Comput. Methods Appl. Mech. Engrg. 200 (2011) 2710–2726

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Adaptive finite element heterogeneous multiscale method for homogenization problems

A. Abdulle *, A. Nonnenmacher

Section of Mathematics, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 27 August 2009
Received in revised form 29 April 2010
Accepted 8 June 2010
Available online 18 June 2010

Keywords:
Adaptive mesh refinement
A posteriori error estimate
Finite element method
Multiscale method
Heterogenenous multiscale method
Homogenization

ABSTRACT

In this paper we present an a posteriori error analysis for elliptic homogenization problems discretized by the finite element heterogeneous multiscale method. Unlike standard finite element methods, our discretization scheme relies on macro- and microfinite elements. The desired macroscopic solution is obtained by a suitable averaging procedure based on microscopic data. As the macroscopic data (such as the macroscopic diffusion tensor) are not available beforehand, appropriate error indicators have to be defined for designing adaptive methods. We show that such indicators based only on the available macro- and microsolutions (used to compute the actual macrosolution) can be defined, allowing for a macroscopic mesh refinement strategy which is both reliable and efficient. The corresponding a posteriori estimates for the upper and lower bound are derived in the energy norm. In the case of a uniformly oscillating tensor, we recover the standard residual-based a posteriori error estimate for the finite element method applied to the homogenized problem. Numerical experiments confirm the efficiency and reliability of the adaptive multiscale method.

1. Introduction

The importance of adaptive numerical methods for the solution of partial differential equations (PDEs) cannot be overemphasized. As many problems exhibit local variations or singularities it is essential to design mesh refinement strategies capable of equidistributing the approximation error and in turn optimizing the computational effort. Such strategies are based on a posteriori error analysis allowing to define computable (local) error indicators for mesh adaptation depending on the available numerical solution. A large amount of literature concerned with a posteriori error analysis is nowadays available for the numerical solution of elliptic PDEs solved by the finite element method (FEM) (see [42,11] and the references therein). However for some classes of problems such as problems with a rapidly oscillating tensor, the standard a posteriori error analysis does not usually apply, as standard FEM cannot be used. Indeed, the standard FEM usually fails to converge to the true solution, unless the smallest scale is discretized by the FE mesh. Computations with such meshes involve a huge number of degrees of freedom and are often too costly and sometimes even impossible with today's computer resources.

Considerable effort has been devoted in the past few years to design multiscale methods capable of discretizing elliptic PDE with multiple scales. In this paper we focus on homogenization problems. Multiscale FEMs for elliptic problems (based on multiscale basis functions) have been pioneered by Babŭska and Osborn [13,14]. Among the recent literature and strategies for such problems, without attempting to be exhaustive, we mention the multiscale finite element method (MsFEM) by Hou et al. [27], the two-scale FEM developed by Matache and Schwab [31], the sparse FEM proposed by Hoang and Schwab [26], the multigrid homogenization method proposed by Neuss et al. [35] and the heterogeneous multiscale method (HMM) introduced by E and Engquist [24] (see also the review [7] for additional references). We also notice that there is a huge literature in the structural mechanics and engineering communities concerned with micro–macro methods for multiscale PDEs. Again without being exhaustive, we mention Yu and Fish [44], Terada and Kikuchi [41], Kouznetsova et al. [29] and Miehe et al. [32].

Despite this flourishing activity around the design and control of multiscale solvers, rigorous analysis of adaptive methods for FEMs applied to multiscale problems has only rarely been addressed and is still an underdeveloped research area. It is fair to say that success in application of many if not all the aforementioned multiscale methods will also depend on efficient and reliable adaptive strategies. As we will see throughout this work, the savings in computational cost achieved when using adaptive methods are even more dramatic than in problems featuring a single scale as costly microscale computations can be avoided or re-used.

* Corresponding author.
E-mail addresses: assyr.abdulle@epfl.ch (A. Abdulle), achim.nonnenmacher@epfl.ch (A. Nonnenmacher).

0045-7825/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2010.06.012

In this paper we present an a posteriori error analysis for the so-called finite element heterogeneous multiscale method (FE-HMM). A priori error analysis for this method for elliptic problems has been obtained in [25,2] (semi-discrete analysis) and in [1,5,4] (fully discrete analysis). Discontinuous Galerkin FEM has been studied in [6,10] and problems in elasticity in [3]. We refer to [7] for an extensive review. Concerning a posteriori error analysis, we note that first results for the FE-HMM have been obtained in [39]. We briefly put in perspective the results obtained in [39] with the results of the present paper which are obtained in a completely different manner by using techniques much closer to the standard techniques for residual based a posteriori error analysis [11,42]. The analysis performed in [39] relies on a reformulation of the FE-HMM in a two-scale framework [36]. In this framework, one adds the microscopic variable as a supplementary variable, doubling the size of the limiting problem. The reformulation of the FE-HMM in this framework is based on a tensor product FEM with quadrature in the slow variable. As a consequence, the a posteriori estimates in [39] are obtained in a two-scale norm over $\Omega \times Y$ (here Ω is the physical domain and Y the domain of the microscopic variable) and *not in a norm related to the physical domain*. Deriving optimal a priori or a posteriori error estimates in the energy norm for the physical domain from the results obtained in [39] are not straightforward and such results have not yet been obtained. In our approach the a posteriori error estimates for the upper and lower bounds are derived in the energy norm of the physical domain. The analysis of [39] is also restricted to the class of homogenization problems with tensor given in an explicit two-scale form, i.e., $a^\varepsilon = a(x,x/\varepsilon)$ and periodicity in the fast variables x/ε (see [39, Section 3]). We do not need an explicit decomposition and our analysis applies to general (non-periodic) tensors (though it involves a data approximation error which can only be explicitly estimated with additional spatial assumptions for the fast variable such as periodicity or random homogeneity). Finally, the analysis in [39] relies on the knowledge (a priori) of the exact periodicity of the problem as the sampling domains are assumed to span exactly one period in each spatial direction (see [39, Remark 3.6]) while our estimates are derived and hold for general sampling domains of size $\delta > \varepsilon$. We close this comparison by noting that in the special case of a periodic tensor $a^\varepsilon = a(x,x/\varepsilon)$ and assuming exact micro-problems we recover the classical a posteriori analysis results that one would obtain by applying the available theory [42,11,18] to the homogenized problem. Thus, our estimates seem to be consistent with usual adaptive procedures for the macroproblem.

The outline of this paper is as follows. In Section 2 we describe our model problem and recall briefly the classical adaptivity theory. In Section 3 we recall the FE-HMM, the numerical method for which we will provide the a posteriori error analysis. Our main results consisting of upper and lower a posteriori bounds are stated in Section 4 and the full analysis is performed in Section 5. We illustrate in Section 6 the a posteriori analysis by a series of numerical experiments quantifying the efficiency and reliability of our bounds and conclude with some remarks on future work based on the analysis presented here. The results presented in this paper have been announced in [9] (without proofs). In this paper we give the detailed proofs of our a posteriori error estimates for a more general setting than considered in [9] and many additional numerical experiments.

Notation. In what follows, $C > 0$ denotes a generic constant, independent of ε, whose value can change at any occurrence but depends only on the quantities which are indicated explicitly. For $r = (r_1, \dots, r_d) \in \mathbb{N}^d$, we denote $|r| = r_1 + \cdots + r_d$, $D^r = \partial_1^{r_1} \cdots \partial_d^{r_d}$. We will consider the usual Sobolev space $H^1(\Omega) = \{u \in L^2(\Omega); D^r u \in L^2(\Omega), |r| \leqslant 1\}$, with norm $\|u\|_{H^1(\Omega)} = (\sum_{|r|\leqslant 1} \|D^r u\|^2_{L^2(\Omega)})^{1/2}$. We will also consider $H^1_0(\Omega)$ the closure of $C^\infty_0(\Omega)$ for the $\|\cdot\|_{H^1(\Omega)}$ norm and the spaces $W^{l,\infty}(\Omega) = \{u \in L^\infty(\Omega); D^r u \in L^\infty(\Omega), |r| \leqslant l\}$. For the unit cube $Y = (0,1)^d$, we will consider $W^1_{per}(Y) = \{v \in H^1_{per}(Y); \int_Y v\,dx = 0\}$, where $H^1_{per}(Y)$ is defined as the closure of $C^\infty_{per}(Y)$ (the subset of $C^\infty(\mathbb{R}^d)$ of periodic functions in Y). Finally, we will use the Frobenius matrix norm $\|a\|_F := \sqrt{\sum_i \sum_j |a_{ij}|^2}$.

2. Model problem and classical adaptivity theory

In this section we briefly discuss the class of multiscale problems that we consider in this paper and review classical coarse-graining techniques such as homogenization. We also discuss the standard adaptive strategy used for (single scale) elliptic problems discretized by the FEM.

2.1. Model problem and homogenization

We consider the second-order elliptic problem in the domain $\Omega \subset \mathbb{R}^d$

$$\begin{aligned} &-\nabla\cdot(a^\varepsilon \nabla u^\varepsilon) = f \quad \text{in } \Omega,\\ &u^\varepsilon = g_D \quad \text{on } \partial\Omega_D,\\ &(a^\varepsilon \nabla u^\varepsilon)\cdot n = g_N \quad \text{on } \partial\Omega_N, \end{aligned} \tag{1}$$

where $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, with Dirichlet boundary conditions $g_D \in H^{1/2}(\partial\Omega)$ imposed on Ω_D (assumed to be of positive measure), and Neumann boundary conditions $g_N \in H^{-1/2}(\partial\Omega)$ imposed on Ω_N.

We assume that the family of tensors, indexed by ε, are symmetric, satisfy $a^\varepsilon(x) \in (L^\infty(\Omega))^{d\times d}$ and are uniformly elliptic and bounded, i.e.,

$$\exists \lambda, \Lambda > 0 \text{ such that } \lambda|\xi|^2 \leqslant a^\varepsilon(x)\xi\cdot\xi \leqslant \Lambda|\xi|^2 \quad \forall \xi \in \mathbb{R}^d \text{ and } \forall \varepsilon. \tag{2}$$

The multiscale nature of the tensor $a^\varepsilon(x)$ is emphasized by the superscript ε. We will assume $f \in L^2(\Omega)$ (our results are also valid for $f \in H^{-1}(\Omega)$). By applying the Lax–Milgram theorem to the weak form of (1), we obtain a family of solutions $\{u^\varepsilon\}$, bounded in $H^1_0(\Omega)$. Without making any further assumption on the spatial structure of the tensor $a^\varepsilon(x)$ (e.g., the form of the heterogeneities) using the notions of G-convergence introduced by De Giorgi and Spagnolo [23],[1] one can show that there exists a symmetric tensor $a^0(x)$ and a subsequence of $\{u^\varepsilon\}$ which weakly converges to an element $u^0 \in H^1_0(\Omega)$, where u^0 is the solution of the so-called homogenized or upscaled problem

$$\begin{aligned} &-\nabla\cdot(a^0 \nabla u^0) = f \quad \text{in } \Omega,\\ &u^0 = g_D \quad \text{on } \partial\Omega_D,\\ &(a^0 \nabla u^0)\cdot n = g_N \quad \text{on } \partial\Omega_N. \end{aligned} \tag{3}$$

The homogenized tensor $a^0(x)$ can be shown to be symmetric and to satisfy $\lambda|\xi|^2 \leqslant a^0(x)\xi\cdot\xi \leqslant \Lambda|\xi|^2$, $\forall \xi \in \mathbb{R}^d$ and thus, (3) has a unique solution. If some additional spatial structure is assumed on $a^\varepsilon(x)$, (for example if $a^\varepsilon(x) = a(x,x/\varepsilon)$ and is periodic in its second argument), then classical results in homogenization theory (see [28,22]) show that the whole sequence $\{u^\varepsilon\}$ weakly converges to an element $u^0 \in H^1_0(\Omega)$. In this case, we can characterize the homogenized tensor $a^0(x)$ at $x \in \Omega$ by an average (integral) involving d boundary value problems (the cell problems). Notice that in this case, one has in general infinitely many cell problems to solve in order to compute the matrix-valued function $a^0(x)$ in Ω.

[1] A generalization of this notion, called H-convergence has been introduced by Murat and Tartar [34].

2.2. Adaptive FEM for single scale problems

Essential ingredients for adaptivity are a posteriori estimators. These are computable quantities depending on the actual FE solution and data that give information on the distribution of the error over the FE mesh and allow to refine or coarsen the triangulation where appropriate. As mentioned in the introduction, the ultimate goal is to find a mesh which equidistributes the approximation error over all elements and to optimize the computational effort. Among the various approaches for a posteriori error estimators for *single scale* elliptic problems we mention estimators based on local residuals, on local boundary value problems or on local superconvergent properties (see [42,11] for an overview and references). More recently, error estimators in *quantities of interest* have been introduced (see [11, Chapter 8]), where the mesh is refined in terms of a certain (physical) quantity $Q(u)$ relevant for a specific design decision. We also mention the recent effort concerning model adaptivity for multiphysics problems (see [38] and the references therein).

In this work we use a residual-based estimator as presented in [42,11,33,37]. For such estimators, one defines on every element K of a triangulation a computable local error indicator $\eta_H(K)$ consisting of an error residual and a jump residual. The sum over all elements of $\eta_H(K)^2$ gives the global error indicator $\eta_H(\Omega)^2$, which has to be related to the error $u-u^H$ in a suitable manner. An upper bound of the error in terms of $\eta_H(\Omega)$ ensures the effectivity of the estimate, while a lower bound ensures its efficiency. By using the local error indicator, one can identify and refine those elements which contribute the most to the global error. The marked elements are then refined in order to equidistribute the error among the FE mesh.

The overall procedure can be described as follows. Starting with a conforming mesh $\mathcal{T}_H$ we iteratively apply the following procedure in order to adapt the mesh

SOLVE → ESTIMATE → MARK → REFINE.

In the SOLVE procedure we solve the PDE to obtain the numerical solution u^H. Based on u^H we estimate the error indicators $\eta_H(K)$ for every element in the ESTIMATE procedure. We mark some elements $\widetilde{\mathcal{T}_H}$ of the triangulation $\mathcal{T}_H$ in the MARK procedure with the aim of equi-distributing the global error among the elements. Finally, in the REFINE procedure, we refine the marked elements together with some neighboring elements in order to preserve the mesh conformity.

3. The FE-HMM

In this section we briefly recall the FE-HMM, the numerical method considered in this paper for the numerical homogenization of (1). In a finite element context, the FE-HMM was analyzed in [1,25,2,5]. The method is based on a macroFEM defined upon quadrature points and microFEMs on sampling domains centered around the macroquadrature points. As input data, the method takes only the multiscale data (e.g., a^ε) and it does not rely on precomputed homogenized (upscaled) coefficents (e.g., $a^0(x)$).

Macrofinite element space. We consider

$$V^p(\Omega,\mathcal{T}_H)=\left\{v^H\in H_0^1(\Omega);\ \ v^H|_K\in\mathcal{R}^p(K),\ \forall K\in\mathcal{T}_H\right\},\tag{4}$$

with macroelements $K\in\mathcal{T}_H$, where $\mathcal{T}_H$ (a macro partition of Ω) is assumed to be shape regular. The diameter of an element $K\in\mathcal{T}_H$ is denoted by H_K and we define $H=\max_{K\in\mathcal{T}_H}H_K$. Here $\mathcal{R}^p=\mathcal{P}^p$ is the space of piecewise polynomials on the element K of *total* degree p, if K is a simplicial element (triangle if $d=2$, tetrahedron if $d=3$). If K is a quadrilateral element (quadrilateral if $d=2$, hexahedron if $d=3$), then $\mathcal{R}^p=\mathcal{Q}^p$ is the space of piecewise polynomials on the element K of degree p *in each variable*.

We emphasize that H, the size of the macrotriangulation, is allowed to be much larger than ε.

Quadrature formulas. Within each macroelement $K\in\mathcal{T}_H$ we consider a quadrature formula $\{x_{K_\ell},\omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$ where x_{K_ℓ} are the macroquadrature points and ω_{K_ℓ} are the quadrature weights. We further consider sampling domains $K_{\delta_\ell}=x_{K_\ell}+\delta I$, where $I=(-1/2,1/2)^d$ and $\delta\geqslant\varepsilon$, which are centered at the macroquadrature points. To have optimal convergence results, we need to choose the quadrature formula $\{x_{K_\ell},\omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$ appropriately on K (we refer to [7] for details).

Macrobilinear form. For $v^H,\ w^H\in V^p(\Omega,\mathcal{T}_H)$ we define

$$B(v^H,w^H)=\sum_{K\in\mathcal{T}_H}\sum_{\ell=1}^{\mathcal{L}}\frac{\omega_{K_\ell}}{|K_{\delta_\ell}|}\int_{K_{\delta_\ell}}a^\varepsilon(x)\nabla v^h_{K_\ell}\cdot\nabla w^h_{K_\ell}\,dx,\tag{5}$$

where $v^h_{K_\ell}$, $w^h_{K_\ell}$ are microfunctions defined in (6) on the sampling domains K_{δ_ℓ}. As the integrals in (5) are defined on K_{δ_ℓ} instead of K, we need to multiply the contribution of the sampling domains by the appropriate weight, the factor $1/|K_{\delta_\ell}|$.

Microsolver. To assemble the macrostiffness matrix, the computation of the microfunctions $v^h_{K_\ell}$ (and $w^h_{K_\ell}$) is needed in the sampling domains K_{δ_ℓ}, $\ell=1,\dots,\mathcal{L}$ of each macroelement K.

The microproblems are computed as follows: find $v^h_{K_\ell}$ such that $(v^h_{K_\ell}-v^H_{\mathrm{lin},K_\ell})\in S^q(K_{\delta_\ell},\mathcal{T}_h)$ and

$$\int_{K_{\delta_\ell}}a^\varepsilon(x)\nabla v^h_{K_\ell}\cdot\nabla z^h\,dx=0\quad\forall z^h\in S^q(K_{\delta_\ell},\mathcal{T}_h),\tag{6}$$

where

$$v^H_{\mathrm{lin},K_\ell}(x)=v(x_{K_\ell})^H+(x-x_{K_\ell})\cdot\nabla v^H(x_{K_\ell})\tag{7}$$

is a linearization of the macrofunction v^H at the integration point x_{K_ℓ} (see [7,25] for details). Here, $S^q(K_{\delta_\ell},\mathcal{T}_h)$ is defined as

$$S^q(K_{\delta_\ell},\mathcal{T}_h)=\{z^h\in W(K_{\delta_\ell});\ z^h|_T\in\mathcal{R}^q(T),\ T\in\mathcal{T}_h\},\tag{8}$$

where $W(K_{\delta_\ell})$ determines the boundary conditions for the coupling condition used for computing the microfunctions $v^h_{K_\ell}$ (or $w^h_{K_\ell}$). Here, $\mathcal{T}_h$ denotes a (micro) partition of the sampling domain K_{δ_ℓ}. The diameter of an element $T\in\mathcal{T}_h$ is denoted by h_T and $h=\max_{T\in\mathcal{T}_h}h_T$. Various choices are possible for the coupling conditions and we will consider

$$W(K_{\delta_\ell})=W^1_{per}(K_{\delta_\ell})\tag{9}$$

or

$$W(K_{\delta_\ell})=H^1_0(K_{\delta_\ell}).\tag{10}$$

We will say that the boundary conditions for (6) are given by (9) if $S^q(K_{\delta_\ell},\mathcal{T}_h)\subset W^1_{per}(K_{\delta_\ell})$ (periodic coupling), and that they are given by (10) if $S^q(K_{\delta_\ell},\mathcal{T}_h)\subset H^1_0(K_{\delta_\ell})$ (Dirichlet coupling).

Variational problem. To determine the macrosolution of the FE-HMM we finally solve the following macrovariational problem: find $u^H\in V^p(\Omega,\mathcal{T}_H)$ such that

$$B(u^H,v^H)=\int_\Omega f v^H\,dx+\int_{\partial\Omega_N}g_N v^H\,dx-B(g_D,v^H)\quad\forall v^H\in V^p(\Omega,\mathcal{T}_H).\tag{11}$$

3.1. Sources of error in the HMM

The first objective of this method is to numerically capture the effective (homogenized) solution u^0 of (3). As the numerical

method depends on macro- and microFEMs, and in turn on macro- and micromeshes, the convergence results will also depend on these quantities [1]. In addition, it will also depend on a so-called modeling error, quantifying the upscaling procedure itself and depending on the coupling condition, the sampling domain size and the spatial structure of the oscillating tensor a^ε.

In what follows we provide a list of all the sources of errors in the numerical scheme which can be obtained from an a priori estimate [1,7,25]. Following the framework of [7], we have the following decomposition of the error

$$\|u^0 - u^H\| \leqslant \underbrace{\|u^0 - u^{0,H}\|}_{err_{mac}} + \underbrace{\|u^{0,H} - \bar{u}^H\|}_{err_{mod}} + \underbrace{\|\bar{u}^H - u^H\|}_{err_{mic}},$$

where err_{mac}, err_{mod} and err_{mic} are macro, modeling and microerror, respectively. Here, u^0 is the solution of (11), $\bar{u}^H$ is the FE-HMM solution of (11) using exact microsolution (i.e., we solve (6) in $W(K_{\delta_\ell})$) and $u^{0,H}$ is the FEM solution of the (usually unknown) problem (3) in the space $V^1(\Omega, \mathcal{T}_H)$.

We have the following estimates for the various components (for simplicity, we describe them for piecewise linear macro- and microFEs only).

- err_{mac} is the error from a standard FE approximation (with numerical quadrature) of (3) in $V^1(\Omega, \mathcal{T}_H)$. We have $err_{mac} \leqslant CH$ in the H^1 norm and $err_{mac} \leqslant CH^2$ in the L^2 norm.
- err_{mic} is the error from the FE approximation of the microproblems in $S^1_h(K_\delta, \mathcal{T}_h)$. The microerror contributes to the global error through the macrobilinear form. We get (under appropriate regularity of a^ε and appropriate boundary conditions) $err_{mic} \leqslant C\left(\frac{h}{\varepsilon}\right)^2$ for both the H^1 and L^2 norm (see [10] for details).
- err_{mod} is the error introduced by the upscaling procedure and the coupling of macro- and microFEMs. This error depends on the structure of the tensor a^ε, the boundary conditions of the microproblem (i.e., the coupling of the micro- to the macroFE spaces) and the sampling domain size δ, but is independent of H and h. In general this error reads $err_{mod} \leqslant C\sup_{K\in\mathcal{T}_H, x_{j,K}\in K} \|a^0(x_{j,K}) - \bar{a}^0_K\|_F$ where $\bar{a}^0_K$ is defined similarly as in (14), but relying on exact micro functions. In case of a periodic tensor $a^\varepsilon = a(x, x/\varepsilon)$, periodic coupling, sampling domains covering an integer number of periods and "collocation" in the slow variable of the oscillating tensor, $err_{mod} = 0$.

These various errors are further described in Section 5.5.2 and in [7].

3.1.1. Fine scale error

In the L^2 norm one can estimate $\|u^\varepsilon - u^H\|_{L^2(\Omega)}$ by using the triangle inequality and

$$\|u^\varepsilon - u^0\|_{L^2(\Omega)} \leqslant C\varepsilon.$$

For the H^1 norm, u^H does not converge to u^ε as already u^0 does not converge to u^ε. Indeed, u^0 (or u^H) does not capture the microoscillation of u^ε resulting in an $\mathcal{O}(1)$ mismatch when estimating their gradients. One can nevertheless post-process the solution u^H and the microsolutions u^h_K (available in the sampling domains $K_\delta \subset K$) using ideas from the construction of correctors in homogenization theory [15]. The post-processed FE-HMM solution $u^{H,\varepsilon}$ satisfies [1,7]

$$\|u^\varepsilon - u^{H,\varepsilon}\|_{\overline{H}^1(\Omega)} \leqslant C\left(H + \sqrt{\varepsilon} + \frac{h}{\varepsilon}\right),$$

where $\overline{H}^1(\Omega)$ is a broken norm defined by $\|u\|_{\overline{H}^1(\Omega)} = \left(\sum_{K\in\mathcal{T}_H} \|\nabla u\|^2_{L^2(K)}\right)^{1/2}$ as the reconstructed solution may have jumps on the interfaces of two neighboring elements. Notice the additional factor $\sqrt{\varepsilon}$ coming from a boundary layer as the corrector does not satisfy the right boundary conditions on $\partial\Omega$.

4. Main results

The goal is to adapt the macromesh according to potential singularities which may be caused by reentrant corners or high contrast in macroscopic coefficients. Localized "macroscopic" residuals determine how the macroscopic mesh has to be adapted. We have to recover these residuals from suitably averaged microscopic quantities as they are not readily available. We will see that the overhead for deriving these macroscopic residuals is minimal as they are based on microscopic solutions already required for the computation of the macrostiffness matrix.

For simplicity, piecewise linear macroFE (simplicial elements) will be used. As a consequence, we will use a quadrature formula with $\mathcal{L} = 1$ and integration node $x_{K_\ell} = x_K$ localized at the barycenter of the macroelement K, and a weight $\omega_{K_\ell} = \omega_K = |K|$ (see Section 3). Moreover, we choose piecewise linear or bilinear microFE, i.e., $q = 1$ in (8) and $\mathcal{R}^1 = \mathcal{P}^1$ or $\mathcal{Q}^1$.

Furthermore, we consider in our analysis only homogeneous Dirichlet boundary conditions $u_D = 0$ and $u_N = \emptyset$, but emphasize that the a posteriori estimates can be derived for more general boundary conditions (such as Neumann or Robin) following the lines of the results presented in this article. The a posteriori error analysis could be generalized for higher order FEs or quadrilateral macroFEs using the ideas in [10].

Let $\mathcal{T}_H$ denote a conformal mesh and let $\mathcal{E}_H$ be the set of interfaces. We label the two elements sharing an interface $e \in \mathcal{E}_H$ as K^+ and K^-. We consider the microfunctions $u^h_{K^+}$ and $u^h_{K^-}$, solutions of the microproblems (6), which correspond to the two sampling domains K^+_δ and K^-_δ of the elements K^+ and K^-, respectively. These microfunctions are constrained by the macrosolution $u^H \in V^1(\Omega, \mathcal{T}_H)$ of problem (11). We then introduce the following jump of *multiscale fluxes*

$$[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e := \begin{cases} \left(\frac{1}{|K^+_\delta|}\int_{K^+_\delta} a^\varepsilon(x)\nabla u^h_{K^+}\,dx - \frac{1}{|K^-_\delta|}\int_{K^-_\delta} a^\varepsilon(x)\nabla u^h_{K^-}\,dx\right)\cdot n_e & \text{for } e \notin \partial\Omega, \\ 0 & \text{for } e \in \partial\Omega, \end{cases} \tag{12}$$

where the unit outward normal n_e is chosen to be $n_e = n^+$. We omit the index K_δ for the microsolutions u^h in $[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e$ as the jump over e involves two sampling domains in adjacent elements. The multiscale fluxes are the building blocks in the derivation of our estimates and were first introduced in the context of multiscale discontinuous Galerkin methods [6,10].

For each vector $\mathbf{e}_i \in \mathbb{R}^d$, $i = 1, \ldots, d$ we consider $\psi^{i,h}_{K_\delta} \in S^1_h(K_\delta, \mathcal{T}_h)$, the solution of the problem

$$\int_{K_\delta} a^\varepsilon(x)\nabla\psi^{i,h}_{K_\delta}\cdot\nabla z^h\,dx = -\int_{K_\delta} a^\varepsilon(x)\mathbf{e}_i\cdot\nabla z^h\,dx \quad \forall z^h \in S^1_h(K_\delta, \mathcal{T}_h). \tag{13}$$

It will be convenient for our analysis to introduce a numerically homogenized tensor (see [10]). We define this tensor a^0_K, constant on each macroelement K, by

$$a^0_K = \frac{1}{|K_\delta|}\int_{K_\delta} a^\varepsilon(x)\left(I + J^T_{\psi^h_{K_\delta}(x)}\right)dx, \tag{14}$$

where $J_{\psi^h_{K_\delta}(x)}$ is a $d\times d$ matrix whose entries are given by $\left(J_{\psi^h_{K_\delta}(x)}\right)_{ij} = \left(\partial\psi^{i,h}_{K_\delta}\right)/(\partial x_j)$. We emphasize that the above tensor is only used as a tool in the derivation of the a posteriori error bounds and is never used for the computation of our error indicators.

Definition 1. Let f^H be a piecewise constant approximation of f. Then the *local error indicator* $\eta_H(K)$ on an element K is defined by

$$\eta_H(K)^2 := H_K^2 \|f^H\|^2_{L^2(K)} + \frac{1}{2}\sum_{e\subset\partial K} H_e \left\| [\![\overline{a^\varepsilon \nabla u^h}]\!]_e \right\|^2_{L^2(e)}.$$

We furthermore define the *data approximation error* $\xi_H(K)$ on an element K by

$$\xi_H(K)^2 := H_K^2 \|f^H - f\|^2_{L^2(K)} + \left\| (a_K^0 - a^0(x)) \nabla u^H \right\|^2_{L^2(K)},$$

where $a^0(x)$ is the unknown homogenized tensor of problem (3).

We will sometimes consider the indicators and data approximation terms on a subset $\omega = K_{i_1} \bigcup K_{i_2} \bigcup \cdots \bigcup K_{i_n}$, $K_{i_j} \in \mathcal{T}_H$ of the domain Ω. In this case, we denote the expression obtained by summing the above quantities over all elements $K \in \omega$ by $\eta_H(\omega)^2$ and $\xi_H(\omega)^2$.

Our first result establishes an a posteriori upper bound for the error between the macroscopic FE-HMM solution u^H and the homogenized solution u^0.

Theorem 2 (A posteriori upper bound). *There exists a constant $C > 0$ depending only on the shape regularity constant γ, the coercivity and continuity bound (2) , the dimension d and the domain Ω such that*

$$\|u^0 - u^H\|^2_{H^1(\Omega)} \leqslant C\left(\eta_H(\Omega)^2 + \xi_H(\Omega)^2\right).$$

The next result gives an a posteriori lower bound.

Theorem 3 (A posteriori lower bound). *There exists a constant $C > 0$ depending only on the shape regularity constant γ, the coercivity and continuity bound (2) and the dimension d such that*

$$\eta_H(K)^2 \leqslant C\left(\|u^0 - u^H\|^2_{H^1(\omega_K)} + \xi_H(\omega_K)^2\right),$$

where the domain ω_K consists of all elements sharing at least one side with K.

Micro–macro refinement. The above two theorems for the a posteriori lower and upper bound do not require any structure assumption on the oscillating tensor (such as periodicity) and only minimal assumptions on regularity. As we assume singularities in the macroscale, we do not consider explicit a posteriori estimates for the microproblem (6). As the microsampling domains have simple geometries (typically squares or cubes), singularities could only arise due to singularities in the microscale of the conductivity tensor. In that case, standard a posteriori methods could be used to refine the micromeshes in a non-uniform way.

We emphasize that the error indicator η_H *does* depend on the microsolutions and hence on the micromesh. Thus, a criterion is needed to determine an appropriate size of the micromesh as we refine the macromesh through our adaptive procedure. Such a criterion can be deduced from the following theorem.

Theorem 4 (Micro–macrorefinement coupling). *Assume that (31) and (32) hold. Assume further that the cell problem (6) is solved with periodic boundary conditions if a^ε is periodic and $\delta/\varepsilon \in \mathbb{N}^*$ and solved with Dirichlet boundary conditions otherwise. Then*

$$\sup_{x\in K} \left\|a_K^0 - a^0(x)\right\|_F \leqslant C\left(H_K + \left(\frac{h}{\varepsilon}\right)^2\right) + err_{mod}, \tag{15}$$

where C is independent of H_K, h, ε, and err_{mod} is independent of H_K, h.

Remark 5. From estimate (15) we deduce that in order to minimize the error originating from the microFEM, we have to refine the micromesh in each sampling domain K_δ as $\frac{h}{\varepsilon} \propto \sqrt{H_K}$. Here H_K is the size of the macroelement K of the mesh $\mathcal{T}_H$ obtained by the Algorithm 1 described below. Please note that the size h of the micromesh may vary in different sampling domains. Thus, $h = h(K)$ is a function of $K \in \mathcal{T}_H$. To simplify the notation, we will remove the dependency of h on K.

Comparison with single scale results. Our upper and lower bounds stated above are consistent with the usual (single scale) residual-based a posteriori estimates. Indeed, suppose $a^\varepsilon = a(\frac{x}{\varepsilon})$, that exact microfunctions are used in (6) and that $\delta/\varepsilon \in \mathbb{N}^*$. Then $\left\|a_K^0 - a^0\right\|_F \equiv 0$ (see Section 5.5.2) and we recover the usual residual-based indicator and data estimator [42,18].

4.1. Algorithm

The adaptive algorithm for the FE-HMM follows the adaptive algorithm for standard FEM. It consists of loops of the form

SOLVE → ESTIMATE → MARK → REFINE

in order to generate the new, refined computational grid. However, due to the multiscale nature of the problem, we need to modify the procedure accordingly as presented in the following algorithm.

Algorithm 1 (*Adaptive FE-HMM*).

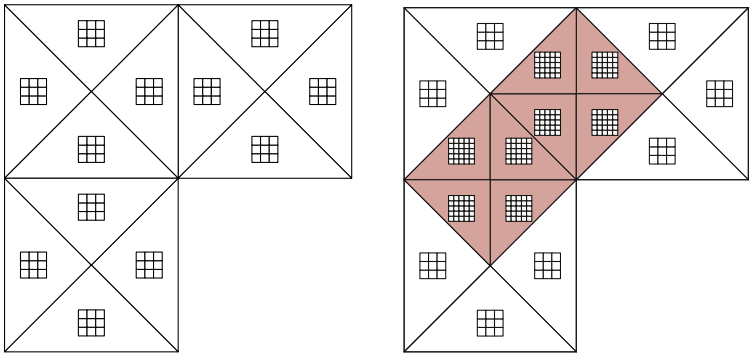
- **Solve**. For the macro- and micromeshes obtained by REFINE, compute the microsolutions (only for the refined macroelements) and the macrosolution u^H of (11) and compute and store the multiscale jumps $[\![\overline{a^\varepsilon(x)\nabla\varphi_{i,K}^h}]\!]_e$ (based on the macroFE basis functions) for the refined elements during the macroassembly process.
- **Estimate**. Reconstruct the full multiscale jumps $[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e$ using the macrosolution u^H of SOLVE and estimate the error by computing the indicators $\eta_H(K)$ for all $K \in \mathcal{T}_H$.
- **Mark**. Identify a subset $\widetilde{\mathcal{T}}_H$ of $\mathcal{T}_H$ based on the indicators $\eta_H(K)$ following Dörfler's bulk-chasing strategy (Marking Strategy E, see [42, Chapter 4.1]).
- **Refine**. Refine the elements in the subset $\widetilde{\mathcal{T}}_H$ and some neighboring elements in order to preserve mesh conformity. Update the micromesh in the sampling domains of the refined macroelements according to $h \propto \sqrt{H_K}\varepsilon$ (see Remark 5).

We discuss two important details for the above algorithm, namely the carry-over of information and the computation of the multiscale flux.

Carry-over of information. In contrast to standard (single scale) adaptive FE methods, in the adaptive FE-HMM most of the computation time is used for solving the microproblems to obtain the entries of the macro stiffness matrix. A fundamental feature of an efficient implementation of an adaptive FE-HMM should thus be the carry-over of reusable micro data from one iteration to the next.

In particular, this means that for every element we store the contributions $A_{ij} = B\left(\varphi_{i,K}^H, \varphi_{j,K}^H\right)$ ($\varphi_{i,K}^H$ are the macrobasis functions) of the microproblems to the macrostiffness matrix and the components of the corresponding multiscale flux $[\![\overline{a^\varepsilon(x)\nabla\varphi_{i,K}^h}]\!]_e$. This is done with a small memory overhead in every iteration and we compute new data A_{ij} and $[\![\overline{a^\varepsilon(x)\nabla\varphi_{i,K}^h}]\!]_e$ only for those elements which are marked for refinement and carry-over the existing data A_{ij} and $[\![\overline{a^\varepsilon(x)\nabla\varphi_{i,K}^h}]\!]_e$ for the remaining, unrefined macrotriangles to the next iteration (see Fig. 1 for an illustration). In the numerical experiments in Section 6 we illustrate the amount of work that can be saved following this strategy.

Computation of the multiscale flux. Let $\left\{\varphi_{i,K}^H\right\}$ be the basis functions of $V^1(\Omega, \mathcal{T}_H)$ and let $\varphi_{i,K}^h$ be the microsolution of (6)



(a) Iteration k (b) Iteration $k+1$

Fig. 1. Adaptive mesh shown for two iterations. In Fig. (b), for all macroelements shown in white, the multiscale flux $[\![\overline{a^\varepsilon(x)\nabla\varphi^h_{i,K}}]\!]_e$ and the contributions A_{ij} to the macrostiffness matrix can be carried over from the previous iteration and re-used, whereas for all macroelements shown in red, new solutions of the microproblems and corresponding multiscale fluxes must be computed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

constrained by $\varphi^H_{i,K}$, i.e., $\left(\varphi^h_{i,K}-\varphi^H_{i,K}\right)\in S^1(K_\delta,\mathcal{T}_H)$, where $K_\delta\subset K$. In SOLVE, we compute for every element K

$$\varrho^h_{i,K}:=\frac{1}{|K_\delta|}\int_{K_\delta}a^\varepsilon(x)\nabla\varphi^h_{i,K}\,dx$$

right after computing the solutions $\varphi^h_{i,K}$ of the microproblem (6) and store the three corresponding $\varrho^h_{i,K}$ for later use (in the case of two dimensions when using macrotriangles this represents a 2×3 matrix in each macroelement K corresponding to the two components of $\varrho^h_{i,K}$ for each macrobasis function $\varphi^H_{i,K}$, $i=1,\dots,3$). The advantage of computing this quantity in the SOLVE instead of the REFINE step is that we do not need to store the full microsolution of each macroelement (if we are not interested in a reconstructed full solution, this significantly reduces the memory requirement).

Denote by $u^H_K(x)=\sum_{i=1}^3\alpha_i\varphi^H_{i,K}(x)$ the representation of u^H in the element K with respect to the nodal basis. As the microsolution $u^h_{i,K}(x)$ corresponding to u^H_K is linear with respect to $u^H_K(x)$, we can reconstruct $u^h_K(x)=\sum_{i=1}^3\alpha_i\varphi^h_{i,K}(x)=\sum_{i=1}^3u^H(x_i)\varphi^h_{i,K}(x)$ where x_i are the nodes of the macrotriangle K. In the ESTIMATE step, where the macrosolution u^H is known from the previous SOLVE step, we reconstruct

$$\left(\frac{1}{|K_\delta|}\int_{K_\delta}a^\varepsilon(x)\nabla u^h_K\,dx\right)\cdot n_e=\sum_{i=1}^3u^H(x_i)\varrho^h_{i,K}\cdot n_e$$

and assemble $[\![\overline{a^\varepsilon(x)\nabla u^h_{i,K}}]\!]_e$ using the previously stored information. As can be seen from the aforementioned procedure, the computation of $[\![\overline{a^\varepsilon(x)\nabla u^h_{i,K}}]\!]_e$ can be done with minimal memory overhead and very small extra computational cost.

Remark 6 (*Coarsening*). We have not considered a coarsening strategy for two reasons. First, for linear elliptic problems, a quasi-optimal mesh can usually be obtained without the need of coarsening strategies (see [40, Chapter 1.5.3]). Second, due to the macro–micro coupling, new extra microproblems must be solved when elements are coarsened, whereas information can be re-used if we omit coarsening the macroelements.

5. Proof of the main results

5.1. Interpolation, trace and inverse estimates

Before proving the estimates for the upper and lower bound, we recall some interpolation, trace and inverse estimates that we will need for our analysis.

Clément interpolation operator (see [21]).

Let $I^H:H^1(\Omega)\to V^1(\Omega,\mathcal{T}_H)$ be the Clément interpolation operator. This is a linear operator with the property that for all $v\in H^1(\Omega)$ and $K\in\mathcal{T}_H$

$$\|v-I^Hv\|_{L^2(K)}\leqslant CH_K\|\nabla v\|_{L^2(N(K))}\tag{16}$$

and

$$\|\nabla(v-I^Hv)\|_{L^2(K)}\leqslant C\|\nabla v\|_{L^2(N(K))},\tag{17}$$

where $N(K)$ is a neighborhood of K that consists of all elements of $\mathcal{T}_H$ which have a non-empty intersection with K.

Trace inequality (see [12, Theorem 3.10]).

Consider an element K_e of the triangulation $\mathcal{T}_H$ with side $e\in\mathcal{E}_H$. Then, for $v\in H^1(K_e)$ we have

$$\|v\|_{L^2(e)}\leqslant CH_e^{1/2}\|\nabla v\|_{L^2(K_e)}+CH_e^{-1/2}\|v\|_{L^2(K_e)}.\tag{18}$$

Inverse inequality (see for example [20]).

For $v^H\in V^p(\Omega,\mathcal{T}_H)$ we have

$$\|\nabla v^H\|_{L^2(K)}\leqslant CH_K^{-1}\|v^H\|_{L^2(K)}.\tag{19}$$

Remark 7. The combination of the Clément interpolation estimates and the trace inequality yields for an element K_e with side $e\in\mathcal{E}_H$

$$\begin{aligned}\|v-I^Hv\|_{L^2(e)}&\leqslant CH_e^{1/2}\|\nabla(v-I^Hv)\|_{L^2(K_e)}+CH_e^{-1/2}\|v-I^Hv\|_{L^2(K_e)}\\&\leqslant CH_K^{1/2}\|\nabla v\|_{L^2(N(K_e))}.\end{aligned}\tag{20}$$

5.2. Error representation formula

The representation formula (22) in Lemma 9 is the central tool to derive our a posteriori bounds, as it allows to link the bilinear form for the homogenized solution with the FE-HMM. We first prove Lemma 8, needed to derive the representation formula.

Lemma 8. *Let v^h_K,w^h_K be the solutions of (6) constrained by $v^H,w^H\in V^1(\Omega,\mathcal{T}_H)$ with boundary conditions given by (9) or (10). Then*

$$\frac{1}{|K_\delta|}\int_{K_\delta}a^\varepsilon(x)\nabla v^h_K\cdot\nabla w^h_K\,dx=\frac{1}{|K|}\int_K a^0_K\nabla v^H\cdot\nabla w^H\,dx.$$

Proof. The proof is similar to the proof of (A.1) in [2, Appendix A] or formula (63) in [7] (for the convenience of the reader we will recall it). While in the aforementioned results, a specific structure of the tensor a^ε was used, we prove the result without any assumption on a^ε (except for the positivity and ellipticity).

First, we notice that the (unique) solution of (6) can be written as

$$v_K^h(x) = v^H(x) + \sum_{i=1}^{d} \psi_{K_\delta}^{i,h}(x) \frac{\partial v^H(x)}{\partial x_i}, \tag{21}$$

where $\psi_{K_\delta}^{i,h} \in S^1(K_\delta, \mathcal{T}_h)$, $i = 1, \dots, d$ are the solutions of (13).

From this we deduce

$$\begin{aligned}
&\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla v_K^h \cdot \nabla w_K^h \, dx \\
&= \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla v_K^h \cdot \nabla \left(w^H + \sum_{i=1}^{d} \psi_{K_\delta}^{i,h}(x) \frac{\partial w^H(x)}{\partial x_i} \right) dx \\
&= \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla \left(v^H + \sum_{i=1}^{d} \psi_{K_\delta}^{i,h}(x) \frac{\partial v^H(x)}{\partial x_i} \right) \cdot \nabla w^H \, dx,
\end{aligned}$$

where we used that $\frac{\partial w^H(x)}{\partial x_i}$ is constant, $\psi_{K_\delta}^{i,h} \in S^1(K_\delta, \mathcal{T}_h)$ and Eq. (6). Recalling the definition (14) of a_K^0 we obtain

$$\begin{aligned}
&\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla v_K^h \cdot \nabla w_K^h \, dx \\
&= \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \left(I + J_{\psi^h(x)}^T \right) dx (\nabla v^H \cdot \nabla w^H) \\
&= a_K^0 \nabla v^H \cdot \nabla w^H = \frac{1}{|K|} \int_K a_K^0 \nabla v^H \cdot \nabla w^H \, dx,
\end{aligned}$$

where we used again that ∇v^H and ∇w^H are constant. □

We define the error as $e^H := u^0 - u^H$ where u^0 is the homogenized solution of (3) and u^H is the FE-HMM solution of problem (11). We shall now obtain an *error representation formula* which is crucial for the derivation of the a posteriori bounds.

Lemma 9. *For all $v \in H_0^1(\Omega)$, we have*

$$\begin{aligned}
B_0(e^H, v) = &\int_\Omega f v \, dx - \sum_{e \in \mathcal{E}} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e v \, ds \\
&+ \sum_{K \in \mathcal{T}_H} \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla v \, dx
\end{aligned} \tag{22}$$

where u^H is the solution of (11) and u_K^h are the corresponding microsolutions (6) and where the multiscale jump $[\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e$ is defined in (12).

Proof. We proceed in two steps. First, we need the following formula

$$\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla u_K^h \, dx = a_K^0 \nabla u_K^H, \tag{23}$$

which is obtained by using the expansion (21) and similar arguments as used for the proof of Lemma 8. Second, we prove that for all $v \in H_0^1(\Omega)$ we have

$$\sum_{K \in \mathcal{T}_H} \int_K a_K^0 \nabla u^H \cdot \nabla v \, dx = \sum_{e \in \mathcal{E}_H} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e v \, ds. \tag{24}$$

Integration by parts and the use of (23) gives

$$\begin{aligned}
&\sum_{K \in \mathcal{T}_H} \int_K a_K^0 \nabla u^H \cdot \nabla v \, dx \\
&= \sum_{K \in \mathcal{T}_H} \int_{\partial K} (a_K^0 \nabla u^H) \cdot n v \, ds - \underbrace{\int_K \nabla \cdot (a_K^0 \nabla u^H) v \, dx}_{=0} \\
&= \sum_{K \in \mathcal{T}_H} \int_{\partial K} \left(\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla u_K^h \, dx \right) \cdot n v \, ds \\
&= \sum_{e \in \mathcal{E}_H} \int_e \left[\left(\frac{1}{|K_\delta^+|} \int_{K_\delta^+} a^\varepsilon(x) \nabla u_{K^+}^h \, dx - \frac{1}{|K_\delta^-|} \int_{K_\delta^-} a^\varepsilon(x) \nabla u_{K^-}^h \, dx \right) \cdot n_e \right] v \, ds \\
&= \sum_{e \in \mathcal{E}_H} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e v \, ds,
\end{aligned}$$

where we used the definition (12) of the multiscale flux. Finally, we obtain the error representation formula

$$\begin{aligned}
B_0(e^H, v) &= B_0(u^0, v) - B_0(u^H, v) \\
&= \int_\Omega f v \, dx - \sum_{K \in \mathcal{T}_H} \int_K a^0(x) \nabla u^H \cdot \nabla v \, dx \\
&= \int_\Omega f v \, dx - \sum_{e \in \mathcal{E}_H} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e v \, ds \\
&\quad + \sum_{K \in \mathcal{T}_H} \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla v \, dx. \qquad \square
\end{aligned}$$

5.3. Upper bound (Proof of Theorem 2)

To proceed with the proof of Theorem 2, we consider the error representation formula (22) and choose the test function $v := e^H$. We recall that I^H denotes the Clément interpolation operator (see (16) and (17)) and f^H denotes a piecewise constant approximation of f over $\mathcal{T}_H$. By noting that

$$B(u^H, I^H e^H) = \sum_{K \in \mathcal{T}_H} \int_K f (I^H e^H) dx,$$

we obtain

$$\begin{aligned}
B_0(e^H, e^H) = &\int_\Omega f e^H \, dx - \sum_{e \in \mathcal{E}_H} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e e^H \, ds \\
&+ \sum_{K \in \mathcal{T}_H} \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla e^H \, dx \\
= &\int_\Omega f e^H \, dx - \sum_{e \in \mathcal{E}_H} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e e^H \, ds \\
&+ \sum_{K \in \mathcal{T}_H} \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla e^H \, dx + B(u^H, I^H e^H) \\
&- \sum_{K \in \mathcal{T}_H} \int_K f (I^H e^H) dx \\
= &\int_\Omega f^H (e^H - I^H e^H) dx + \int_\Omega (f - f^H)(e^H - I^H e^H) dx \\
&- \sum_{e \in \mathcal{E}_H} \int_e [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e (e^H - I^H e^H) ds \\
&+ \sum_{K \in \mathcal{T}_H} \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla e^H \, dx.
\end{aligned}$$

We define $\phi^H := e^H - I^H e^H$. Using the Cauchy–Schwarz inequality we obtain the following estimate

$$\begin{aligned}
B_0(e^H, e^H) \leqslant C \Bigg(&\sum_{K \in \mathcal{T}_H} \|f^H\|_{L^2(K)} \|\phi^H\|_{L^2(K)} + \sum_{K \in \mathcal{T}_H} \|f - f^H\|_{L^2(K)} \|\phi^H\|_{L^2(K)} \\
&+ \sum_{e \in \mathcal{E}_H} \left\| [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e \right\|_{L^2(e)} \|\phi^H\|_{L^2(e)} \\
&+ \sum_{K \in \mathcal{T}_H} \left\| (a_K^0 - a^0(x)) \nabla u^H \right\|_{L^2(K)} \|\nabla e^H\|_{L^2(K)} \Bigg).
\end{aligned}$$

With the help of the Clément interpolation estimates (16) and (20) we deduce that

$$\begin{aligned}
B_0(e^H, e^H) \leqslant C \Bigg(&\sum_{K \in \mathcal{T}_H} H_K \|f^H\|_{L^2(K)} \|\nabla e^H\|_{L^2(N(K))} \\
&+ \sum_{K \in \mathcal{T}_H} H_K \|f - f^H\|_{L^2(K)} \|\nabla e^H\|_{L^2(N(K))} \\
&+ \sum_{e \in \mathcal{E}_H} H_e^{1/2} \left\| [\![\overline{a^\varepsilon(x) \nabla u^h}]\!]_e \right\|_{L^2(e)} \|\nabla e^H\|_{L^2(N(K_e))} \\
&+ \sum_{K \in \mathcal{T}_H} \left\| (a_K^0 - a^0(x)) \nabla u^H \right\|_{L^2(K)} \|\nabla e^H\|_{L^2(K)} \Bigg).
\end{aligned}$$

The finite overlapping property of the neighborhoods $N(K)$ allows us to estimate $\sum_{K\in\mathcal{T}_H}\|\nabla e^H\|_{L^2(N(K))}\leqslant C\sum_{K\in\mathcal{T}_H}\|\nabla e^H\|_{L^2(K)}$, where C depends only on the shape regularity of the triangulation and the dimension d. Using the coercivity of $B_0(\cdot,\cdot)$ and the triangle inequality yields

$$\|\nabla e^H\|^2_{L^2(\Omega)}\leqslant C\Bigg(\sum_{K\in\mathcal{T}_H}H_K^2\|f^H\|^2_{L^2(K)}+\sum_{e\in\mathcal{E}_H}H_e\Big\|[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\Big\|^2_{L^2(e)}+\sum_{K\in\mathcal{T}_H}H_K^2\|f-f^H\|^2_{L^2(K)}+\sum_{K\in\mathcal{T}_H}\big\|(a^0_K-a^0(x))\nabla u^H\big\|^2_{L^2(K)}\Bigg).$$

Using Poincaré inequality leads to

$$\|u^0-u^H\|^2_{H^1(\Omega)}\leqslant C(\eta_H(\Omega)^2+\xi_H(\Omega)^2)$$

as stated in Theorem 2.

5.4. Lower bound (Proof of Theorem 3)

To derive the lower bound, we will use a construction involving bubble functions in a space $\widetilde{V}^1_H\supset V^1_H$ which is defined over a refinement $\widetilde{\mathcal{T}}_H$ of $\mathcal{T}_H$. We assume that the refinement $\widetilde{V}^1_H$ is chosen such that every $K\in\mathcal{T}_H$ has an interior node $\tilde{x}_K\in K\setminus\partial K$ in $\widetilde{\mathcal{T}}_H$ and every edge e of $\mathcal{T}_H$ not on the boundary $\partial\Omega$ has an interior node in $\widetilde{\mathcal{T}}_H$. We emphasize once again that the use of the representation formula (22) allows largely to follow the classical construction of [42]. In what follows we estimate successively interior and jump residuals.

5.4.1. Interior residual

For any $K\in\mathcal{T}_H$ consider an interior bubble function, i.e., a function $\psi_K\in\widetilde{V}^1_H$ such that $0\leqslant\psi_K\leqslant 1,\ \psi_K(\tilde{x}_K)=1$ and $\psi_K\equiv 0$ on $\Omega\setminus K$.

We choose $v:=\psi_K f^H\in H^1_0(\Omega)$ as a test function in the error representation formula (22) and obtain

$$\int_K f^H(\psi_K f^H)dx=B_0(e^H,\psi_K f^H)-\int_K(f-f^H)\psi_K f^H\,dx-\int_K(a^0_K-a^0(x))\nabla u^H\cdot\nabla(\psi_K f^H)dx.$$

Using the equivalence of norms on a finite-dimensional space we have (see [11, Theorem 2.2] for details)

$$C\|f^H\|^2_{L^2(K)}\leqslant\int_K f^H(\psi_K f^H)dx.$$

Furthermore, the continuity of $B_0(\cdot,\cdot)$, the Cauchy–Schwarz inequality and the inverse inequality (19) give

$$\begin{aligned}\|f^H\|^2_{L^2(K)}&\leqslant C\Big(\|\nabla e^H\|_{L^2(K)}\|\nabla(\psi_K f^H)\|_{L^2(K)}+\|f-f^H\|_{L^2(K)}\|\psi_K f^H\|_{L^2(K)}\\&\quad+\|(a^0_K-a^0(x))\nabla u^H\|_{L^2(K)}\|\nabla(\psi_K f^H)\|_{L^2(K)}\Big)\\&\leqslant C\Big(H_K^{-1}\|\nabla e^H\|_{L^2(K)}+\|f-f^H\|_{L^2(K)}\\&\quad+H_K^{-1}\|(a^0_K-a^0(x))\nabla u^H\|_{L^2(K)}\Big)\|\psi_K f^H\|_{L^2(K)}.\end{aligned}$$

Finally, since $0\leqslant\psi_K\leqslant 1$ we have $\|\psi_K f^H\|_{L^2(K)}\leqslant\|f^H\|_{L^2(K)}$ and we obtain the *interior residual*

$$H_K^2\|f^H\|^2_{L^2(K)}\leqslant C\Big(\|\nabla e^H\|^2_{L^2(K)}+H_K^2\|f-f^H\|^2_{L^2(K)}+\|(a^0_K-a^0(x))\nabla u^H\|^2_{L^2(K)}\Big).\tag{25}$$

5.4.2. Jump residual

Let $e\in\mathcal{E}_H$ be an interior interface and let $K_1\in\mathcal{T}_H,\ x_e\in e$ be an interior node and $K_2\in\mathcal{T}_H$ such that $K_1\cap K_2=e$. Let $\psi_e\in\widetilde{V}^1_H$ be a bubble function such that $\psi_e(x_e)=1,\ \psi_e\equiv 0$ on $\Omega\setminus(K_1\cup K_2)$. Using again the equivalence of norms on a finite-dimensional space we have

$$\int_e\psi_e\,ds\geqslant C|e|\geqslant CH_e^{d-1},$$

where $|e|$ denotes the measure of e and where the constant C depends only on the shape regularity and the dimension d. As the multiscale jump $[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e$ is constant, we have

$$\begin{aligned}\int_e[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\psi_e ds&=[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\int_e\psi_e\,ds\\&\geqslant C|e|^{-1/2}\Big\|[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\Big\|_{L^2(e)}H_e^{d-1}\\&\geqslant CH_e^{\frac{d-1}{2}}\Big\|[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\Big\|_{L^2(e)}.\end{aligned}\tag{26}$$

Next, we set $v=\psi_e$ in the representation formula (22) (recall that $v\equiv 0$ on $\Omega\setminus(K_1\cup K_2)$) and obtain

$$\begin{aligned}&\int_e[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\psi_e\,ds\\&=\sum_{K_1,K_2}\Big(\int_{K_i}f\psi_e\,dx-\int_{K_i}a^0(x)\nabla e^H\nabla\psi_e\,dx\\&\quad+\int_{K_i}\big(a^0_{K_i}-a^0(x)\big)\nabla u^H\nabla\psi_e\,dx\Big)\\&\leqslant C\sum_{K_1,K_2}\Big(H_{K_i}\|f^H\|_{L^2(K_i)}+\|\nabla e^H\|_{L^2(K_i)}+H_{K_i}\|f-f^H\|_{L^2(K_i)}\\&\quad+\Big\|\big(a^0_{K_i}-a^0(x)\big)\nabla u^H\Big\|_{L^2(K_i)}\Big)H_{K_i}^{-1}\|\psi_e\|_{L^2(K_i)},\end{aligned}$$

where we used $\|\nabla\psi_e\|_{L^2(K)}\leqslant CH_K^{-1}\|\psi_e\|_{L^2(K)}\leqslant C(H_K)^{\frac{d-2}{2}}$, which follows form the inverse inequality (19). The inequality (26) and the above estimate for $\|\psi_e\|_{L^2(K)}$ yield

$$H_e\Big\|[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\Big\|^2_{L^2(e)}\leqslant C\sum_{K_1,K_2}\Big(H_{K_i}^2\|f^H\|^2_{L^2(K_i)}+\|\nabla e^H\|^2_{L^2(K_i)}+H_{K_i}^2\|f-f^H\|^2_{L^2(K_i)}+\Big\|\big(a^0_{K_i}-a^0(x)\big)\nabla u^H\Big\|^2_{L^2(K_i)}\Big).\tag{27}$$

5.4.3. Combining interior and jump residuals

We use the interior residual (25) to eliminate $\|f^H\|_{L^2(K)}$ from the jump residual

$$H_e\Big\|[\![\overline{a^\varepsilon(x)\nabla u^h}]\!]_e\Big\|^2_{L^2(e)}\leqslant C\Big(\|\nabla e^H\|^2_{L^2(\omega_e)}+H_{\omega_e}^2\|f-f^H\|^2_{L^2(\omega_e)}+\|(a^0_K-a^0(x))\nabla u^H\|^2_{L^2(\omega_e)}\Big),\tag{28}$$

where $H_{\omega_e}=\max_{i=1,2}H_i$ and $\omega_e=K_1\cup K_2$. Adding the interior residual (25)–(28) leads to the desired upper bound

$$\eta_H(K)^2\leqslant C(\|u^0-u^H\|^2_{H^1(\omega_K)}+\xi_H(\omega_K)^2)$$

as stated in Theorem 3.

5.5. Data approximation

So far, to derive our a posteriori upper and lower bound, we did not make any specific spatial assumption on the oscillating tensor (e.g., periodicity, random stationarity in the fast variable). In addition, the sampling domain size as well as boundary conditions of the micro solution for the HMM were quite general. We notice that upper and lower bounds involve the data approximation term

$$\xi_H(K)^2:=H_K^2\|f^H-f\|^2_{L^2(K)}+\|(a^0_K-a^0(x))\nabla u^H\|^2_{L^2(K)}.$$

The first term of the right-hand side of this equality involves the usual data approximation term. The second term quantifies the accuracy of the macro–micro algorithm and it depends on the macro- and micromeshes of the macro- and microFEMs, on the structure of the oscillating tensor a^ε and on the coupling condition between micro and macrosolver. We first notice that

$$\left\|\left(a_K^0 - a^0(x)\right)\nabla u^H\right\|_{L^2(K)}^2 \leqslant \sup_{x\in K}\left\|a_K^0 - a^0(x)\right\|_F^2 \|\nabla u^H\|_{L^2(K)}^2,$$

where we recall that for a given tensor $\|\cdot\|_F$ denotes its Frobenius norm. Let us then introduce the following tensor

$$\bar{a}_K^0 = \frac{1}{|K_\delta|}\int_{K_\delta} a^\varepsilon(x)(I + J_{\psi_K(x)}^T)dx, \tag{29}$$

where $J_{\psi(x)}$ is a $d\times d$ matrix with entries $(J_{\psi_K(x)})_{ij} = \left(\partial\psi_K^i\right)\big/(\partial x_j)$. This tensor is computed similarly to the tensor a_K^0 in (14) but with functions $\psi_K^i(x)$ solving (13) in the exact Sobolev space $W(K_\delta)$ instead of its FE approximation $S^1(K_\delta, \mathcal{T}_h)$. We then consider the following decomposition

$$\left\|a_K^0 - a^0\right\|_F \leqslant \underbrace{\left\|a^0 - a^0(x_K)\right\|_F}_{err_{mac}} + \underbrace{\left\|a^0(x_K) - \bar{a}_K^0\right\|_F}_{err_{mod}} + \underbrace{\left\|\bar{a}_K^0 - a_K^0\right\|_F}_{err_{mic}}, \tag{30}$$

where $x_K \in K$ is the quadrature node located at its barycenter. In the above equation, err_{mac} and err_{mic} stands for macroscopic and microscopic error, respectively. The analysis of the macroscopic and microscopic errors relies on regularity assumptions on the homogenized and fine scale tensors. This will be analyzed in Section 5.5.1. The term denoted by err_{mod} (the modeling error) needs additional assumptions on the structure of a^ε in order to be quantified (e.g., periodicity and random stationarity). This part does not depend on the discretization parameters. As it does not depend on the specific form of the numerical method, previously derived results can be used to analyze this contribution to the error. We will examine in Section 5.5.2 the case of a non-uniformly periodic tensor and comment on the case of a random tensor.

Regularity assumptions. For the oscillating tensor a^ε we assume

$$a^\varepsilon|_K \in W^{1,\infty}(K),\ \forall K\in\mathcal{T}_H \quad \text{and} \quad \left|a_{ij}^\varepsilon\right|_{W^{1,\infty}(K)} \leqslant C_K\varepsilon^{-1} \text{ for } i,j = 1,\dots,d. \tag{31}$$

In the analysis we will often use a constant $C = \max_K C_K$ independent of K. It is clear that if (31) is valid for an initial mesh, assumed to be aligned with the possible discontinuities of a^ε, it is still valid (with the same value of C) for every mesh obtained by refining the initial one. In view of (31) we see that a^ε is allowed to be discontinuous in different macroelements but we assume that the macromesh (i.e., the interface between two neighboring elements) is aligned with these discontinuities. For the homogenized tensor we assume

$$a_{ij}^0 \text{ are Lipschitz continuous in } K \text{ for any } K\in\mathcal{T}_H. \tag{32}$$

Remark 10. Without further knowledge about the structure of the oscillating tensor a^ε, we will impose Dirichlet boundary conditions for (6) (or (14)). Assuming (31) one can show $\left|\psi_K^i\right|_{H^2(K_\delta)} \leqslant C\varepsilon^{-1}\sqrt{|K_\delta|}$, with C independent of ε, of the quadrature points x_K and the domain K_δ (this follows from classical H^2 regularity results, see for example [30, Chapter 2.6]). If $a^\varepsilon = a(x,x/\varepsilon) = a(x,y)$ is Y-periodic in y, then assuming (31), $\delta/\varepsilon\in\mathbb{N}$ and periodic boundary condition for (6) (or (14)) one can show $\left|\psi_K^i\right|_{H^2(K_\delta)} \leqslant C\varepsilon^{-1}\sqrt{|K_\delta|}$ (this follows from classical regularity results for solutions of periodic boundary value problems (see [17, Chapter 3]). Notice that in the periodic case, for more regular tensors $a^\varepsilon(x)$, one can obtain higher order estimates $\left|\psi_K^i\right|_{H^{q+1}(K_\delta)} \leqslant C\varepsilon^{-q}\sqrt{|K_\delta|},\ q\in\mathbb{N}$ (see [5]).

5.5.1. Micro- and macrodata approximation (Proof of Theorem 4)

We start with the macroerror. Assuming (32) directly gives the following estimate

$$\sup_{x\in K}\left\|a^0(x) - a^0(x_K)\right\|_F \leqslant CH_K, \tag{33}$$

where C only depends on the dimension d and the Lipschitz constant. For the microerror, we follow the ideas of [1]. In the precise form of (34), the microerror estimate has been proved in [10]. Assuming that (2) and (31) hold, we have,

$$\left\|\bar{a}_K^0 - a_K^0\right\|_F \leqslant C\left(\frac{h}{\varepsilon}\right)^2, \tag{34}$$

where C depends only on the constant in (31) and the bound (2). Combining (2), (33) and (8) we obtain

$$\sup_{x\in K}\left\|a^0(x) - a_K^0\right\|_F \leqslant C\left(H_K + \left(\frac{h}{\varepsilon}\right)^2 + err_{mod}\right), \tag{35}$$

where C only depends on the coercivity and continuity bound (2), the dimension d, the Lipschitz constant (31) and where err_{mod} is independent of the discretization parameter of the micro- and macroFE spaces.

Remark 11 (*Micromesh refinement*). From the estimates (33) and (34), we obtain the criterion for the micromesh refinement used in Algorithm 1 in Section 4.

5.5.2. Modeling error

In this section we discuss the error term $\|a^0(x_K) - \bar{a}_K^0\|$ in (30). Additional assumption on the spatial structure of the oscillating tensor a^ε is required in order to give convergence rates for the modeling error. We will consider the case of non-uniformly periodic tensor and assume

$$a^\varepsilon = a(x,x/\varepsilon) = a(x,y) \quad Y-\text{periodic in } y, \tag{36}$$

where for simplicity we set $Y = (0,1)^d$. We will sometimes refer to the variables x and y as slow and fast variables, respectively. Other tensors a^ε could be considered. For example, following the results in [25] we could also consider an appropriate random field $a^\varepsilon = a(\omega,y)$ (with invariant statistics under integer shifts). The modeling error estimates from [16] or [25, Theorem 2.1] or could be used directly to estimate $\|a^0(x_K) - \bar{a}_K^0\|_F$. We thus see that our a posteriori error analysis applies to a variety of tensors and is not restricted to the periodic case.

In the periodic case, the homogenized tensor a^0 can be computed explicitly (see e.g., [15]). For $x_K\in K$ it reads

$$a_{ij}^0(x_K) = \int_Y a_{ij}(x_K,y) + \sum_{k=1}^d a_{ik}(x_K,y)\frac{\partial\chi^j(x_K,y)}{\partial y}dy, \quad i,j = 1,\dots,d, \tag{37}$$

where the functions $\chi^j(\cdot,y)$ are solutions of the cell problem

$$\int_Y a(x_K,y)\nabla\chi^j(x_K,y)\cdot\nabla z\,dy = -\int_Y a(x_K,y)\mathbf{e}_j\cdot\nabla z\,dy \quad \forall z\in W_{per}^1(Y), \tag{38}$$

where $\mathbf{e}_j$ is the jth basis vector of $\mathbb{R}^d$. We will consider separately the coupling conditions (periodic and Dirichlet) for the microFE space. To analyze the modeling error, we can use results obtained for the FE-HMM in [25,1,2,7].

Periodic microboundary conditions. We assume that $S^1(K_\delta, \mathcal{T}_h) \subset W^1_{per}(K_\delta)$. In this case we assume that $\delta/\varepsilon \in \mathbb{N}^*$, i.e., the sampling domains cover an integer number of the exact period of the tensor $a(\cdot, x/\varepsilon)$. We then can derive the following error estimates (see [10, Theorem 5.8]). Assume that (36) and (31) hold. Then

$$\left\|a^0(x_K) - \bar{a}^0_K\right\|_F \leqslant C\delta \tag{39}$$

for the modeling error, where C is independent of H, h, and ε. If the decomposition in fast and slow variable of the tensor $a(x, x/\varepsilon)$ is explicitly known, we can slightly modify the FE-HMM macrobilinear form (5) and microproblems (6) by replacing a^ε with $a(x_K, x/\varepsilon)$. In this case, performing a similar modification of the tensor a^ε of (14) and (29), one can show that the modeling error vanishes (see [10, Theorem 5.8] for details)

$$\left\|a^0(x_K) - \bar{a}^0_K\right\|_F = 0. \tag{40}$$

Dirichlet microboundary conditions. In the case where the exact period is not known but an estimation of the size of the periodicity is available, the idea is to embed the periodic sampling domain K_ε in a larger cube K_δ with $\delta > \varepsilon$. Here we do not assume that $\delta/\varepsilon \in \mathbb{N}^*$. Artificial boundary conditions are chosen for the microsolver. Various conditions are possible and we assume $S^1(K_\delta, \mathcal{T}_h) \subset H^1_0(K_\delta)$, see (8). Assume that (36) and (31) hold. Then

$$\left\|a^0(x_K) - \bar{a}^0_K\right\|_F \leqslant C\left(\delta + \frac{\varepsilon}{\delta}\right), \tag{41}$$

where C only depends on the domain Ω and the bound (2). This estimate can be obtained following along the line of the proof of [25, Theorem 17].

6. Numerical experiments

We present in this section a series of numerical experiments which verify the sharpness of the theoretical a posteriori estimates and confirm that our adaptive scheme is efficient and effective.

We will present various elliptic problems with a two-scale, quasi 1d tensor on a square domain (Section 6.1), a crack problem with a highly oscillating 2d tensor and a random tensor (Section 6.2), and finally an L-shaped problem with a highly oscillating 2d tensor (Section 6.3).

We emphasize that the oscillating tensor of the various problems (with the exception of the random one) is chosen such that an analytical solution of the homogenized problem is available. As various contributions to the error arise in our a posteriori estimator (due to the multiscale nature of the numerical scheme), avoiding to use a refined FEM reference solution allows to assess accurately the quality of our estimator. Of course, in practice our theory and method apply to general tensors for which there is no need to derive an analytic tensor beforehand. This is further illustrated with the random tensor problem.

Notation. We will use the following notation for various quantities measured in our numerical experiments.

- $\#el$, number of macroelements for the specific mesh. It reflects the degrees of freedom.
- $EOC(e^H)$, experimental order of convergence. $EOC := d\frac{\log\left(e^H_{k-1}/e^H_k\right)}{\log(\#el(k)/\#el(k-1))}$, where e^H_k refers to the error $\|u_0 - u^H_k\|_{H^1(\Omega)}$ in the kth refinement step and d is the physical dimension of the problem (the scaling with d allows to get a convergence rate independent of the dimension).
- $EOC(\eta)$, experimental order of convergence of the indicator $\eta_H(\Omega)$. $EOC(\eta) := d\frac{\log\left(\eta^{k-1}_H/\eta^k_H\right)}{\log(\#el(k)/\#el(k-1))}$, where η^k_H refers to the indicator in the kth refinement step and d is the physical dimension of the problem.
- Z_e, the reduction factor $Z_e := \frac{e^H_k}{e^H_{k-1}}$ of the error.
- Eff, the effectivity index $\mathrm{Eff} := \frac{\eta^k_H}{e^H_k}$ allows to estimate the upper bound constant C of Theorem 2.

Furthermore we denote by $\hat{h} := (N_{mic})^{-(1/d)}$ the scaled (i.e., independent of ε) micromeshsize, where N_{mic} denotes the degrees of freedom of the microproblem on K_δ and d is the spatial dimension. Notice that $h/\varepsilon = C\hat{h}$, where $C = \delta/\varepsilon$ is usually of moderate size.

The numerical experiments were performed using the FE-HMM code presented in [8] and the implementation of the mark and refine steps are based in part on the AFEM@Matlab code (see [19]).

Remark 12. In most of the following experiments, as we refine the mesh, we reach the point when $H_K < \delta$ for some elements $K \in \mathcal{T}_H$, i.e., the macroelement K is smaller than the sampling domain. Refining beyond this point is not computationally efficient and one should switch to the fine scale solver for the whole (macro) triangle K. A precise study and analysis for such a modified algorithm will be presented elsewhere. We notice here that in the case of periodic coefficients taking a sampling domain larger than some macroelements still makes sense and allows us to check the efficiency and reliability of our estimates.

6.1. Uniform refinement test

We consider the quasi-1d problem taken from [39],

$$\begin{aligned} -\nabla\cdot(a^\varepsilon(x)\nabla u^\varepsilon) &= -1 \quad \text{in } \Omega := (0,1)^2,\\ u^\varepsilon &= 0 \quad \text{on } \Gamma_D := \{0\}\times(0,1)\cup\{1\}\times(0,1),\\ a^\varepsilon(x)\nabla u^\varepsilon &= 0 \quad \text{on } \Gamma_N := \partial\Omega\setminus\Gamma_D, \end{aligned}$$

where $a^\varepsilon(x) = a\left(x, \frac{x}{\varepsilon}\right) = \frac{2}{3}(1+x_1)\left(1+\cos^2\left(2\pi\frac{x_1}{\varepsilon}\right)\right)\cdot I_2$ and I_2 is the unit matrix. The exact homogenized solution is given by $u^0(x) = \frac{3}{2\sqrt{2}}\left(x_1 - \frac{\log(x_1+1)}{\log(2)}\right)$. We choose $\varepsilon = \delta = 10^{-5}$.

As no singularity appears in the domain, we uniformly refine the macro triangles in every iteration step and compare the error to the indicator $\eta_H(\Omega)$. Parallel to the macrorefinement, we refine the micromesh according to Remark 5. The initial mesh for the microFE spaces is chosen as $\hat{h} = \frac{1}{8}$.

We show in Fig. 2 that the error in the H^1 norm and the indicator $\eta_H(\Omega)$ both converge to zero with rate $\mathcal{O}(H)$ and thus match the

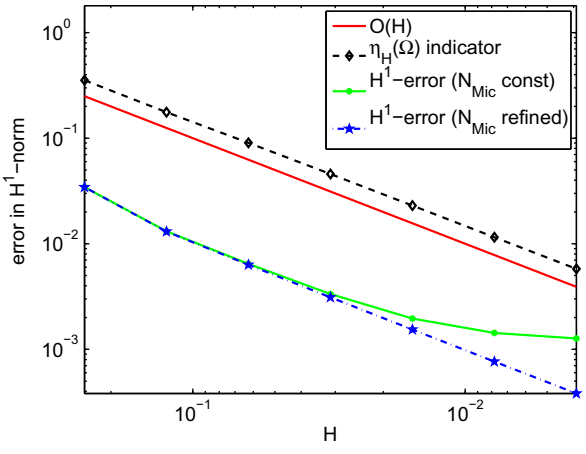


Fig. 2. Errors and error estimate in the H^1-norm for the highly oscillatory problem described in Section 6.1.

Table 1
Grid size H, H^1-error, error indicator, experimental order of convergence for the error, and the indicator and the effectivity of the problem described in Section 6.1.

H	$\|u^0 - u^H\|_{H^1(\Omega)}$	$\eta_H(\Omega)$	EOC (e^H)	EOC (η_H)	Eff $:= \frac{\eta_H}{e^H}$
2^{-2}	3.44e−02	3.54e−01			
2^{-3}	1.31e−02	1.76e−01	1.40	1.00	13.51
2^{-4}	6.33e−03	9.04e−02	1.04	0.97	14.29
2^{-5}	3.11e−03	4.57e−02	1.03	0.98	14.71
2^{-6}	1.54e−03	2.30e−02	1.01	0.99	14.93
2^{-7}	7.65e−04	1.15e−02	1.01	1.00	15.15
2^{-8}	3.82e−04	5.77e−03	1.00	1.00	15.15

prediction obtained by the a priori and posteriori estimates, respectively.

Finally, we indicate the effect of inappropriate macro–micro coupling by plotting the errors obtained with the same adaptive strategy, but computing the microsolutions on a micromesh of fixed size $\hat{h} = \frac{1}{8}$. It can be seen that the correct asymptotic convergence rate is not achieved. Thus an appropriate simultaneous mesh refinement is crucial.

In Table 1, we list various quantities (explained in detail at the beginning of this section) illustrating the quality of the refinement procedure.

The experimental order of convergence for both the error as well as the indicator confirms the theoretical linear convergence. The ratio Eff between them remains nearly constant, which confirms that our (here uniform) refinement is both effective and efficient; it furthermore gives an estimate for the *effectivity index*, the constant in the upper bound of Theorem 2.

6.2. Crack problem

In our next experiment we consider a crack problem based on [33, Example 5.2] that exhibits a singularity in the macro domain. But unlike Example 5.2 of [33], we use here a two-dimensional, highly oscillating conductivity tensor. We consider the following problem

$$-\nabla\cdot(a^\varepsilon(x)\nabla u^\varepsilon) = 1 \quad \text{in } \Omega,$$
$$u^\varepsilon = g_D \quad \text{on } \Gamma_D = \partial\Omega$$

on a domain $\Omega = \{|x_1| + |x_2| < 1\}\backslash\{0 \leqslant x_1 \leqslant 1,\ x_2 = 0\}$ with a crack along the positive x-axis (see Fig. 3). We use the tensor

$$a\left(\frac{x}{\varepsilon}\right) = \frac{64}{9\sqrt{17}}\left(\sin\left(2\pi\frac{x_1}{\varepsilon}\right) + \frac{9}{8}\right)\left(\cos\left(2\pi\frac{x_2}{\varepsilon}\right) + \frac{9}{8}\right)\cdot I_2,$$

where we chose the coefficients of the tensor in such a way that the homogenized tensor coincides with the unit tensor I_2 (see [28, Chapter 1.2]). The Dirichlet boundary conditions $g_D = u^0$ match the exact homogenized solution u^0 of the problem which is given (in polar coordinates) by

$$u^0(r,\vartheta) = r^{\frac{1}{2}}\sin\frac{\vartheta}{2} - \frac{1}{4}r^2,$$

where $x_1 = r\cos(\vartheta)$, $x_2 = r\sin(\vartheta)$. We emphasize that we use an analytically homogenizable tensor only to be able to compare our solution to the exact solution. Any other oscillating tensor could be used (see Section 6.2.2 for an experiment using a random tensor).

A solution of the problem is shown in Fig. 3. We choose $\varepsilon = \delta = 10^{-3}$ and periodic microboundary conditions. This time, we use our adaptive strategy to refine the mesh and select a total of 23 refinement steps (Dörfler's bulk-chasing strategy is used for marking, with a parameter of $\theta = 0.3$, see [42, Chapter 4.1]). We again use the relation $\hat{h}_K = \sqrt{H_K}$ for our microrefinement strategy (see Algorithm 1) with an initial mesh of $\hat{h}_K = \frac{1}{8}$.

In Fig. 4 we show the errors in the H^1 norm. The rate of convergence of the error and the error indicator confirms the theoretical rate of $\mathcal{O}(N_{mac}^{-1/d})$, where N_{mac} denotes the macro degrees of freedom. We again plot the error obtained by using the same adaptive strategy without refining the micromesh (fixed to $\hat{h} = \frac{1}{8}$). As expected, the obtained asymptotic convergence rate is incorrect.

Finally, we provide a comparison with a uniformly refined FE-HMM starting from the same initial mesh. Obviously, the order of convergence is significantly lower than what can be obtained using adaptive methods. In order to get an accuracy of e.g.,

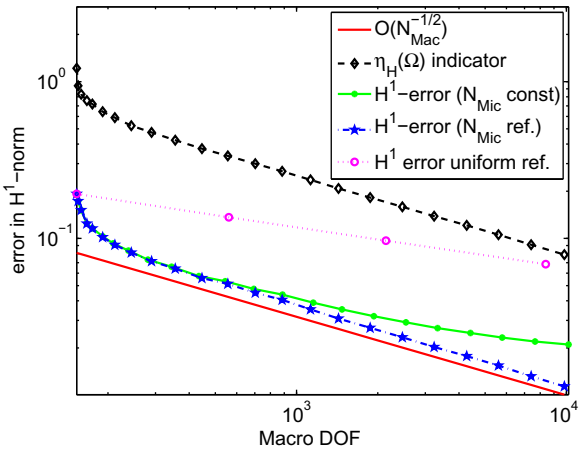


Fig. 4. Errors and error estimate in the H^1-norm for the crack problem described in Section 6.2.

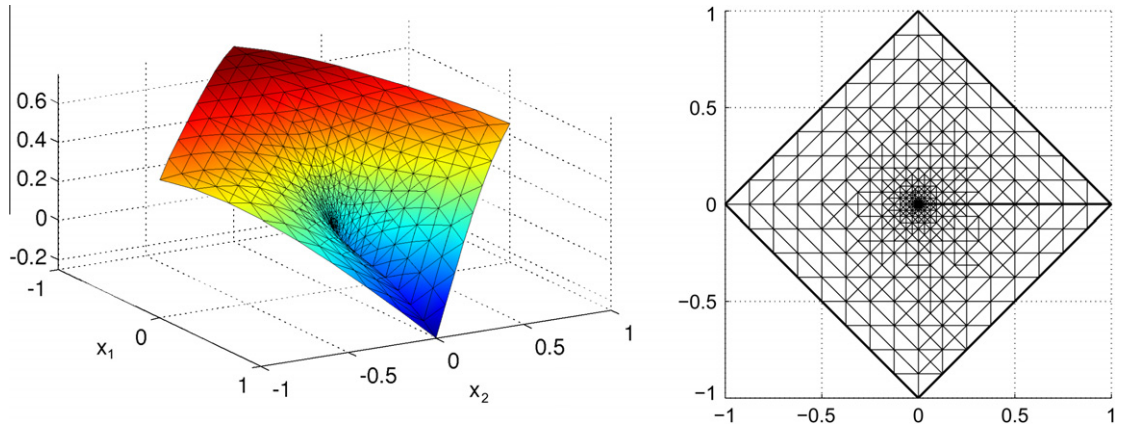


Fig. 3. FE-HMM solution and mesh after 10 iterations for the crack problem described in Section 6.2.

Table 2
Amount of microproblems with various $\hat{h}$ (due to refinement) to be solved for the specific iteration in which we reach the accuracy $\|e^H\|_{H^1(\Omega)} \leqslant 0.07$ in the crack problem described in Section 6.2.

$\hat{h}$	1/8	1/16	1/24	1/32	1/40	1/48
Adaptive FE-HMM, 10th iteration	278	218	60	24	28	40
Uniform FE-HMM, 4th iteration	–	–	–	16,384	–	–

Table 3
Total amount of microproblems with various $\hat{h}$ (due to refinement) to be solved to achieve an accuracy of $\|e^H\|_{H^1(\Omega)} \leqslant 0.07$ in the crack problem described in Section 6.2 (here we take into account all the iterations needed to reach the prescribed accuracy).

$\hat{h}$	1/8	1/16	1/24	1/32	1/40	1/48
Adaptive FE-HMM, total cost	436	360	108	48	48	40
Uniform FE-HMM, total cost	256	1024	4096	16,384	–	–

$\|u^0 - u^H\|_{H^1(\Omega)} \approx 0.07$ we need 16,384 macroelements in the fourth iteration of the uniformly refined FE-HMM, whereas only 796 macroelements are used in the 10th iteration of the adaptive FE-HMM. In order to reach an accuracy of $\mathcal{O}(10^{-2})$, approximatively $\mathcal{O}(10^8)$ DOF would be needed for the uniform scheme.

For uniform refinement, microsolutions in the sampling domain of every triangle of the macromesh have to be refined and recomputed at each step. In contrast, for the adaptive FE-HMM this needs only to be done for the macrotriangles marked for refinement. In Table 2 we list the number of microproblems with a mesh size (complexity) of $\hat{h}$ that need to be solved in the specific iteration when we reach an accuracy of $\|e^H\|_{H^1(\Omega)} \leqslant 0.07$ for the adaptive and uniform refinement, respectively. Looking at the first line of Table 2, we see that with an adaptive scheme, the sampling domains of most of the macroelements need a micromesh with a relative coarse resolution. This is in sharp contrast with a uniformly refined mesh, where all the sampling domains need to be solved with the same (fine) resolution.

As mentioned in Section 4.1, in an efficient implementation, one should store the contributions of the microproblems and the corresponding *multiscale flux* in every iteration and re-use these results for those elements which are not marked to be refined. In this way, the computational cost per iteration can be dramatically reduced, as only a fraction of the elements is refined at each iteration. For the crack problem and the adaptive FE-HMM, we start with 256 microproblems with a micromesh size chosen as $\hat{h} = \frac{1}{8}$ in the first iteration. Four elements are refined into eight elements, which require the solution of eight new microproblems with a micromesh size $\hat{h} = \frac{1}{8}$ ($\hat{h}$ is rounded s.t. $\hat{h} = \frac{1}{n \cdot 8},\ n \in \mathbb{N}^*$). On the other hand, the solutions of the 252 other microproblems will be re-used in the next iteration. After the second iteration, those 8 elements are yet again refined and we need to solve 16 new microproblems with micromesh size $\hat{h} = \frac{1}{16}$ and re-use all of the 252 solutions of the microproblems with micromesh size $\hat{h} = \frac{1}{8}$. Summing over the 10 iterations, we only need to compute 436 different microproblems of micromesh size $\hat{h} = \frac{1}{8}$. In Table 3 we list the *total* amount of different microproblems (with different micromesh) to be solved. For the uniform refinement, *every* triangle is divided into four new triangles, thus no information can be carried over from one iteration to the other.

In Table 4 we list various quantities illustrating the quality of the refinement procedure. It reveals that the error indicator converges with the same rate as the error itself. This is confirmed by the effectivity index which remains nearly constant over the iterations, showing that our adaptive refinement strategy is both effective and efficient.

In Table 5 we show the error and the experimental order of convergence when using a uniform refinement strategy. A comparison between Tables 4 and 5 shows that a desired given accuracy of the numerical solution can be obtained with a much lower computational cost using an adaptive method (compare e.g., the number of elements needed to reach a certain accuracy).

6.2.1. Sampling domain size

We consider the crack problem in the situation when the size δ of the sampling domain is not an integer multiple of ε. This situation might even arise for periodic problems when the exact size of the period is not known. We select Dirichlet boundary conditions in the microproblems and so estimates (41) apply.

We choose the same initial mesh size h for four different sampling domain sizes $\delta_1 = \frac{4}{3}\varepsilon,\ \delta_2 = \frac{5}{3}\varepsilon,\ \delta_3 = \frac{11}{3}\varepsilon$ and $\delta_4 = \frac{17}{3}\varepsilon$ (i.e., $\hat{h}$

Table 4
Iteration number, number of macroelements, H^1-error, error indicator, experimental order of convergence for the error and the indicator, reduction factor and effectivity index for the crack problem described in Section 6.2.

Iteration	#el	$\|u^0 - u^H\|_{H^1(\Omega)}$	$\eta_H(\Omega)$	$EOC(e^H)$	$EOC(\eta_H)$	Z_e	Eff
1	256	1.92e−01	1.22e+00				6.33
2	260	1.73e−01	9.41e−01	13.696	33.123	0.899	5.43
3	268	1.52e−01	8.31e−01	8.715	8.255	0.876	5.49
4	284	1.24e−01	7.58e−01	6.844	3.167	0.820	6.10
5	298	1.16e−01	7.19e−01	2.960	2.170	0.931	6.21
6	329	1.02e−01	6.43e−01	2.584	2.255	0.880	6.33
7	370	9.09e−02	5.90e−01	1.934	1.485	0.893	6.49
8	432	8.12e−02	5.23e−01	1.459	1.543	0.893	6.45
9	522	7.15e−02	4.74e−01	1.340	1.036	0.881	6.62
10	648	6.41e−02	4.23e−01	1.010	1.055	0.897	6.58
11	830	5.58e−02	3.74e−01	1.125	1.005	0.870	6.71
12	1040	5.13e−02	3.36e−01	0.746	0.940	0.919	6.54
13	1326	4.50e−02	3.00e−01	1.082	0.948	0.877	6.67
14	1670	4.05e−02	2.68e−01	0.896	0.964	0.902	6.62
15	2138	3.52e−02	2.36e−01	1.136	1.019	0.869	6.71
16	2738	3.08e−02	2.08e−01	1.085	1.022	0.874	6.76
17	3614	2.69e−02	1.82e−01	0.988	0.965	0.872	6.76
18	4782	2.34e−02	1.59e−01	0.989	0.970	0.871	6.80
19	6268	2.03e−02	1.39e−01	1.057	1.010	0.867	6.85
20	8310	1.78e−02	1.21e−01	0.940	0.961	0.876	6.80
21	10,948	1.55e−02	1.06e−01	1.009	0.998	0.870	6.85
22	14,534	1.32e−02	9.11e−02	1.114	1.036	0.854	6.90
23	19,360	1.14e−02	7.89e−02	1.031	1.010	0.863	6.94

Table 5
Iteration number, number of macroelements and experimental order of convergence of the H^1 error for the crack problem described in Section 6.2 when using uniform refinement instead of an adaptive strategy.

iteration	1	2	3	4
$\#el$	256	1024	4096	16,384
$\|u^0 - u^H\|_{H^1(\Omega)}$	1.92e−01	1.36e−01	9.67e−02	6.85e−02
$EOC(e^H)$		0.496	0.495	0.497

differs), where $\varepsilon = 10^{-5}$. According to (41) one should take $\delta \propto \sqrt{\varepsilon}$ (of course the value of δ is unknown if ε is unknown. Furthermore, taking $\delta = \sqrt{\varepsilon}$ can be computationally too expensive). We verify here that increasing the sampling domain size does improve the convergence rates. This is illustrated in Fig. 5.

6.2.2. Random tensor

Many problems of interest are not periodic. As mentioned earlier our adaptive algorithm does not rely on periodic problems (although the relation between micro- and macromesh does). To illustrate the versatility of the method, we test the behavior of the adaptive FE-HMM on the crack problem in Section 6.2 with a random tensor. This tensor is a log-normal stochastic field generated by the moving ellipse average method (see [43, Section 4.1]). The numerically generated values of the tensor are given at 9000^2 discrete points and we use bilinear interpolation to obtain a smooth representation of $a^\varepsilon(x)$.

We set the correlation lengths of the stochastic field to $\varepsilon_{x_1} = 0.0045$ and $\varepsilon_{x_2} = 0.0065$ and choose the mean equal to be 1 and variance to be $\sigma = 0.25$. A snapshot of this tensor is shown in Fig. 6.

As no analytical solution exists, we use the error e^H between a finescale solution computed using a highly resolved standard adaptive FEM with 10^7 DOF, which acts as a reference solution, and the FE-HMM solution. We denote by $P_H u^H_{FEM}$ the L^2-projection of the resolved FEM solution u^H_{FEM} onto the FE space $V^1(\Omega, \mathcal{T}_{H,FE\text{-}HMM})$, where $\mathcal{T}_{H,FE\text{-}HMM}$ denotes the mesh obtained in the adaptive algorithm for the FE-HMM. We define the error as $e^H := \|P_H u^H_{FEM} - u^H_{FE-HMM}\|_{H^1(\Omega)}$.

We choose an initial micromesh size $h = \frac{1}{8}$ and sampling domains of size $\delta_1 = 0.005$, $\delta_2 = 0.010$, $\delta_3 = 0.020$ and $\delta_4 = 0.040$ (we keep h fixed when increasing the sampling domain size). In Fig. 7 we see that the indicator and the error follow the expected (optimal) convergence rate. Furthermore, increasing the sampling domain size reduces the modeling error and thus leads to a more accurate solution.

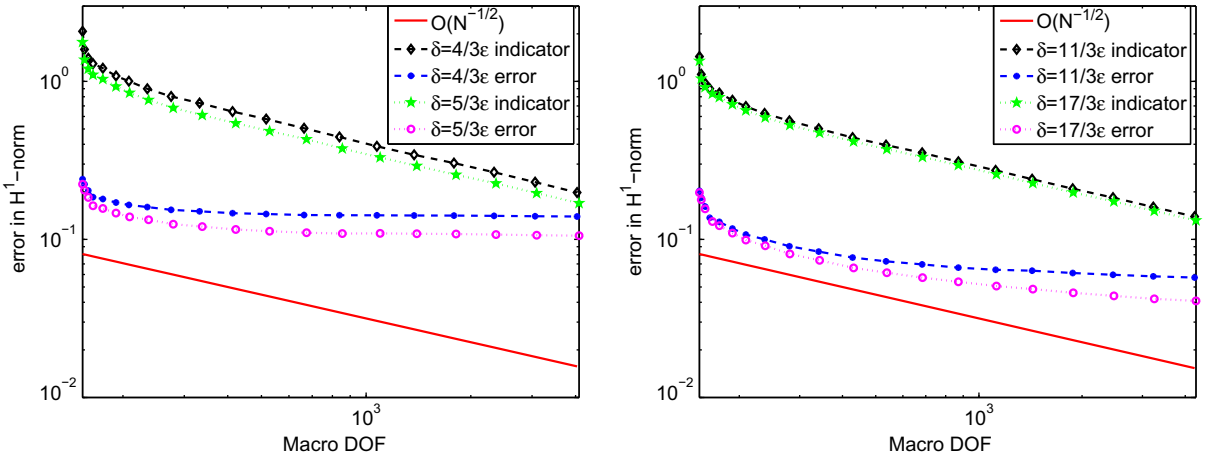
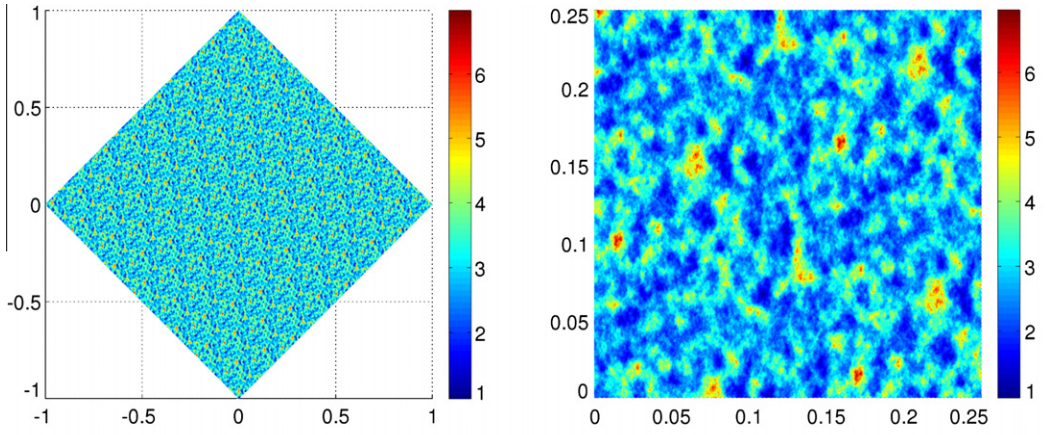


Fig. 5. Errors and error estimate in the H^1-norm for the crack problem with $\delta \neq \varepsilon$ described in Section 6.2.1.



(a) Snapshot of random conductivity tensor.

(b) Zoom into snapshot of random conductivity tensor.

Fig. 6. Snapshot of the random conductivity tensor used in the crack problem described in Section 6.2.2.

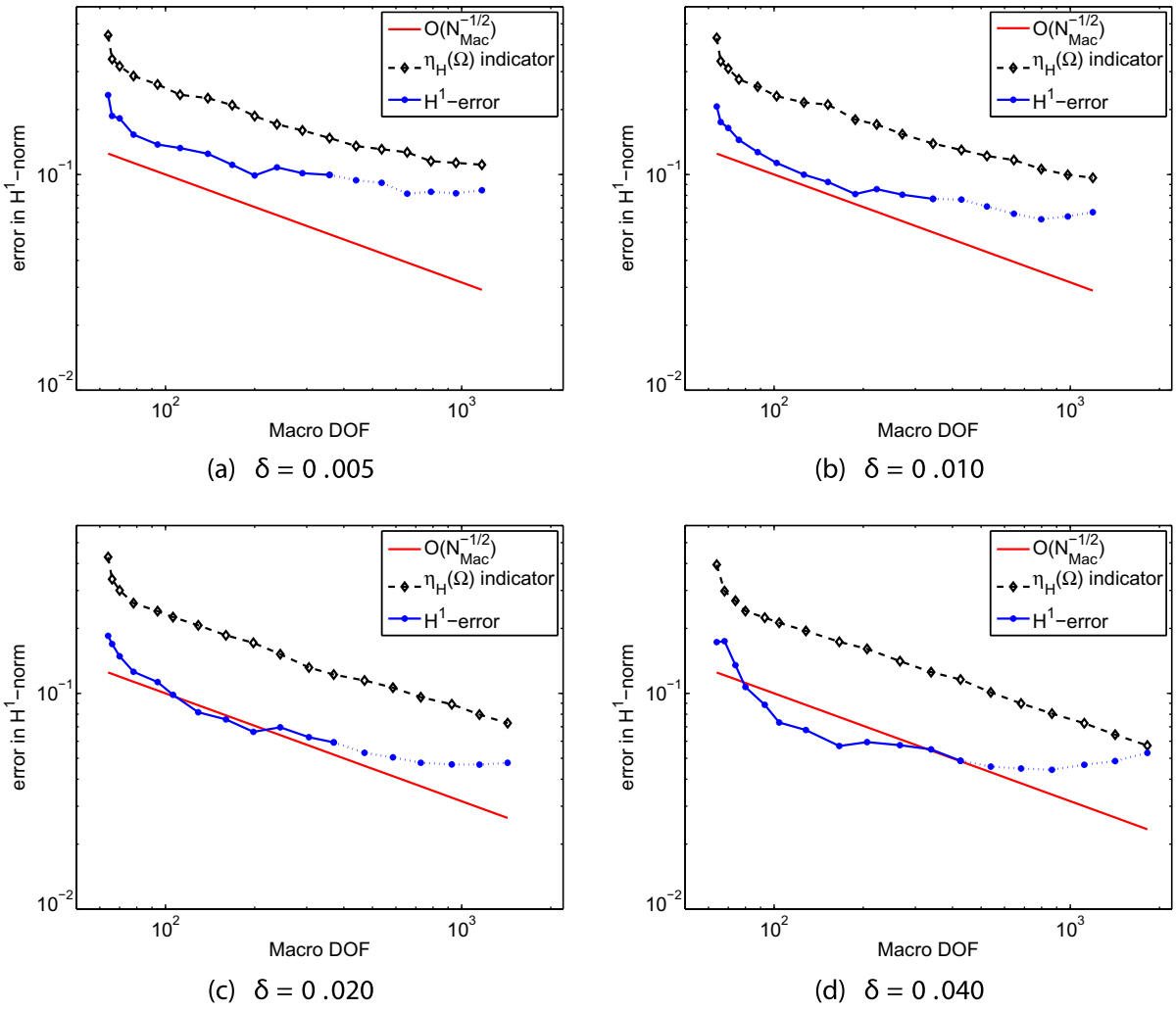


Fig. 7. Error and error estimates in the H^1-norm for the crack problem with a random tensor described in Section 6.2.2 (for plotting reasons the indicator is scaled by a factor of 0.1). For the H^1-error a continuous line is used in the case $H > \varepsilon$ and a dotted line is used when some K reach $H < \varepsilon$, see Remark 13.

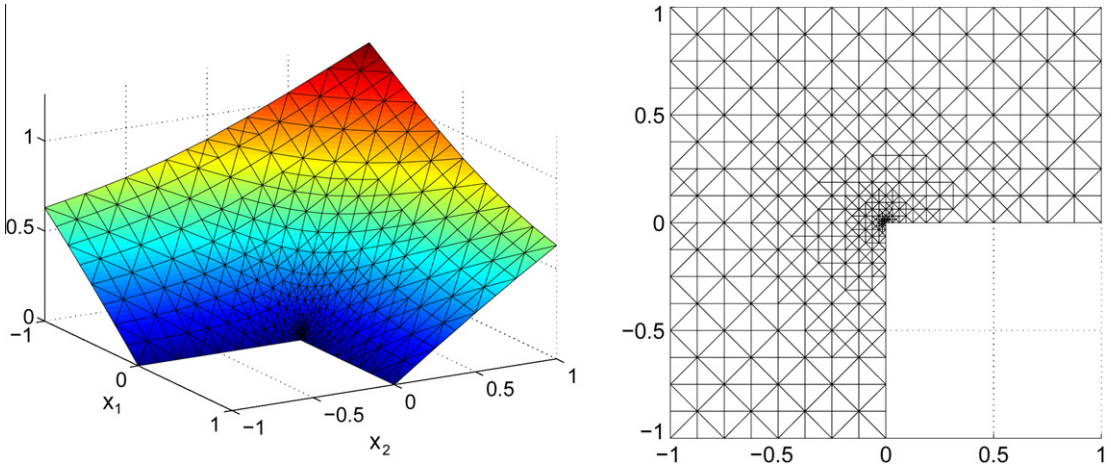
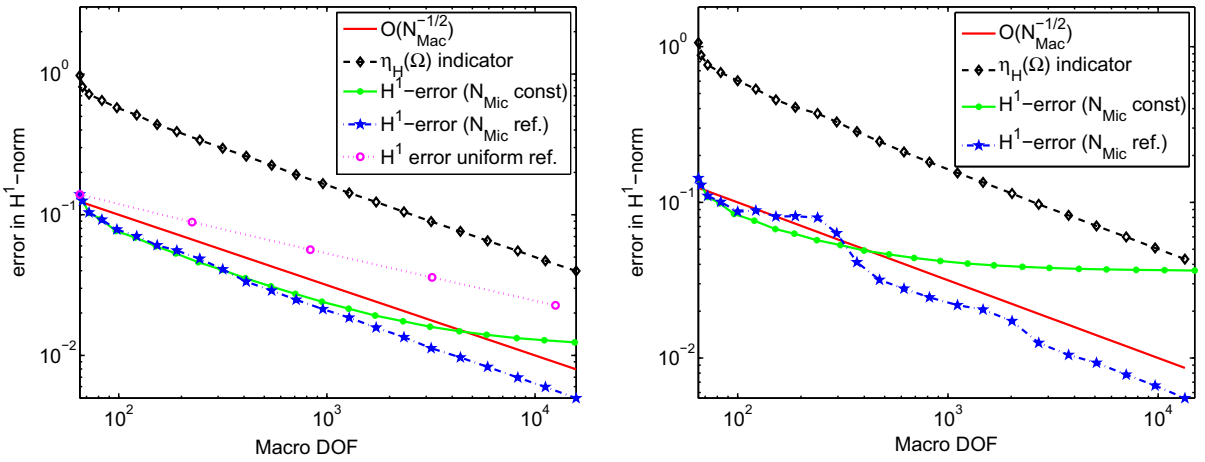


Fig. 8. FE-HMM solution and refined grid after 10 iterations using tensor a_1^ε of L-shape problem described in Section 6.3.

Remark 13. For $H < \varepsilon$, there is less than a period of the fine scale solution that is averaged (by the L^2-projection) on each macroelement of the mesh $\mathcal{T}_{H,FE\text{-}HMM}$ and as $H \to 0$ the L^2-projection restores the behavior of the fine scale solution whose gradient has a $\mathcal{O}(1)$ discrepancy with the solution obtained by the FE-HMM (see Section 3.1.1 and recall, e.g., Remark 12).



(a) L-shape problem using tensor $a_1^\varepsilon(x)$. (b) L-shape problem using tensor $a_2^\varepsilon(x)$.

Fig. 9. Errors and error estimate in the H^1-norm for the L-shape problem described in Section 6.3. For the left plot we used tensor a_1^ε, for the right-hand side we used a_2^ε.

Table 6
Iteration number, number of macroelements, H^1-error, error indicator, experimental order of convergence for the error and the indicator, reduction factor and effectivity of the L-shape problem using tensor a_1^ε described in Section 6.3.

Iteration	#el	$\|u^0 - u^H\|_{H^1(\Omega)}$	$\eta_H(\Omega)$	$EOC(e^H)$	$EOC(\eta_H)$	Z_e	Eff $:= \frac{\eta_H}{e^H}$
1	96	1.43e−01	1.061			0.135	
2	100	1.30e−01	0.880	4.886	9.145	0.905	6.80
3	110	1.10e−01	0.763	3.423	2.995	0.850	6.94
4	130	1.01e−01	0.682	1.064	1.343	0.915	6.76
5	164	8.72e−02	0.604	1.235	1.048	0.866	6.94
6	206	8.85e−02	0.534	−0.133	1.082	1.015	6.02
7	265	8.10e−02	0.456	0.708	1.263	0.915	5.62
8	335	8.15e−02	0.408	−0.052	0.943	1.006	5.00
9	435	7.97e−02	0.372	0.166	0.707	0.979	4.67
10	542	6.34e−02	0.329	2.090	1.128	0.795	5.18
11	670	4.13e−02	0.285	4.043	1.347	0.651	6.90
12	866	3.18e−02	0.246	2.036	1.143	0.770	7.75
13	1156	2.79e−02	0.211	0.907	1.064	0.877	7.58
14	1545	2.45e−02	0.181	0.882	1.039	0.880	7.41
15	2123	2.19e−02	0.154	0.729	1.014	0.891	7.04
16	2787	2.05e−02	0.134	0.483	1.022	0.936	6.58
17	3876	1.73e−02	0.114	1.030	0.998	0.844	6.58
18	5238	1.25e−02	0.097	2.148	1.060	0.724	7.75
19	7312	1.05e−02	0.083	1.062	0.977	0.838	7.87
20	9934	9.30e−03	0.071	0.765	1.020	0.889	7.58
21	13,772	7.80e−03	0.060	1.074	0.998	0.839	7.69
22	19,010	6.70e−03	0.051	0.996	1.008	0.852	7.63
23	26,430	5.50e−03	0.043	1.129	1.011	0.830	7.81

6.3. L-shape problem

In our final example we consider a non-convex L-shape domain with two different conductivity tensors. We consider the following problem

$$-\nabla\cdot(a^\varepsilon(x)\nabla u^\varepsilon) = 1 \quad \text{in } \Omega,$$
$$u^\varepsilon = g_D \quad \text{on } \Gamma_D = \partial\Omega,$$

with $\Omega = (-1,1)^2\backslash[0,1]\times[-1,0]$ (see Fig. 8). As a first test, we consider the tensor

$$a_1^\varepsilon(x) = a_1\left(\frac{x}{\varepsilon}\right) = \frac{64}{9\sqrt{17}}\left(\sin\left(2\pi\frac{x_1}{\varepsilon}\right)+\frac{9}{8}\right)\left(\cos\left(2\pi\frac{x_2}{\varepsilon}\right)+\frac{9}{8}\right)\cdot I_2,$$

introduced in Section 6.2, where I_2 is the 2×2 unit matrix. The tensor has coefficients chosen such that the homogenized tensor matches the identity tensor. We will further investigate the same tensor with different coefficients

$$a_2^\varepsilon(x) = a_2\left(\frac{x}{\varepsilon}\right) = \frac{400}{21\sqrt{41}}\left(\sin\left(2\pi\frac{x_1}{\varepsilon}\right)+\frac{21}{20}\right)\left(\cos\left(2\pi\frac{x_2}{\varepsilon}\right)+\frac{21}{20}\right)\cdot I_2.$$

The homogenized problem corresponding to this latter tensor again leads to a problem with a homogenized tensor equal to I_2. As the tensor a_2^ε has a coercivity bound closer to zero, one expects a larger error than with the tensor a_1^ε (recall that our a posteriori error estimates depend on the bound (2)). For both a_1^ε and a_2^ε, an analytical homogenized solution exists and is given by $u^0(r) = r^{\frac{2}{3}}\sin\left(\frac{2}{3}\vartheta\right)$ where $x_1 = r\cos(\vartheta)$ and $x_2 = r\sin(\vartheta)$. We take the value of this exact solution for the Dirichlet boundary condition $g_D = u^0$ (in the computation below we use Algorithm 1 with Dörfler's bulk-chasing marking strategy with parameter $\theta = 0.3$).

In Fig. 9 we compare the error and the indicator $\eta_H(\Omega)$ in the H^1 norm for the tensors a_1^ε and a_2^ε. We choose an initial meshsize of $H = \frac{1}{8}$ and $\delta = \varepsilon = 10^{-5}$ with periodic boundary conditions in the

Table 7
Amount of microproblems with various $\hat{h}$ (due to refinement) to be solved for the specific iteration in which we reach an accuracy $\|e^H\|_{H^1(\Omega)} \leqslant 0.025$ in the L-shape problem described in Section 6.3.

$\hat{h}$	1/8	1/16	1/24	1/32	1/40	1/48	1/56	1/64	1/80
Adaptive FE-HMM, 13th iteration	0	862	246	64	72	30	18	24	24
Uniform FE-HMM, 5th iteration	–	–	–	–	–	24,576	–	–	–

Table 8
Total amount of microproblems with various $\hat{h}$ (due to refinement) to be solved to reach an accuracy of $\|e^H\|_{H^1(\Omega)} \leqslant 0.025$ in the L-shape problem described in Section 6.3 (here we take into account all the iterations needed to reach the prescribed accuracy).

$\hat{h}$	1/8	1/16	1/24	1/32	1/40	1/48	1/56	1/64	1/80
Adaptive FE-HMM, total cost	288	1556	388	112	96	48	36	36	24
Uniform FE-HMM, total cost	96	384	1536	6144	0	24,576	–	–	–

microproblems. For both problems we find that the corresponding indicator converges with the same (optimal) rate of $\mathcal{O}(N_{mac}^{-1/d})$ as the errors between the FE-HMM solution and the exact homogenized solution, thus confirming numerically our theoretical estimates. As noted earlier we verify again that the asymptotic convergence rate is incorrect when using a constant micromesh size $\hat{h} = \frac{1}{8}$. The asymptotic limit of the macroscopic error in Fig. 9b (green[2] continuous line with disks) reflects the error introduced through the microproblems and illustrates that the error stagnates even though we keep refining our macrogrid. A comparison between Fig. 9a and b shows that with constant micromesh size the error is significantly larger when using a_2^ε as compared to a_1^ε (the tensor a_2^ε is closer to being singular and thus needs a higher resolution of the micromesh in order to avoid a singular stiffness matrix).

Various quantities illustrating the quality of the refinement procedure and confirming the correct experimental order of convergence are reported in Table 6. In particular, we see that the effectivity index is approximatively constant indicating that our adaptive scheme is both effective and efficient. It is furthermore robust with respect to the change of tensors in the problem. In Tables 7 and 8 we compare again the number of sampling domains and the resolution of the mesh needed to solve the microproblems with adaptive and non-adaptive strategies. The results illustrate once more the importance of adaptive methods for multiscale problems.

7. Conclusion

In this paper we have given an a posteriori error analysis for a multiscale FE method, the FE-HMM, and derived explicit localized error indicators for robust and reliable adaptive mesh refinement. These are the first rigorous a posteriori results for the FE-HMM derived in the *energy norm* of the physical variables. Our numerical results confirm that the adaptive strategy is both reliable and efficient. Up to a data approximation term, upper and lower bounds are obtained without specific structure assumptions (as periodicity, random stationarity) on the oscillating tensor of the elliptic problem. A (non-uniform) refinement of the macromesh should be coupled to a refinement of the micromesh covering the sampling domain. A strategy for such a micro refinement has been proposed and justified in the case of non-uniform periodic coefficients. The adaptive algorithm does not rely on a fixed size of the sampling domain. Estimates of the error introduced by artificial (Dirichlet) boundary conditions and domain size larger than a typical length of the small scale have been derived for the case of non-uniformly periodic oscillating coefficients. The framework that we used to derive our results allowed us to use the strategy developed for single scale problem. Furthermore, the derived a posteriori estimates are consistent with classical explicit residual-based a posteriori error estimators applied to the homogenized problems in the case of periodic tensors and resolved microcalculations.

We did not address in this paper some important topics. We mention the estimation of the error in quantities of interests which might be needed for a specific design purpose [11, Chapter 8]. For such estimates, it is known that the error measured in the energy norm can be used to obtain error estimates for other quantities of interests. We also mention that for macro-to-micro multiscale methods as considered in this paper, it is desirable to change the physical model in those regions of the computational domain where the size of the macroelements becomes smaller than the corresponding sampling domain. Combining model adaptivity [38] with the adaptive strategy proposed in this paper is of high interest. Extensions of our adaptive algorithm to model adaptivity and refinements related to quantities of interests are currently under investigation and detailed discussion will be reported elsewhere.

[2] For interpretation of color in Fig. 9, the reader is referred to the web version of this article.

References

[1] A. Abdulle, On a-priori error analysis of fully discrete heterogeneous multiscale FEM, SIAM Multiscale Model. Simul. 4 (2) (2005) 447–459.
[2] A. Abdulle, C. Schwab, Heterogeneous multiscale FEM for diffusion problem on rough surfaces, SIAM Multiscale Model. Simul. 3 (1) (2005) 195–220.
[3] A. Abdulle, Analysis of a heterogeneous multiscale FEM for problems in elasticity, Math. Models Methods Appl. Sci. 16 (2) (2006) 615–635.
[4] A. Abdulle, Heterogeneous multiscale methods with quadrilateral finite elements, in: Numerical Mathematics and Advanced Applications, Springer, Berlin, 2006, 743–751.
[5] A. Abdulle, B. Engquist, Finite element heterogeneous multiscale methods with near optimal computational complexity, SIAM Multiscale Model. Simul. 6 (4) (2007) 1059–1084.
[6] A. Abdulle, Multiscale method based on discontinuous Galerkin methods for homogenization problems, C.R. Acad. Sci. Paris, Ser. I 346 (1–2) (2008) 97–102.
[7] A. Abdulle, The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs, GAKUTO Int. Ser. Math. Sci. Appl. 31 (2009) 135–184.
[8] A. Abdulle, A. Nonnenmacher, A short and versatile finite element multiscale code for homogenization problems, Comput. Methods Appl. Mech. Engrg. 198 (37–40) (2009) 2839–2859.
[9] A. Abdulle, A. Nonnenmacher, A posteriori error analysis of the heterogeneous multiscale method for homogenization problems, C.R. Acad. Sci. Paris, Ser. I 347 (17–18) (2009) 1081–1086.
[10] A. Abdulle, Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales, preprint.
[11] M. Ainsworth, J.T. Oden, A Posteriori Error Estimation in Finite Element Analysis, John Wiley & Sons, New York, 2000.
[12] S. Agmon, Lectures on Elliptic Boundary Value Problems, D. Van Nostrand Co., Princeton, NJ – Toronto – London, 1965.
[13] I. Babŭska, E. Osborn, Generalized finite element methods: their performance and their relation to mixed methods, SIAM J. Numer. Anal. 20 (1983) 510–536.

[14] I. Babüska, G. Caloz, E. Osborn, Special finite element methods for a class of second order elliptic problems with rough coefficients, SIAM J. Numer. Anal. 31 (1994) 945–981.
[15] A. Bensoussan, J.-L. Lions, G. Papanicolaou, Asymptotic Analysis for Periodic Structures, North Holland, Amsterdam, 1978.
[16] A. Bourgeat, A. Piatnitski, Approximations of effective coefficients in stochastic homogenization, Ann. Inst. H. Poincaré Prob. Stat. 40 (2) (2004) 153–165.
[17] L. Bers, F. John, M. Schechter, Partial differential equations, in: Lectures in Applied Mathematics, Proceedings of the Summer Seminar, Boulder, CO, 1957.
[18] C. Bernardi, Y. Maday, F. Rapetti, Discrétisations variationnelles de problèmes aux limites elliptiques, Mathématiques & Applications, vol. 45, Springer-Verlag, Berlin, 2004.
[19] L. Chen, C.S. Zhang, AFEM@matlab: A Matlab Package of Adaptive Finite Element Methods, Technical Report, 2006.
[20] P. Ciarlet, The finite element method for elliptic problems, Classics in Applied Mathematics, vol. 40, SIAM, Philadelphia, PA, 2002.
[21] P. Clément, Approximation by finite element functions using local regularization, Rev. Française Automat. Informat. Recherche Opérationnelle, Sér. RAIRO Analyse Numérique 9 (R-2) (1975) 77–84.
[22] D. Cioranescu, P. Donato, An introduction to homogenization, Oxford Lecture Series in Mathematics and its Applications, vol. 17, Oxford University Press, New York, 1999.
[23] E. De Giorgi, S. Spagnolo, Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine, Boll. Un. Mat. Ital. 4 (8) (1973) 391–411.
[24] W. E, B. Engquist, The heterogeneous multiscale methods, Commun. Math. Sci. 1(1) (2003), 87–132.
[25] W. E, P. Ming, P. Zhang, Analysis of the heterogeneous multiscale method for elliptic homogenization problems, J. Am. Math. Soc. 18 (1) (2005) 121–156.
[26] V.H. Hoang, C. Schwab, High-dimensional finite elements for elliptic problems with multiple scales, SIAM Multiscale Model. Simul. 3 (1) (2005) 168–194.
[27] T.Y. Hou, X-H. Wu, Z. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, Math. Comput. 68 (227) (1999) 913–943.
[28] V.V. Jikov, S.M. Kozlov, O.A. Oleinik, Homogenization of Differential Operators and Integral Functionals, Springer-Verlag, Berlin, Heidelberg, 1994.
[29] V. Kouznetsova, W.A.M. Brekelmans, F.P.T. Baaijens, An approach to micro–macro modeling of heterogeneous materials, Comput. Mech. 27 (2001) 37–48.
[30] O.A. Ladyzhenskaya, The boundary value problems of mathematical physics, Applied Mathematical Sciences, vol. 49, Springer-Verlag, New York Inc., 1985.
[31] A.M. Matache, C. Schwab, Two-scale FEM for homogenization problems, M2AN Math. Model. Numer. Anal. 36 (4) (2002) 537–572.
[32] C. Miehe, J. Schröder, C. Bayreuther, On the homogenization analysis of composite materials based on discretized fluctuations on the micro-structure, Acta Mech. 135 (2002) 1–16.
[33] P. Morin, R.H. Nochetto, K.G. Siebert, Convergence of adaptive finite element methods, SIAM Rev. 44 (4) (2002) 631–658.
[34] F. Murat, L. Tartar, H-convergence, Topics in the Mathematical Modelling of Composite Materials, Progr. Nonlinear Differential Equations Appl., vol. 31, Birkhäuser Boston, Boston, MA, 1997, 21–43.
[35] N. Neuss, W. Jäger, G. Wittum, Homogenization and multigrid, Computing 66 (1) (2001) 1–26.
[36] G. Nguetseng, A general convergence result for a functional related to the theory of homogenization, SIAM J. Math. Anal. 20 (3) (1989) 608–623.
[37] R.H. Nochetto, Adaptive finite element methods for elliptic PDEs, Lecture Notes, <http://www-users.math.umd.edu/rhn/lectures.html>.
[38] J.T. Oden, S. Prudhomme, A. Romkes, P. Bauman, Multiscale modeling of physical phenomena: adaptive control of models, SIAM J. Sci. Comput. 28 (6) (2006) 2359–2389.
[39] M. Ohlberger, A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems, Multiscale Model. Simul. 4 (1) (2005) 88–114.
[40] A. Schmidt, K.G. Siebert, Design of adaptive finite element software: the finite element toolbox ALBERTA, Lecture Notes in Computational Science and Engineering, vol. 42, Springer-Verlag, Berlin, 2005.
[41] K. Terada, N. Kikuchi, A class of general algorithms for multi-scale analyses of heterogeneous media, Comput. Methods Appl. Mech. Engrg. 190 (40–41) (2001) 5427–5464.
[42] R. Verfürth, A review of a posteriori error estimation and adaptive mesh-refinement techniques, Wiley-Teubner, New York, 1996.
[43] T.C. Wallstrom, S. Hou, M.A. Christie, L.J. Durlofsky, D.H. Sharp, Accurate scale up of two phase flow using renormalization and nonuniform coarsening, Comput. Geosci. 3 (1) (1999) 67–87.
[44] Q. Yu, J. Fish, Multiscale asymptotic homogenization for multiphysics problems with multiple spatial and temporal scales: a coupled thermo-viscoelastic example problem, Int. J. Solids Struct. 39 (26) (2002) 6429–6452.