, if the h-mesh matches the particle boundaries, the rate is $O(h)$ for piecewise linear elements and the critical mesh size is no longer present.

Concerning the modeling error, a critical length also appears, the same as for the numerical error: the subdomain size should be *bigger* than three to four times the size of the particles to allow an efficient decomposition into subdomains, for our 1-D model problem.

Using the ECL concept, we are able to give a mechanical interpretation of the main results of the HDPM and to extend them to nonlinear constitutive laws.

Finally, computable upper bounds for the total and numerical error are obtained and the effectivity indices obtained for our 1-D model problem are close to one. A simple adaptive strategy is also proposed to choose the size of the subdomain and the h-mesh size.

Acknowledgment

The authors gratefully acknowledge the support of this work by the U.S. Office of Naval Research under Grant N00014-95-1-0401 and the National Science Foundation under grant ECS-9422707.

References

[1] Tarek I. Zohdi, J.T. Oden, and Gregory J. Rodin, Hierarchical modeling of heterogeneous bodies, TICAM report 96-21, Texas Institute for Computational and Applied Mathematics, Austin, 1996. Also to appear in Comput. Methods Appl. Mech. Engrg.

[2] J.T. Oden and Tarek I. Zohdi, Analysis of elastic structures composed of highly heterogeneous materials, TICAM report 96-56, Texas Institute for Computational and Applied Mathematics, Austin, 1996. Also to appear in Comput. Methods Appl. Mech. Engrg.

[3] P. Ladevèze and D. Leguillon, Error estimate procedure in the finite element method and application, SIAM J. Numer. Anal. 20(3) (1983) 485–509.

[4] P. Ladevèze and M. Reynier, A localisation method of stiffness errors for the adjustment of F.E. models, In Special issue, 12th ASME Mechanical Vibration and Noise Conference, Montreal, 1989.

[5] W. Prager and J.L. Synge, Approximation in elasticity based on the concept of functions space, Quart. Appl. Math. 5 (1947) 261–269.