

Optimization based methods for highly heterogeneous multiscale problems and multiscale methods for elastic waves

THÈSE N° 7467 (2017)

PRÉSENTÉE LE 3 MARS 2017

À LA FACULTÉ DES SCIENCES DE BASE

CHAIRE D'ANALYSE NUMÉRIQUE ET MATHÉMATIQUES COMPUTATIONNELLES
PROGRAMME DOCTORAL EN MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Orane Camille JECKER

acceptée sur proposition du jury:

Prof. F. Eisenbrand, président du jury

Prof. A. Abdulle, directeur de thèse

Dr M. D'Elia, rapporteuse

Prof. P. Lin, rapporteur

Prof. M. Picasso, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

À ma famille

Remerciements

Après quatre années passées à l'EPFL, dans le groupe ANMC, ma thèse de doctorat touche à sa fin et j'aimerais remercier toutes les personnes qui ont rendu cette expérience si positive et agréable.

Tout d'abord, je souhaite remercier mon directeur de thèse, Prof. Assyr Abdulle, pour m'avoir permis de réaliser cette incroyable aventure que représente une thèse de doctorat. Je le remercie pour toutes les discussions et précieux conseils qu'il a su partager avec moi tout au long de mon travail. Toutes ses remarques et commentaires, notamment sur la rédaction de textes mathématiques ou sur la présentation orale des sujets ma thèse, m'ont été très bénéfiques et m'ont permis d'acquérir des connaissances qui me seront utiles dans le futur.

Je suis également reconnaissante envers les membres du jury, présidé par le Prof. Friedrich Eisenbrand (EPFL) et composé des rapporteurs Dr Marta D'Elia (Sandia, Albuquerque, USA), Prof. Ping Lin (Université de Dundee, Écosse) et Prof. Marco Picasso (EPFL). Je les remercie d'avoir accepté de faire partie du jury et d'avoir pris le temps de lire et de commenter ma thèse. Je remercie aussi Alexander Shapeev pour sa collaboration effectuée dans le travail sur les méthodes de couplage, ainsi que pour toutes les discussions et échanges faits durant ma thèse.

Ces années passées à l'EPFL sont remplies de moments conviviaux partagés avec mes collègues et plus en particulier avec les membres du groupe MATHICSE. Je remercie mes amis du groupe ANMC, Yun, Andrea, Patrick, Virginie, Ondrej, Martin, Adrian, Simon, Edoardo, Giacomo et Giacomo pour avoir créé une ambiance dynamique et amicale au sein de notre groupe. Un grand merci à tous mes amis qui ont rendu mon parcours plaisant et amusant. Je remercie spécialement Tristan, Mathieu et Mikael.

Cette thèse n'aurait pas été possible sans le soutien infaillible de mes parents et de ma famille. Je tiens finalement à remercier Timothée pour son support, son aide et son amour. Je vous suis infiniment reconnaissante, car grâce à vous, j'ai pu réaliser ce travail, m'épanouir et vivre pleins d'expériences enrichissantes.

Lausanne, janvier 2017

Orane Jecker

Abstract

Multiscale or multiphysics partial differential equations are used to model a wide range of physical systems with various applications, e.g. from material and natural science to problems in biology or engineering. When the ratio between the smallest scale in the problem and the size of the physical domain (also the size of the solution) is very small, the numerical approximation of the effective behaviour with classical numerical methods, such as the finite element method (FEM), can become computationally prohibitive. Indeed, as the smallest scale in the problem has to be fully resolved, one obtains a discretization of the computational domain with a very large number of degrees of freedom.

In the first part of the thesis, we derive a finite element heterogeneous multiscale method (FE-HMM) applied to the wave equation in a linear elastic medium. We state the FE-HMM and give robust a priori error estimates with explicit convergence rates for the macro and micro discretizations. For simplicity, we start with the static highly heterogeneous linear problem and, then, add the time dependency and consider the wave propagation in a highly heterogeneous linear elastic medium.

In the second part of the thesis we are interested in problems in which the scales are well separated only in some regions of the computational domain, with possibly a continuum of scales in the complementary domain. Such problems arise in various situations, for example in heterogeneous composite materials whose effective properties can be well captured by assuming a (locally) periodic microstructure that can however not be valid near defects of the material. In our modeling, the smallest scale is supposed to be still discretized at the continuum level, but for some applications atomistic scale should be considered. Our coupling method is based on a domain decomposition into a family of overlapping domains. Virtual (interface) controls are introduced as boundary conditions, and act as unknown traces or fluxes. Our method is formulated as a minimization problem with states equations as constraints. The optimal boundary controls of two overlapping domains are found by an heterogeneous optimization problem that is based on minimizing the discrepancy between the two models on, at first, the overlapping region, and at second, over the boundary of the overlapping region. The fully discrete optimization based method couples the continuous or discontinuous Galerkin FE-HMM with the FEM. The well-posedness of our method, in continuous and discrete forms, are established and (fully discrete) a priori error estimates are derived.

Key words: multiscale problems, heterogeneous multiscale method, homogenization, linear elasticity problems, wave equation, global to local methods, domain decompositions, optimization based methods.

Résumé

Les équations aux dérivées partielles à multi-échelles ou multi-physiques sont utilisées pour modéliser une large classe de systèmes physiques présents par exemple dans les sciences matérielles et naturelles ou en biologie et ingénierie. Lorsque le rapport entre la plus petite échelle du problème et la taille du domaine physique (ou aussi la taille de la solution) est très petit, l'approximation numérique du comportement effectif de la solution avec des méthodes numériques classiques, telles que la méthode des éléments finis (MEF), peut devenir computationnellement prohibitive. En effet la plus petite échelle du problème doit être résolue et l'on obtient une discrétisation du domaine computationnel avec un très grand nombre de degrés de liberté.

Dans la première partie de la thèse, nous donnons une méthode multi-échelle d'éléments finis appliquée à l'équation d'ondes dans un milieu élastique linéaire. Nous établissons la méthode et donnons des estimations d'erreur a priori avec des ordres de convergence explicites pour la discrétisation macroscopique et microscopique. Nous commençons par traiter un problème statique dans un milieu élastique linéaire multi-échelles et ensuite nous considérons la propagation d'une onde dans ce milieu.

Dans la deuxième partie de la thèse, nous nous intéressons à des problèmes dans lesquels les échelles ne sont bien séparées les unes des autres que dans certaines régions du domaine. De tels problèmes sont fréquents et on les trouve, par exemple, dans des matériaux composites dont les propriétés effectives peuvent être capturées en assumant une microstructure (localement) périodique, mais qui n'est peut-être pas valide autour de défauts présents dans le matériau. Dans notre modélisation, nous supposons que la plus petite échelle est toujours discrétisée au niveau continu, mais notons que dans certaines applications, des échelles atomistiques doivent être considérées. Notre méthode est basée sur une décomposition du domaine en une famille de sous-domaines qui se chevauchent. Des contrôles virtuels (d'interface) sont introduits comme conditions inconnues aux bords et agissent comme traces ou comme flux. Notre méthode est reformulée comme un problème de minimisation sous contraintes, avec pour contraintes des équations d'état. Ici, les contrôles optimaux de deux sous-domaines qui s'intersectent sont trouvés en minimisant la différence entre les deux modèles soit sur le chevauchement soit le bord du chevauchement. La méthode d'optimisation est ensuite donnée dans sa forme discrète et couple la méthode d'éléments finis multi-échelles hétérogène avec la MEF. L'existence et l'unicité de la méthode, dans sa forme continue et discrète, sont prouvées, et nous donnons une analyse a priori de notre méthode.

Mots clefs : problèmes multi-échelles, méthodes multi-échelles hétérogènes, homogénéisation, problèmes d'élasticité linéaire, méthodes globales à locales, décomposition de domaines, méthodes basées sur l'optimisation.

Contents

Remerciements	i
Abstracts (English/Français)	iii
Notations	xi
1 Introduction	1
I Finite element multiscale method for heterogeneous linear elastic waves	7
2 Multiscale methods for elliptic problems	11
2.1 Standard numerical methods for elliptic problems	12
2.1.1 Finite element method (FEM)	13
2.1.2 Discontinuous finite element method (DG-FEM)	15
2.2 Homogenization theory	17
2.2.1 Periodic homogenization	18
2.3 A brief review of numerical homogenization methods	19
2.3.1 Numerical methods based on global extractions of the small scales . . .	20
2.3.2 Numerical methods based on homogenization	23
2.4 Finite Element Heterogeneous Multiscale Method	26
2.4.1 Discontinuous Galerkin FE-HMM	32
2.5 A priori error analysis for the (DG-)FE-HMM	33
2.6 Numerical results	37
2.7 Summary	37
3 Multiscale method for the wave equation in linear elastic heterogeneous media	41
Heterogeneous multiscale method for the linear elasticity problem	43
3.1 Homogenization of linear elasticity problems and basic results	44
3.2 FE-HMM for linear elasticity	45
3.3 A priori error analysis	48
3.4 Numerical results for the FE-HMM applied to linear elasticity problems	53
3.5 Summary	57
Heterogeneous multiscale method for the wave equation in a linear elastic medium	59
3.6 FE-HMM for the wave equation in a linear elastic medium	61
3.7 A priori error analysis	62
3.8 Numerical examples	67

Contents

3.8.1	Periodic and locally periodic tensor	68
3.8.2	Horizontally layered material	71
3.8.3	Arbitrarily heterogeneous media.	73
3.9	Summary	78
4	Conclusion and outlook of Part I	83
II	An optimization based coupling method for highly heterogeneous multi-scale elliptic PDEs	85
5	Homogenization based global to local methods	89
5.1	Caccioppoli inequalities	90
5.2	L^2 global to local projection method	91
5.3	Goal-oriented adaptive method	92
5.4	Summary	94
6	Optimization based coupling method	95
6.1	Literature overview	96
6.2	Optimal control of partial differential equations	97
	Continuous optimization based method	99
6.3	Existence and uniqueness	103
6.4	Optimality systems	109
6.4.1	À la Lions	110
6.4.2	Lagrangian functional	110
6.4.3	Transposition method	112
6.5	A priori error analysis	115
6.5.1	A priori error estimates to the fine scale solver in ω^+	116
6.5.2	A priori error estimates to the reconstructed coarse scale solver in $\Omega \setminus \omega^+$	120
6.6	Summary	121
	Fully discrete optimization based coupling method	123
6.7	Numerical method for the optimization based coupling.	124
6.7.1	Numerical method for the fine scale solver	124
6.7.2	Discontinuous Galerkin (DG) method for the coarse scale problem	125
6.7.3	Numerical Algorithm	127
6.8	Discrete inequalities	130
6.9	Well-posedness of the discrete coupling method	134
6.10	Fully discrete error estimates	137
6.10.1	A priori error estimates in the fine scale region	140
6.10.2	A priori error estimates in the scale separated region	141
6.11	Numerical experiments	145
6.11.1	Computational costs of the methods	146
6.11.2	Influence of ε and τ in the convergence rates	146
6.11.3	Influence of micro HMM error	150
6.11.4	A domain with a crack	150
6.11.5	Singular source term	152

6.11.6 A domain with a defect	154
6.12 Summary	157
7 Improvement of meshing and coupling strategy	159
7.1 Optimization based coupling with minimization on $\Gamma_1 \cup \Gamma_2$	160
7.1.1 A priori error analysis	163
7.2 Fully discrete coupling method	167
7.2.1 Numerical method for the fine scale problem.	167
7.2.2 Numerical method for the coarse scale problem.	168
7.2.3 Numerical Algorithm	169
7.2.4 Fully discrete error estimates	170
7.3 Partitions used in the numerical coupling method	171
7.4 Numerical experiments	172
7.4.1 Comparison of matching and non-matching grids on the overlap	173
7.4.2 Minimization with interface controls	177
7.4.3 Minimization with interface controls on non-matching grids	181
7.5 Summary	182
8 Conclusion and outlook of Part II	185
Bibliography	189

Notations

d	spatial dimension
ε	one or several microscopic scales
$C > 0$	generic constant whose value can change at any occurrence
Y	d -dimensional unit cube $(0, 1)^d$
Ω	spatial domain $\Omega \subset \mathbb{R}^d$ with $\varepsilon \ll \Omega $
O, D	sets $O, D \subset \Omega$
$ O $	measure of the set O
B	Banach space
$[0, T]$	finite time interval $T \in \mathbb{R}, T > 0$
e_i	vectors of the canonical basis of $\mathbb{R}^d, i = 1, \dots, d$
$ r $	Euclidean norm of the vector $r \in \mathbb{R}^d$
$\ a\ _F$	Frobenius norm for matrice $a \in \mathbb{R}^{d \times d}, \ a\ _F = \sqrt{\sum_{i,j} a_{ij} ^2}$
I_d	identity matrix in $\mathbb{R}^{d \times d}$

Functions spaces

$\mathcal{C}^k(O)$ k -times continuously differentiable functions $f : O \rightarrow \mathbb{R}, O \subset \mathbb{R}^d, k \in \mathbb{N}$

$\mathcal{C}_{per}^\infty(Y)$ subset of $\mathcal{C}^\infty(\mathbb{R}^d)$ of periodic functions on Y

$W^{k,p}(\Omega)$ usual Sobolev space on $\Omega \subset \mathbb{R}^d$ with $k \in \mathbb{N}$ and $1 \leq p \leq \infty$

$H^k(\Omega)$ Sobolev space $W^{k,2}(\Omega)$ on $\Omega \subset \mathbb{R}^d, k \in \mathbb{N}$

$H^k(\mathcal{T}_H)$ broken Sobolev space $H^k(\mathcal{T}_H) = \{v \in L^2(\Omega) \mid v \in H^k(K), \forall K \in \mathcal{T}_H\}$
for a partition \mathcal{T}_H of $\Omega, k \in \mathbb{N}$

$H_0^1(\Omega)$ subspace of $H^1(\Omega)$ with vanishing trace on the boundary of Ω

$H_{loc}^1(D)$ defined as $H_{loc}^1(\omega_1) = \{u \in H^1(O) \mid \text{for any open set } O \text{ with } \overline{O} \subset D\}$

$H_{per}^1(Y)$ defined as the closure of $\mathcal{C}_{per}^\infty(Y)$ for the H^1 norm

$W_{per}^1(Y)$ defined as $W_{per}^1(Y) = \{v \in H_{per}^1(Y) \mid \int_Y v dy = 0\}$

$L^p(0, T; B)$ Bochner space of functions $w : (0, T) \rightarrow B$ equipped with the norm

$$\|w\|_{L^p(0,T;B)} = \begin{cases} (\int_0^T \|w(t)\|_B^p dt)^{1/p}, & \text{for } 1 \leq p < \infty \\ \text{ess sup } \|w\|_B, & \text{for } p = \infty \end{cases}$$

1 Introduction

In seismic modeling engineers are often interested in the macroscopic behavior of the solid undergoing deformation; for example, it can be the shaking and displacement of the Earth's surface as a response to an earthquake. If small heterogeneities are present in the solid, e.g., if the solid is a fractured medium or if it is made of a layered material, engineers need to have a reliable method that includes the heterogeneities in the macroscopic outcome. This is an example of a multiscale problem, or multiphysic problem, as it involves physical laws at different levels. Indeed, the propagation of the earthquake happens at a macroscopic level, whereas the heterogeneities are microscopic. In Figure 1.1, we illustrate three heterogeneous multiscale media. Multiphysics or multiscale problems regroup a wide range of problems arising in physical science, biology, chemistry or geoscience.

One scale methods, that are traditionally used in modeling, might produce inaccurate results or might even be impossible to process numerically. Assuming that a microscopic model is available, using all the microscopic information to model the problem is in practice not feasible. Furthermore, considering a macroscopic model without including the small heterogeneities is computationally efficient but might give an unrealistic outcome. One scale modeling is therefore not a suitable option. As stated by E in [53], "[i]t would be nice to have a strategy that combines the efficiency of macroscale models and the accuracy of microscale models." Multiscale modeling proposes to consider simultaneously the models at the different levels in order to describe the effective properties without resolving all the small features of the medium. This offers a better understanding of physical processes by including the different levels of details in "on-the-fly" computations.

Numerical multiscale methods. In the past decades, various methods have been developed to resolve partial differential equations (PDEs) with highly heterogeneous coefficients, see Chapter 2 for a brief review. Multiscale methods can be characterized in two categories. In the first one, we regroup methods that are based on a global extraction of the fine scale information, and where all the fine scale information are used and processed. Such methods are therefore computationally expensive, but they do not require structural assumptions on the fine scales.

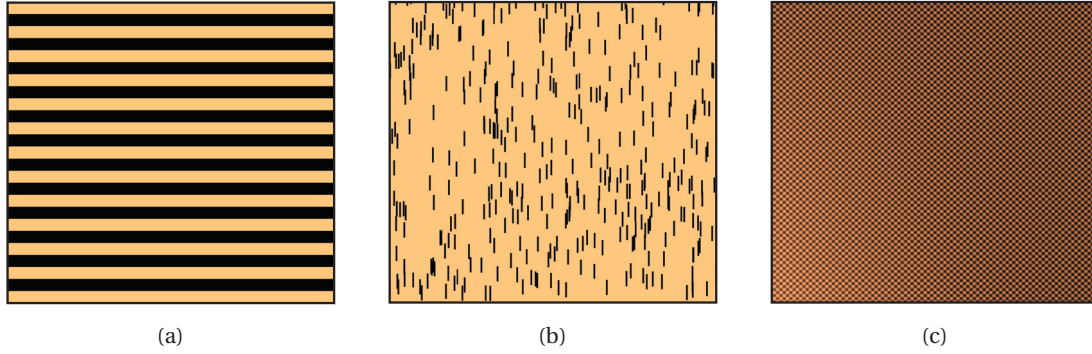


Figure 1.1 – Illustration of multiscale media, (a) layered, (b) fractured, and (c) locally periodic media.

The second class of numerical methods is closely related to the homogenization theory. They are based on representative local computations and on a global to local downscaling. As a consequence, they are much more efficient than the methods in the first class, but they rely on a clear separation of scales.

In this thesis, we focus on multiscale problems belonging to the class of homogenization problems. Homogenization is the tool for finding the best averaging process that can be used to model the macroscopic description of the multiscale system. Let $\mathcal{L}^\varepsilon(u^\varepsilon) = f$ be a heterogeneous multiscale PDE, and when ε goes to zero, the homogenization theory ensures that u^ε converges (in a weak sense) towards u^0 , the solution of a homogenized PDE $\mathcal{L}^0(u^0) = f$. The coefficients of the homogenized operator \mathcal{L}^0 are independent of the small scale ε but are, however, often unknown.

In Part I, we developed a numerical multiscale method for the wave equation in a linear elastic medium with highly heterogeneous coefficients. The method is based on the framework of the heterogeneous multiscale method (HMM) developed by E and Engquist [54] (see [55, 8] for a review). The HMM can be summed up in two components,

an effective system: the global behavior of a heterogeneous system should be described by an effective macroscopic system. The homogenized problem $\mathcal{L}^0(u^0) = f$ is a good effective model, but as the data of \mathcal{L}^0 are generally unknown, one needs to use a numerical homogenized method for the problem $\mathcal{L}^0(u^0) = f$.

a macro to micro modeling: the unknown macroscopic data used in the effective numerical method are needed at macro nodes. Micro simulations are conducted in small sampling domains around the macro nodes, using the available microscopic problems \mathcal{L}^ε . The macroscopic data are recovered from the micro simulations using averaging.

These two steps are based on the assumptions that the scales are well separated; that is when the fine scale information can be described by a quantity ε and the ratio of the length scale of the macroscopic behavior of the problem to the size ε is large. The HMM framework has

been applied to, amongst others, multiscale elliptic or parabolic or Stokes equations, or to the multiscale wave equation and the linear elastic equation. In Part I, we focus on the wave propagation, for short time, in a heterogeneous linear elastic medium with well-separated scales.

In Part II, we consider problems that are between scale separation and non-scale separation, and, aim at designing a new multiscale method for elliptic problems with non-separated small scales. This is the case, for example, when the coefficients oscillate at several scales that are indistinguishable. Figure 1.2 illustrates two tensors, one with scale separation and the other without scale separation. Such problems arise in many situations, for example in composite material design, in which the effective model might not be valid near a crack or a defect. Our aim is to couple a homogenized solver with a heterogeneous solver using domain decomposition, minimization, and optimization techniques. The HMM framework is used in regions with scale separation, whereas in regions without scale separation, a fine scale solver is chosen.

Main contributions. In this thesis, we design and analyze a numerical multiscale method, based on the HMM framework, for two types of multiscale problems.

In **Part I**, we derive and analyze the finite element heterogeneous multiscale method (FE-HMM), for the wave equation in a linear elastic medium with highly heterogeneous coefficients with explicit separation of scales. In Chapter 2, we review the FE-HMM for elliptic multiscale problems derived in [3] and state the a priori error analysis. In Chapter 3, we consider, at first, a highly heterogeneous linear elasticity problem and extend the FE-HMM with piecewise macro and micro finite elements given in [2] to finite element spaces of order $p \geq 1$. A priori error estimates are given and numerical experiments are proposed to assess the convergence rates. In the second part of Chapter 3, we are interested in the propagation of a wave in a heterogeneous linear elastic medium for short times. We extend the FE-HMM for the heterogeneous wave equation given by Abdulle and Grote in [9, 10] to heterogeneous linear elastic waves. A priori error estimates are proved, and we verify the sharpness of the error bounds through various numerical experiments. A conclusion and outlook are given at the end of Part I.

In **Part II**, we develop a numerical multiscale method, based on optimization techniques, for problems with and without scale separation. Our model problem is an elliptic problem with highly heterogeneous coefficients. We address the following issue.

First issue. When the fine scales are not well separated in some subregions of the computational domain, no explicit small lengths ε are available to characterize the highly heterogeneous structure of the media. Therefore, the HMM procedure, which needs the fine scale information only inside sampling domains of (small) size δ , is similar (in efficiency and accuracy) to a fine scale solver. Indeed, without the characteristic length ε , we often set δ to the size of the partition used in the discretization. As a

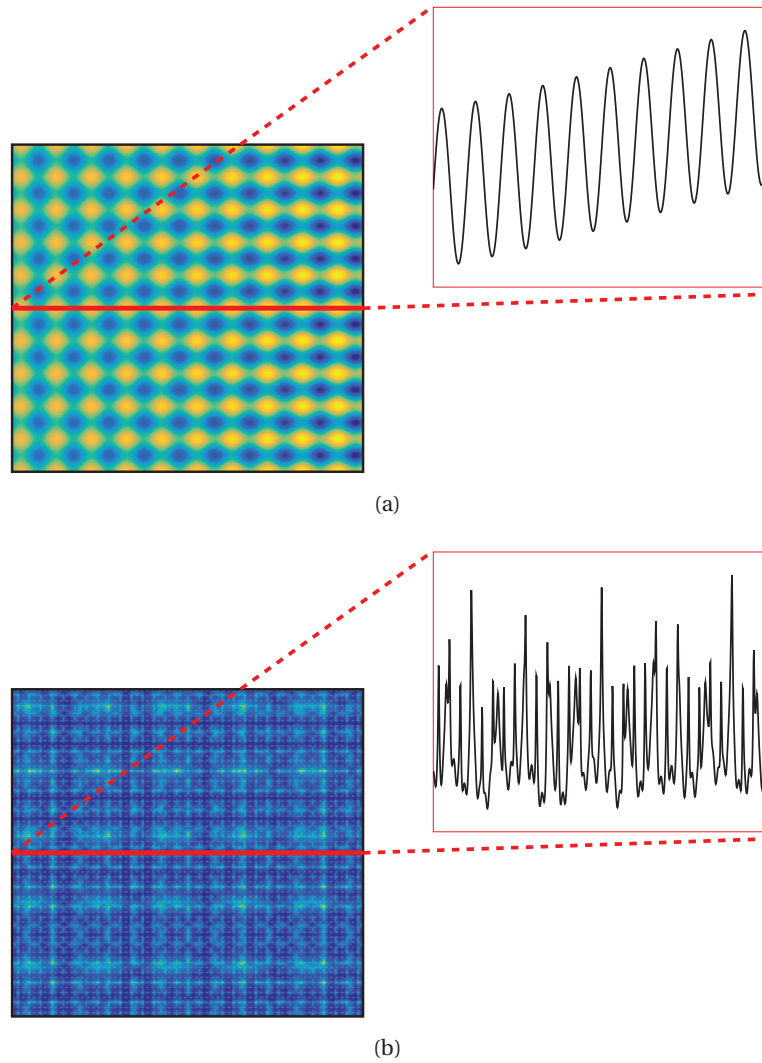


Figure 1.2 – (a) Locally periodic tensor with scale separation, and (b) tensor without scale separation.

consequence, all fine scales are needed and processed.

In Chapter 5, we mention two global to local methods that give good a H^1 approximation of the fine scale solution in subregions of interest. We review the L^2 global to local projection method derived by Babuška and Lipton in [27] and the goal-oriented adaptive method given by Oden and Vemaganti in [95]. Nevertheless, in such methods, a second issue can be addressed.

Second issue. Global to local methods rely on the existence of a good precomputed L^2 approximation of the fine scale solution. The computation of this L^2 approximation can be computationally expensive and the quality of the numerical solution, obtained from the global to local method, depends on the quality of the L^2 approximation.

In Chapter 6, we consider an elliptic problem with a highly heterogeneous tensor that exhibits scale separation only in some regions of the computational domain. In the regions with scale separation, homogenization methods can be used, whereas in the regions without explicit separation of scales, a fine scale solver is preferred. We propose a multiscale method that couples a homogenized PDE with a heterogeneous PDE. The method is inspired by virtual control methods and overlapping domain decomposition [71, 87], and is formulated as a minimization problem under constraints. Such ideas have appeared earlier in the literature for coupling of different type of partial differential equations [65, 49, 46] and for atomistic-to-continuum coupling [97, 98]. Chapter 6 is based on the minimization problem over the overlapping regions using the L^2 norm. We prove existence and uniqueness of the optimal solution of the coupling, following Lions' theory of optimal controls [85], and derive a priori error estimates for the method. Using classical periodic correctors in the regions with scale separation, we prove strong convergence in H^1 of order $\mathcal{O}(\varepsilon)$ in the regions without scale separation, and of order $\mathcal{O}(\varepsilon^{1/2})$ for the regions with scale separation.

The fully discrete optimization based method is given in the second part of Chapter 6. We couple the DG-FE-HMM with the FEM, prove the well-posedness of the method and give the fully discrete a priori error analysis. The continuous and discrete analysis of the method are based on Caccioppoli inequalities and on a strong version of the Cauchy–Schwarz inequality. We give explicit convergence rates with respect to the macro and micro mesh sizes, and propose numerical examples to verify the convergence rates. Further, we compare our coupling method with other coupling strategies. In contrast to [27] and [95], our method does not rely on the existence of a good precomputed L^2 approximation of the heterogeneous solution.

In Chapter 7, we propose numerical improvements of our method. The first numerical improvement is to use another cost function in the minimization problem.

Minimization over the boundary. We consider an optimization based coupling method with a minimization over the boundary of the overlapping regions instead of the minimization in the overlapping regions. This reduces the number of degrees of freedom in the optimization based problem.

Chapter 1. Introduction

We prove the well-posedness and derive a priori error estimates for the coupling method based on this new minimization problem.

The second improvement is done in the partitions used in the fine and coarse scale solvers. At first, in Chapter 6, we consider a fine and a coarse scale partitions with identical finite elements in the overlapping regions. As the mesh size in the fine scale regions should be small to resolve the heterogeneities of the problem, the number of degrees of freedom in the overlapping regions and the computational cost of the coupling method become large. Indeed, we recall that in the overlap, the FE-HMM is used and thus that we need to resolve fine scale problems in sampling domains in each element of the overlapping region. This is computationally expensive as the fine scale are resolved everywhere in the domain.

Interpolation in the overlap. We use an interpolation in the overlapping between the fine and coarse partitions used to resolve the heterogeneous and homogenized PDEs respectively.

Through various numerical examples, we show that the computational costs of the method is significantly reduced without affecting the efficiency of the method. Part II ends with a conclusion and an outlook of perspective for future research.

Finite element multiscale method for heterogeneous linear elastic waves

Part I

In the first part of the thesis, we develop a numerical homogenization method for the wave equation in a heterogeneous linear elasticity medium

$$\partial_{tt} u^\varepsilon - \operatorname{div}(a^\varepsilon(x) : e(u^\varepsilon)) = f,$$

where e is the linearized strain tensor, u^ε corresponds to the displacement of the domain, and where the tensor a^ε is highly heterogeneous and varies at the small scale $\varepsilon > 0$. Solving such problems with standard methods, such as the finite element method or the finite volume method, is often not feasible. Indeed, a good approximation of u^ε requires that the fine scales of the problem are fully resolved by the spatial mesh, which is computationally expensive. Numerical homogenization methods should be used instead.

Outline of Part I. The outline of Part I is as follows.

In **Chapter 2**, we consider an elliptic heterogeneous multiscale problem $-\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) = f$, and give an introduction to the multiscale methods available to resolve such problems. Without attempting to be exhaustive, we recall two important approaches to resolve multiscale problems. The chapter starts with the standard finite element method (FEM) and the discontinuous Galerkin finite element method (DG-FEM), which are efficient tools to solve single-scale problems. Then, the theory of homogenization is given, with details about periodic homogenization, and an overview of different multiscale methods is proposed. We mention methods based on coarse oscillatory basis functions that encode the high variations of the data in the multiscale problem; we review the multiscale finite element method (MsFEM) and the localized orthogonal decomposition method (LOD). Then, we mention numerical methods that are based on homogenization theory, and dedicate one section to the finite element heterogeneous multiscale method (FE-HMM), which is the numerical multiscale method used in this thesis. Then, its discontinuous counterpart, the DG-FE-HMM, is given, and their a priori error analysis are recalled. Numerical experiments and limitations of the FE-HMM (DG-FE-HMM) are presented, as a step towards global to local methods presented in the Part II of this thesis.

In **Chapter 3**, we consider a wave propagation through an heterogeneous linear elasticity medium $\partial_{tt} u^\varepsilon - \operatorname{div}(a^\varepsilon(x) : e(u^\varepsilon)) = f$, and we separate the chapter in two parts. In the first part we give the finite element heterogeneous multiscale method applied to a linear multiscale static elasticity problem $-\operatorname{div}(a^\varepsilon(x) : e(u^\varepsilon)) = f$. We recall the a priori analysis and give numerical experiments to assess the convergence rates. In the second part, we consider the wave equation in a domain filled with an heterogeneous linear elasticity medium. We then give the FE-HMM and the a priori analysis. The chapter ends with numerical experiments applied to various examples.

Publications. The first part of Chapter 3 about linear elasticity is based on [A. Abdulle, O. Jecker, ENUMATH 2015 Proceedings, 2016].

2 Multiscale methods for elliptic problems

In this chapter, we review the heterogeneous multiscale methods (HMM) based on the finite element method, using an elliptic multiscale equation as a model problem (see amongst others [54, 56, 20, 3]). We relate the HMM method to other multiscale methods.

Let $\Omega \subset \mathbb{R}^d$, with $d = 1, 2, 3$, be a convex polygonal domain with Lipschitz boundary Γ , and consider an elliptic model problem

$$\begin{aligned} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, & \text{in } \Omega, \\ u^\varepsilon &= 0, & \text{on } \Gamma, \end{aligned} \quad (2.1)$$

where the right-hand side f is a source term in $L^2(\Omega)$, and the tensor a^ε describes the multiscale structure of the problem. The index $\varepsilon > 0$ represents the small scale of the problems and determines the rapid variations in the tensor a^ε . We assume that the tensor a^ε is in $(L^\infty(\Omega))^{d \times d}$ and that a^ε is uniformly elliptic and bounded, i.e.,

$$\exists 0 < \lambda \leq \Lambda : \lambda |\xi|^2 \leq a^\varepsilon(x) \xi \cdot \xi, \text{ and } |a^\varepsilon(x) \xi| \leq \Lambda |\xi|, \quad \forall \xi \in \mathbb{R}^d, \text{ a.e. } x \in \Omega, \varepsilon > 0. \quad (2.2)$$

A weak solution u^ε of the model problem (2.1) is obtained from the variational formulation

$$B^\varepsilon(u^\varepsilon, w) = F(w), \quad \forall w \in H_0^1(\Omega), \quad (2.3)$$

where the bilinear form $B^\varepsilon : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ and the right-hand side $F : H^1(\Omega) \rightarrow \mathbb{R}$ are given by

$$B^\varepsilon(v, w) = \int_{\Omega} a^\varepsilon(x) \nabla v \cdot \nabla w \, dx \quad \text{and} \quad F(w) = \int_{\Omega} f w \, dx. \quad (2.4)$$

Thanks to the Lax–Milgram lemma, the problem (2.3) admits a family $\{u^\varepsilon\}$, indexed by the microscopic scale ε , of unique solutions which are bounded by the data of the problem (2.1); i.e.,

$$\|u^\varepsilon\|_{H^1(\Omega)} \leq C \|f\|_{L^2(\Omega)},$$

where the constant C is independent of ε (but depends on the coercivity constants of a^ε (2.2)).

Chapter 2. Multiscale methods for elliptic problems

The size of the microstructure prevents us to solve problem (2.1) with a numerical scheme such as the finite element method (FEM) or the finite volume method. Indeed, assuming that $u^\varepsilon \in H^{l+1}(\Omega)$, $l > 0$, and that u^H is an approximation of u^ε , obtained by a p -order FEM on a discretization of Ω with mesh size H ; classical FEM theory (see Ciarlet in [41]) gives a standard a priori error estimate

$$\|u^\varepsilon - u^H\|_{H^1(\Omega)} \leq CH^s |u^\varepsilon|_{H^{s+1}(\Omega)}, \quad s = \min\{l, p\}.$$

However, it holds $\|u^\varepsilon\|_{H^{s+1}(\Omega)} \leq C\varepsilon^{-s} \|f\|_{H^{s-1}(\Omega)}$, leading to an a priori error estimate bounded by $C(H/\varepsilon)^s$, which gives a good approximation of u^ε only if $H \ll \varepsilon$. As the size of the multiscale structure is microscopic, it leads to a discretization with a very large number of degrees of freedom, and thus the cost of such methods is prohibitive.

Numerical methods based on homogenization theory [54, 3] give a good L^2 approximation of the fine scale solution u^ε , i.e., such methods numerically capture the global behavior of u^ε . Techniques exist to add the missing fine scale information to the coarse scale numerical solution and obtain a good approximation of u^ε in the H^1 norm, as well.

Outline. In Section 2.1, we give standard numerical methods for elliptic problems with a single-scale tensor. In Section 2.2, we briefly give the homogenization theory with a focus on periodic homogenization. A review of some numerical homogenization methods is presented in Section 2.3. The numerical homogenization method used in this thesis is the finite element heterogeneous multiscale method (FE-HMM) and is the subject of Section 2.4. A priori error estimates and a numerical example are given in Sections 2.5 and 2.6, respectively.

2.1 Standard numerical methods for elliptic problems

Standard numerical methods for elliptic problems play a major role in the design of a numerical heterogeneous multiscale scheme. In this section, we recall the standard finite element method (FEM) [33, 42, 41] and its discontinuous version (DG-FEM) [76].

Let us lose the multiscale nature of the elliptic model problem (2.1) and consider a single-scale elliptic problem: find $u \in H^1(\Omega)$ such that

$$\begin{aligned} -\operatorname{div}(a(x)\nabla u) &= f, & \text{in } \Omega, \\ u &= 0, & \text{on } \Gamma, \end{aligned} \tag{2.5}$$

where $f \in L^2(\Omega)$ and $a \in (L^\infty(\Omega))^{d \times d}$ verifies (2.2).

An approximation of the solution of problem (2.5) is constructed via the Galerkin method, which consists in finding a solution, in a finite dimensional subspace of $H_0^1(\Omega)$, of a discrete problem similar to the problem (2.5). In order to apply the Galerkin method, one needs to construct finite dimensional subspaces V_H of the space $H_0^1(\Omega)$ (which is the space of the solutions of problem (2.5)). The inclusion $V_H \subset H_0^1(\Omega)$ can be violated and the resulting finite

element method will be non-conforming (DG-FEM, see 2.1.2), as opposed to the conforming finite element method (FEM, see 2.1.1).

2.1.1 Finite element method (FEM)

Let \mathcal{T}_H be a partition of Ω , where Ω is divided into a finite number of elements K ; the mesh size $H > 0$ is given by the maximum diameter of the elements, i.e., $H = \max_{K \in \mathcal{T}_H} H_K$. The family $\{\mathcal{T}_H\}$ of partitions is assumed to be admissible and shape-regular [41],

(T1) *admissible*: $\Omega = \bigcup_{K \in \mathcal{T}_H} K$, and the intersection between two elements is either empty, a vertex, or a common face;

(T2) *shape-regular*: there exists $\sigma > 0$ such that $\frac{h_K}{\rho_K} \leq \sigma$, where ρ_K is the diameter of the larger circle contained in the element K , for all $K \in \mathcal{T}_H$ and for all $\mathcal{T}_H \in \{\mathcal{T}_H\}$.

For each partition \mathcal{T}_H in the family $\{\mathcal{T}_H\}$, a finite element space of degree $p \geq 1$ can be defined by

$$V_0^p(\Omega, \mathcal{T}_H) = \{v^H \in H_0^1(\Omega) \mid v^H|_K \in \mathcal{R}^p(K), \forall K \in \mathcal{T}_H\}, \quad (2.6)$$

where the space $\mathcal{R}^p(K)$ denotes the space $\mathcal{P}^p(K)$ of polynomials on K of total degree at most p if the element K is simplicial, or the space $\mathcal{Q}^p(K)$ of polynomials on K of degree at most p in each variable, if the element K is rectangular.

The numerical approximation of u , the solution of problem (2.5), satisfies the discrete problem: find $u^H \in V_0^p(\Omega, \mathcal{T}_H)$ such that

$$B(u^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (2.7)$$

where the bilinear form $B : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ and the right-hand side $F : H^1(\Omega) \rightarrow \mathbb{R}$ are given by

$$B(v, w) = \int_{\Omega} a(x) \nabla v \cdot \nabla w dx \quad \text{and} \quad F(w) = \int_{\Omega} f w dx. \quad (2.8)$$

The discrete problem (2.7) is well-posed thanks to the Lax–Milgram lemma, and explicit convergence rates between u^H and u can be derived following classical results [41, Ch. 3]. Assuming that the data are smooth enough to have $u \in H^{p+1}(\Omega)$, it holds

$$\|u - u^H\|_{H^1(\Omega)} \leq CH^p |u|_{H^{p+1}(\Omega)}, \quad \|u - u^H\|_{L^2(\Omega)} \leq CH^{p+1} |u|_{H^{p+1}(\Omega)}, \quad (2.9)$$

where the constant C is independent of H .

Finite element method with numerical quadrature

In practice, the approximations of the integrals in the bilinear form B and right-hand side F defined in (2.8), require a numerical quadrature formula. We state results given in [41, 42].

Let $\hat{K} \subset \mathbb{R}^d$ be a reference element, either simplicial or quadrilateral, and consider a continuous diffeomorphism $F_K : \hat{K} \rightarrow K$ that denotes the parametrization of each element $K \in \mathcal{T}_H$. A quadrature formula (QF) on \hat{K} is the set of couples $\{\hat{x}_j, \hat{\omega}_j\}_{j=1}^J$ composed of integration nodes \hat{x}_j and weights $\hat{\omega}_j$, with $J \in \mathbb{N}$. When using the FEM with a quadrature formula, there is no guarantee that the approximation converges to the exact solution with the standard convergence rates given in (2.9). However, with suitable assumptions, one could recover the same rates as the FEM with exact integration [41],

- (Q1) $\hat{\omega}_j > 0$, for $j = 1, \dots, J$, and $\sum_{j=1}^J \hat{\omega}_j |\nabla \hat{p}(\hat{x}_j)|^2 \geq \hat{\alpha} \|\nabla \hat{p}\|_{L^2(\hat{K})}^2$, for all $\hat{p} \in \mathcal{R}^p(\hat{K})$ and some $\hat{\alpha} > 0$;
- (Q2) $\int_{\hat{K}} \hat{p}(\hat{x}) d\hat{x} = \sum_{j=1}^J \hat{\omega}_j \hat{p}(\hat{x}_j)$, for all $\hat{p} \in \mathcal{R}^\sigma(\hat{K})$ with $\sigma = \max\{2p-2, p\}$ if \hat{K} is simplicial, or $\sigma = \max\{2p-1, p+1\}$ if \hat{K} is a rectangle.

For simplicial FE, assumption (Q2) implies (Q1) with an equality and $\hat{\alpha} = 1$. The quadrature formula over the reference element \hat{K} induces a quadrature formula $\{x_{j,K}, \omega_{j,K}\}_{j=1}^J$ over K with $x_{j,K} = F_K(\hat{x}_j)$ and $\omega_{j,K} = \hat{\omega}_j \det(\partial F_K)$, for $j = 1, \dots, J$.

Let u_{QF}^H be the numerical approximation of u the solution of problem (2.5) using the FEM where the integrals are computed using a quadrature formula; i.e., $u_{QF}^H \in V_0^p(\Omega, \mathcal{T}_H)$ satisfies

$$B^H(u_{QF}^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} f(x_{j,K}) w^H(x_{j,K}), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H),$$

where the bilinear form $B^H : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ is defined by

$$B^H(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} a(x_{j,K}) \nabla v^H(x_{j,K}) \cdot \nabla w^H(x_{j,K}),$$

where we assume that f and a can be evaluated at the quadrature points $x_{j,K}$. The well-posedness follows from the Lax–Milgram lemma together with assumption (Q1), see [41, Ch. 4. §4.1.]. Let assumptions (Q1) and (Q2) hold and assume that $u \in H^{p+1}(\Omega)$, and $a \in (W^{p+m,\infty}(\Omega))^{d \times d}$, $m = 0, 1$. Then, the convergence rates are

$$\begin{aligned} \|u - u_{QF}^H\|_{H^1(\Omega)} &\leq CH^p |u|_{H^{p+1}(\Omega)}, & \text{if } m = 0, \\ \|u - u_{QF}^H\|_{L^2(\Omega)} &\leq CH^{p+1} |u|_{H^{p+1}(\Omega)}, & \text{if } m = 1, \end{aligned} \tag{2.10}$$

where the constant C is independent of H .

2.1.2 Discontinuous finite element method (DG-FEM)

Flexibility in the mesh and local conservation properties are sometimes required in many problems, and this is made possible by the use of discontinuous approximations in the finite element spaces. The first discontinuous Galerkin method was introduced by Reed and Hill [76] for hyperbolic equations and the first analysis was given by Lesaint and Raviart [82] for linear problems. It has then be adapted to elliptic and parabolic equations leading to a number of DG methods; we cite the original method for Navier–Stokes equation by Bassi and Rebay [29], or for diffusion problems by Brezzi et al [35]. In this thesis, we focus on interior penalty discontinuous Galerkin FEM [51, 25], where the FE spaces are allowed to be spaces of discontinuous polynomials and where the continuity of the approximation is made by including penalty terms. A review of discontinuous methods is proposed in [26].

In here, for simplicity, we restrict to simplicial finite elements, but notice that the discontinuous FEM can be extended to quadrilateral FE.

Let $\{\mathcal{T}_H\}$ be a shape-regular (T2) family of triangulations over Ω , composed of simplicial elements K with diameter h_K , such that $\Omega = \bigcup_{K \in \mathcal{T}_H} K$ and $H = \max_{K \in \mathcal{T}_H} h_K$. For each triangulation \mathcal{T}_H in $\{\mathcal{T}_H\}$, we denote by E the set of $d - 1$ dimensional interfaces of the mesh \mathcal{T}_H ; it regroups the edges for $d = 2$ or the faces for $d = 3$. For two neighboring finite elements K_+, K_- in the triangulation \mathcal{T}_H , an interior interface e is in the set E if $\bar{e} = \bar{K}_+ \cap \bar{K}_-$. Similarly, for a finite element $K \in \mathcal{T}_H$, a boundary interface e is in E if $\bar{e} = \bar{K} \cap \bar{\Gamma}$. The mesh allows for hanging nodes, thus we assume that E is composed of the smallest common interface between two neighboring elements. For any piecewise smooth functions v , we define the jump and average by

$$\begin{aligned} \llbracket v \rrbracket &= v_+ n_{K_+} + v_- n_{K_-}, \quad \text{and } \{v\} = \frac{1}{2}(v_+ + v_-), \quad \text{for interior interfaces,} \\ \llbracket v \rrbracket &= v n, \quad \text{and } \{v\} = v, \quad \text{for boundary interfaces,} \end{aligned} \tag{2.11}$$

where n, n_{K_\pm} , denote the unit outward normal vectors to Γ and K_\pm respectively.

For each triangulation \mathcal{T}_H in the family $\{\mathcal{T}_H\}$, a FE space of degree $p \geq 1$ can be defined as

$$V^p(\Omega, \mathcal{T}_H) = \{v^H \in L^2(\Omega) \mid v^H|_K \in \mathcal{P}^p(K), \forall K \in \mathcal{T}_H\},$$

where $\mathcal{P}^p(K)$ is the space of polynomials on K of total degree at most p . As opposed to the conforming FE, it holds $V^p(\Omega, \mathcal{T}_H) \not\subset H_0^1(\Omega)$. We can however, define a piecewise Sobolev space

$$H^l(\mathcal{T}_H) = \Pi_{K \in \mathcal{T}_H} H^l(K), \quad l \geq 1,$$

and then obtain $V^p(\Omega, \mathcal{T}_H) \subset H^1(\mathcal{T}_H)$.

The interior penalty (IP) discontinuous Galerkin FE method reads: find $u^H \in V^p(\Omega, \mathcal{T}_H)$ such that

$$B_{DG}(u^H, w^H) = F(w^H), \quad \forall w^H \in V^p(\Omega, \mathcal{T}_H), \tag{2.12}$$

Chapter 2. Multiscale methods for elliptic problems

where the right hand side F is given by equation (2.8), and the bilinear form $B_{DG} : H^1(\mathcal{T}_H) \times H^1(\mathcal{T}_H) \rightarrow \mathbb{R}$ is defined by

$$\begin{aligned} B_{DG}(v^H, w^H) &= \int_{\Omega} a(x) \nabla v^H \cdot \nabla w^H dx + \sum_{e \in E} \int_e \mu_e \llbracket v^H \rrbracket \llbracket w^H \rrbracket ds \\ &\quad - \sum_{e \in E} \int_e (\{a(x) \nabla v^H\}) \llbracket w^H \rrbracket ds, \end{aligned}$$

where the second term is called the interior penalty term and μ_e , the penalty weighting functions, that penalize the jumps of the functions at the interface e . The penalty weighting functions are given by $\mu_e = \alpha h_e^{-1}$, with $\alpha > 0$ and h_e is the size of the interface e .

The bilinear form B_{DG} can be made symmetric by adding a term $\{a(x) \nabla w^H\} \llbracket v^H \rrbracket$ in the last integral; this leads to the symmetric IP method with a bilinear form given by

$$\begin{aligned} B_{DG}(v^H, w^H) &= \int_{\Omega} a(x) \nabla v^H \cdot \nabla w^H dx + \sum_{e \in E} \int_e \mu_e \llbracket v^H \rrbracket \llbracket w^H \rrbracket ds \\ &\quad - \sum_{e \in E} \int_e (\{a(x) \nabla v^H\} \llbracket w^H \rrbracket + \{a(x) \nabla w^H\} \llbracket v^H \rrbracket) ds. \end{aligned} \tag{2.13}$$

An appropriate space must be considered for the analysis (see [25]), and we define a space $V(\mathcal{T}_H)$ as

$$V(\mathcal{T}_H) = V^p(\Omega, \mathcal{T}_H) + H_0^1(\Omega) \cap H^2(\Omega), \quad V(\mathcal{T}_H) \subset H^2(\mathcal{T}_H),$$

equipped with a mesh-dependent norm

$$\|v\| = \left(\|\nabla v\|_{L^2(\Omega)}^2 + \sum_{K \in \mathcal{T}_H} h_K^2 |v|_{2,K}^2 + |v|_*^2 \right)^{1/2}, \tag{2.14}$$

where

$$|v|_{m,K}^2 = \sum_{|s|=m} \|\partial^s v\|_{L^2(K)}^2, \quad \|\nabla v\|_{L^2(\Omega)}^2 = \sum_{K \in \mathcal{T}_H} |v|_{1,K}^2, \quad \text{and} \quad |v|_*^2 = \sum_{e \in E} \|\mu_e \llbracket v \rrbracket\|_{L^2(e)}^2.$$

Thanks to a discrete Poincaré inequality [25, Lemma 2.1]; i.e.,

$$\|v\|_{L^2(\Omega)} \leq C(\|\nabla v\|_{L^2(\Omega)}^2 + |v|_*^2)^{1/2},$$

it holds that (2.14) is a norm over the space $V(\mathcal{T}_H)$. The well-posedness of the discrete problem (2.12) follows from the Lax–Milgram lemma, provided that the bilinear form B_{DG} is bounded and stable. The latter holds if the parameter $\alpha > 0$ in the penalty weighting functions μ_e is chosen sufficiently large [61]. Optimal convergence rates are obtained; assume that $u \in H^{p+1}(\Omega)$, it holds

$$\|u - u^H\| \leq CH^p |u|_{H^{p+1}(\Omega)}, \quad \text{and} \quad \|u - u^H\|_{L^2(\Omega)} \leq CH^{p+1} |u|_{H^{p+1}(\Omega)}.$$

2.2 Homogenization theory

The aim of homogenization theory is to capture the global (macroscopic or coarse) behavior of a heterogeneous system. To this end, we consider the heterogeneous problem (2.1) where the multiple scale are represented by $\varepsilon > 0$. When the heterogeneities are made smaller and smaller, i.e., when ε tends to zero, the material is progressively replaced by a homogenized one whose solution is a good approximation of the global behavior of the solution u^ε of (2.1). Homogenization theory [80, 30] is the tool to establish the global behavior of u^ε , when ε tends to zero.

The family of solutions $\{u^\varepsilon\}$ is bounded by the data in $H_0^1(\Omega)$, thus there exist a subsequence $\{u^{\varepsilon'}\}$, and an element $u^0 \in H_0^1(\Omega)$ such that weak convergence is obtained; i.e.,

$$u^{\varepsilon'} \rightharpoonup u^0, \text{ weakly in } H_0^1(\Omega).$$

It can be easily established that the vector $a^\varepsilon \nabla u^\varepsilon$ is bounded by the data of the problem (2.1); therefore, there exist a subsequence $\{a^{\varepsilon'} \nabla u^{\varepsilon'}\}$ and a quantity $\xi^0 \in L^2(\Omega)$ such that

$$a^{\varepsilon'} \nabla u^{\varepsilon'} \rightharpoonup \xi^0 \text{ weakly in } L^2(\Omega)^d.$$

Further, it holds

$$\operatorname{div}(\xi^0) = f, \text{ in } \Omega. \quad (2.15)$$

The limit u^0 is our quantity of interest and natural questions arise such as: is u^0 a solution of an elliptic boundary value in Ω and if so, can it be uniquely determined? In view of problem (2.15), these questions are answered if we can find a link between ξ^0 and u^0 . Homogenization gives positive answers to the two questions and the key resides in the convergence of the tensor a^ε when ε goes to zero. Notice that the right hand side f can also be varying at the fine scale ε and we refer to [80, 44] for such treatments.

G and H -convergences were introduced by Spagnolo and de Giorgi [106, 45], Tartar [108], and further by Murat and Tartar [90], and deal with the convergence of the solutions of problem (2.1). The notion of G -convergence is concerned with the convergence of u^ε for boundary problems with symmetric tensor a^ε , whereas the H -convergence deals with the convergence of u^ε and $a^\varepsilon \nabla u^\varepsilon$ for any matrices $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$ satisfying (2.2). The H -convergence reads: for any family of tensor $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$, there exists a tensor a^0 such that a subsequence $\{u^{\varepsilon'}\}$ of $\{u^\varepsilon\}$, satisfies

$$u^{\varepsilon'} \rightharpoonup u^0 \text{ weakly in } H_0^1(\Omega), \quad a^{\varepsilon'} \nabla u^{\varepsilon'} \rightharpoonup a^0 \nabla u^0 \text{ weakly in } L^2(\Omega)^d,$$

and where the limit u^0 is the solution of the so-called homogenization problem

$$\begin{aligned} -\operatorname{div}(a^0(x) \nabla u^0) &= f, \text{ in } \Omega, \\ u^0 &= 0, \text{ on } \Gamma. \end{aligned} \quad (2.16)$$

Note that a^0 satisfies (2.2), with a different upper bound Λ , and that the homogenization problem (2.16) is well-posed, thanks to the Lax–Milgram lemma. Even though, there exists a unique u^0 solution of (2.16), the limit u^0 depends, a priori, on the subsequence $\{u^{\varepsilon'}\}$. The uniqueness of the limit would hold if the sequence $\{u^\varepsilon\}$ converges weakly to u^0 , i.e., if the convergence holds for any subsequence $\{u^{\varepsilon'}\}$ of $\{u^\varepsilon\}$. Further, in general, no explicit formulas are available for the homogenized tensor a^0 .

However, in some situations, in particular in the periodic case, one can show that the tensor a^0 is independent of the subsequence $\{u^{\varepsilon'}\}$, which in turn implies that u^0 is unique.

2.2.1 Periodic homogenization

Restricting the class of admissible tensors a^ε to periodic ones leads to explicit equations for the homogenized tensor a^0 . As a result, one obtains a tensor a^0 independently of the subsequence $\{\varepsilon'\}$ which implies that the homogenized solution u^0 is also independent of $\{\varepsilon'\}$, and that the sequence $\{u^\varepsilon\}$ converges to u^0 . The homogenized tensor is described using solutions of auxiliary periodic problems.

Let Y be the unit hypercube in \mathbb{R}^d , i.e., $Y = [0, 1]^d$. Consider $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$ a Y -periodic tensor in the fast variable, i.e., $a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y)$ is Y -periodic in y and assume that $a(x, y) \in \mathcal{C}(\Omega; L^\infty_{per}(Y))$. Explicit equations are available for the homogenized tensor a^0 ; i.e.,

$$a^0(x) = \int_Y a(x, y) (I_d + \nabla \chi(x, y)) dy,$$

where $\nabla \chi(x, y) = (\nabla \chi_1(x, y), \dots, \nabla \chi_d(x, y)) \in \mathbb{R}^{d \times d}$ and I_d is the $d \times d$ identity matrix. The functions $\chi_i(x, y)$ are defined as the unique solutions of the cell problems

$$\int_Y a(x, y) \nabla \chi_i(x, y) \cdot \nabla w dy = - \int_Y a(x, y) e_i \cdot \nabla w dy, \quad \forall w \in W^1_{per}(Y), \quad (2.17)$$

where $\{e_i\}$ is the canonical basis of \mathbb{R}^d .

The sequence u^ε converges towards the homogenized solution u^0 strongly in $L^2(\Omega)$, but weakly in $H^1(\Omega)$. Assuming sufficient regularity on the domain and the data, it holds

$$\|u^\varepsilon - u^0\|_{L^2(\Omega)} \leq C\varepsilon, \quad \text{and} \quad \|u^\varepsilon - (u^0 + \varepsilon u_1(x, x/\varepsilon))\|_{H^1(\Omega)} \leq C\varepsilon^{1/2}, \quad (2.18)$$

where u_1 is given by

$$u_1(x, x/\varepsilon) = \sum_{j=1}^d \chi^j(x, x/\varepsilon) \frac{\partial u^0(x)}{\partial x_j}.$$

The estimate (2.18) holds if $a(\cdot, y) \in W^{1,\infty}(Y)$, and $u^0 \in H^2(\Omega)$, see [80]. Further, thanks to the regularity of a^ε , we have $\chi^j \in W^{1,\infty}(Y)$. The regularity on a^ε can be relaxed to $a(\cdot, y) \in W^{1,p}(Y)$ for $p > 2$, and $\chi^j \in W^{1,p}(Y) \cap \mathcal{C}^{1,s}(\overline{Y})$ for $s = 1 - d/p$. For the proof of (2.18), we refer to [80, 89].

2.3 A brief review of numerical homogenization methods

Many problems are concerned with rapidly varying data with multiple scales. Although the standard methods introduced in Section 2.1 provide us with an efficient tool to solve single-scale problems, i.e., when no multiple scales are involved, such methods fail for multiple scale problems. To obtain a good approximation u^H , the FEM requires a discretization of Ω with a mesh size H satisfying $H \ll \varepsilon$ (see the FEM errors (2.9)). This is consistent with the fact that the multiple scale variations in the data must be fully resolved by the numerical method. This leads to a discretization of the domain Ω with a number of degrees of freedom of the order $\mathcal{O}(\varepsilon^{-d})$, and as ε goes to zero the cost of the method becomes prohibitive. This motivates the need for different numerical methods.

In this section, we give a brief review of some numerical homogenization methods. Such methods are split into two categories; the first one regroups the methods based on a global extraction of the fine scale information, whereas the second one regroups the methods based on a global to local downscaling, which is closely related to the homogenization theory. The latter supplement macroscopic data, computed through micro computations, to obtain a solution of an effective equation, which is solved using a macro solver. The main difference between the two classes of methods lies in the extraction of the fine scale information; in the first class all the information about the multiscale nature of the tensor are used, as opposed to the second class of methods, where the small scales are needed only in some small subdomains of Ω .

In the first category, one can find, amongst others, the multiscale finite element method [78, 79], the generalized multiscale finite element method [57], the wavelet-based numerical homogenization [50], and the localized orthogonal decomposition method [88].

The second class of methods regroups, amongst others, the heterogeneous multiscale method [54], high-dimensional finite element methods [77], and the zero-order regularization of local problems [68, 69, 70]. The HMM is the method used in this thesis and will be explained in details in Section 2.4.

Although different, all numerical homogenization methods deal with the approximation of the solution u^ε (also u^0 in the process), and are based on two different discretization levels on Ω ; one coarse scale level to describe the macroscopic behavior of the fine scale solution, and at least one fine scale level to recover the missing fine scale information. The numerical homogenized solution lives in the coarse finite element space, and is considered as a good L^2 approximation of the fine scale solution u^ε , if the coarse mesh size is not too large. However, as the numerical homogenized solution corresponds to an average of u^ε , it lacks the fine scale information, making it a bad approximation in H^1 .

The number of numerical homogenization methods is large, and the purpose of the thesis is not to review all existing methods. We therefore, chose to focus only on two methods in each category given above.

Outline of Section 2.3. The outline is as follows. In 2.3.1 we discuss two methods based on the global extraction of the fine scales; the multiscale finite element method and the localized orthogonal decomposition method. In 2.3.2 we give one method based on homogenization; the high-dimensional finite element method.

2.3.1 Numerical methods based on global extractions of the small scales

The methods considered here are based on solving fine scale problems locally in an overlapping partition of Ω . The fine scales are recovered globally on the domain Ω , as it is covered by the union of the overlapping subdomains. In that sense, all the available information about the fine scales are needed and used. This leads to numerical methods with a computational complexity depending on the number of degrees of freedom of the fine grid.

In here, we chose to give a detailed description of the MsFEM and the LOD. The LOD aims at reducing a high-dimensional finite element space, and we note that other methods use the same strategy. We cited the generalized MsFEM [57], but refer as well to recent developments [39] and [40].

Multiscale finite element method

The multiscale finite element method (MsFEM) was first introduced by Hou and Wu [78] and further developed in [79]. The idea is to adapt the basis functions to the local properties of the problems; this is done by enriching the nodal basis functions of the finite element space with the local fine scales. The numerical approximation of the solution u^ε is obtained from the Galerkin method applied to the set of enriched functions. For higher order MsFEM, we refer to [24, 75].

The classical formulation of the MsFEM produces resonance errors due to the unnatural forced boundary conditions used in the local problems. Indeed, in each macro element of the partition of Ω , the method requires the new multiscale basis functions to be equal to the coarse nodal basis functions on the boundary of the finite element. This implies strong oscillations close to the boundary of the coarse element, and lead to a pollution in the quality of the multiscale approximation — a term ε/H . In order to decrease these effects, oversampling can be applied; the local problems are not solved in the macro element K , but in a slightly larger domain $U(K)$. The obtained solution is restricted to the finite element K and only this part is used as the new multiscale basis function. Hence, the information coming from $U(K) \setminus K$ is ignored. This strategy improves the final approximation, which is now, non-conforming.

We briefly give the main components of the general MsFEM. Let $\{\mathcal{T}_H\}$ be a family of coarse partitions over Ω , into elements K of mesh size h_K , and set $H = \max_{K \in \mathcal{T}_H} h_K$. The mesh size $H > 0$ is, typically, larger than the fine scale, i.e. $H \gg \varepsilon$. For each partition $\mathcal{T}_H \in \{\mathcal{T}_H\}$, one defines a macro finite element space of degree $p \geq 1$, $V_0^p(\Omega, \mathcal{T}_H)$ following (2.6).

2.3. A brief review of numerical homogenization methods

The fine scale levels are defined as follows. Around each coarse element K , we define a shape regular domain $U(K)$, i.e., $K \subset U(K)$ and consider a family of fine partitions $\{\mathcal{T}_h\}$ with the requirement that the fine scales are fully resolved in the grids, i.e., $h \leq \varepsilon$. For each partition \mathcal{T}_h of $\{\mathcal{T}_h\}$ we define a micro finite element space of degree $p \geq 1$, $V^p(U(K), \mathcal{T}_h)$, defined as in (2.6). Further, we assume that both family of partitions are admissible (T1) and shape-regular (T2). Consider $\{\varphi_{K,i}^H\}_{i=1}^{n_K}$ the set of coarse basis functions in $V_0^p(\Omega, \mathcal{T}_H)$ with support on K , and where n_K denotes the number of nodes in the element K — n_K depends on the dimension d , the degree p of the finite element space, and on whether K is simplicial or quadrilateral. To each basis function, we seek $Q_K^h(\varphi_{K,i}^H) \in V^p(U(K), \mathcal{T}_h)$ such that

$$\int_{U(K)} a^\varepsilon(x) \nabla Q_K^h(\varphi_{K,i}^H) \cdot \nabla w^h dx = - \int_{U(K)} a^\varepsilon(x) \nabla \varphi_{K,i}^H \cdot \nabla w^h dx, \quad \forall w^h \in V^p(U(K), \mathcal{T}_h).$$

The multiscale basis functions can be constructed by $\varphi_{K,i}^H + Q_K^h(\varphi_{K,i}^H)$, for all $i = 1, \dots, n_K$ and for all $K \in \mathcal{T}_H$. The multiscale finite element space $V_{ms}^p(\Omega, \mathcal{T}_H)$ is obtained by the span of the multiscale basis functions. The coarse MsFEM approximation is the solution of the variational formulation: find $u^H \in V_{ms}^p(\Omega, \mathcal{T}_H)$ such that

$$\sum_{K \in \mathcal{T}_H} |K| \int_K a^\varepsilon(x) (\nabla u^H + \nabla Q_K^h(u^H)) \cdot \nabla w^H dx = \int_\Omega f w^H dx, \quad \forall w^H \in V_{ms}^p(\Omega, \mathcal{T}_H).$$

The multiscale approximation of u^ε is defined by

$$u_{ms}^H = \sum_{K \in \mathcal{T}_H} \mathbb{1}_K(u^H + Q_K^h(u^H)).$$

where $\mathbb{1}_K$ is the indicator function of K .

Assume that a^ε is locally periodic in the fast variable; e.g. $a^\varepsilon(x) = a(x, x/\varepsilon)$, where $a(x, y)$ is Y -periodic in y . Further assume that $u^\varepsilon \in H^2(\Omega)$. Then, an a priori error estimate can be derived [79, 24]; i.e.,

$$\left(\sum_{K \in \mathcal{T}_H} \|\nabla(u^\varepsilon - u_{ms}^H)\|_{L^2(K)}^2 \right)^{1/2} \leq C \left(\frac{\varepsilon}{d_H} + \varepsilon^{1/2} + H + \frac{h}{\varepsilon} \right),$$

where $d_H = \min_{K \in \mathcal{T}_H} \text{dist}(K, \partial U(K))$.

Localized orthogonal decomposition method

The localized orthogonal decomposition method (LOD) was initially proposed by Malqvist and Peterseim [88]. The idea behind the LOD is to reduce a high-dimensional finite element space by sorting out functions in a kernel of a coarse scale interpolation operator. Let $\{\mathcal{T}_H\}$ be a family of partitions over Ω , and, to each partition \mathcal{T}_H in $\{\mathcal{T}_H\}$, consider a fine refinement \mathcal{T}_h of \mathcal{T}_H , where each element is at least refined twice. For each \mathcal{T}_H and \mathcal{T}_h , we construct FE spaces of degree $p \geq 1$, $V_0^p(\Omega, \mathcal{T}_H)$ and $V_0^p(\Omega, \mathcal{T}_h)$, as defined in (2.6). The goal is to split

Chapter 2. Multiscale methods for elliptic problems

the fine space $V_0^p(\Omega, \mathcal{T}_h)$ into a coarse FE space $V_{ms}^p(\Omega, \mathcal{T}_H)$ and a fine space V_h containing negligible information, i.e.,

$$V_0^p(\Omega, \mathcal{T}_h) = V_{ms}^p(\Omega, \mathcal{T}_H) \oplus V_h.$$

Let $\{x_i\}_{i=1}^{n_{int}}$ be the set of interior nodes in the mesh \mathcal{T}_H , and consider φ_i^H the nodal basis function such that $\varphi_i^H(x_j) = \delta_{ij}$, for $i, j = 1, \dots, n_{int}$. One consider a Clément-type operator, $I_H : H_0^1(\Omega) \rightarrow V^p(\Omega, \mathcal{T}_H)$, defined as

$$I_H(v) = \sum_{i=1}^{n_{int}} v_{x_i} \varphi_i^H, \quad \text{with} \quad v_{x_i} = \frac{(v, \varphi_i^H)_{L^2(\Omega)}}{(1, \varphi_i^H)_{L^2(\Omega)}}.$$

The fine space V_h is given by the kernel of the operator I_H on $V^p(\Omega, \mathcal{T}_h)$; i.e.,

$$V_h = \{v^h \in V_h \mid I_H(v^h) = 0\},$$

and one notices that all the fine scale information that are not captured by the coarse scale FE space $V^p(\Omega, \mathcal{T}_H)$ are present in V_h . The multiscale basis used to construct the Galerkin approximation of the homogenized solution u^0 , are living in the space $V_{ms}^p(\Omega, \mathcal{T}_H)$, which is given by the B^ε -orthogonal complement of the space V_h onto the fine scale space $V^p(\Omega, \mathcal{T}_h)$. More precisely, $V_{ms}^p(\Omega, \mathcal{T}_H) = \ker(P_h)$, where $P_h : V^p(\Omega, \mathcal{T}_h) \rightarrow V_h$ satisfies

$$\int_{\Omega} a^\varepsilon(x) \nabla P_h(v^h) \cdot \nabla w^h dx = \int_{\Omega} a^\varepsilon(x) \nabla v^h \cdot \nabla w^h dx, \quad \forall w^h \in V_h.$$

Notice that it requires to solve a fine scale problem for each coarse basis function on the whole domain Ω . To solve that issue, one localizes the decompositions by using patches $U_k(K)$ that consists of a coarse element $K \in \mathcal{T}_H$ and k -layers of coarse elements around, i.e., for each $k \in \mathbb{N}$, and $K \in \mathcal{T}_H$, one constructs iteratively a sequence of patches $\{U_k(K)\}$ by

$$U_0(K) = K, \quad U_k(K) = \bigcup \{T \in \mathcal{T}_H \mid T \cup U_{k-1}(K) \neq \emptyset\}, \quad k = 1, 2, \dots$$

Now instead of considering V_h on Ω , the strategy proposed by [88] is to consider

$$V_h(U_k(K)) = \{v^h \in V_h \mid v^h = 0 \text{ in } \Omega \setminus U_k(K)\}.$$

Then, one defines a local corrector $Q_h^K : V^p(\Omega, \mathcal{T}_H) \rightarrow V_h(U_k(K))$ by: for each φ^H in the space $V^p(\Omega, \mathcal{T}_H)$ the corrector $Q_h^K(\varphi^H) \in V_h(U_k(K))$ satisfies

$$\int_{U_k(K)} a^\varepsilon(x) \nabla Q_h^K(\varphi^H) \cdot \nabla w^h dx = - \int_K a^\varepsilon(x) \nabla \varphi^H \cdot \nabla w^H dx, \quad \forall w^h \in V_h(U_k(K)).$$

The multiscale FE space $V_{ms,k}^p(\Omega, \mathcal{T}_H)$ is given by

$$V_{ms,k}^p(\Omega, \mathcal{T}_H) = \{\varphi^H + Q_h(\varphi^H) \mid \varphi^H \in V^p(\Omega, \mathcal{T}_H)\},$$

2.3. A brief review of numerical homogenization methods

where $Q_h(\varphi^H)$ is a global corrector defined by

$$Q_h(\varphi^H) = \sum_{K \in \mathcal{T}_H} Q_h^K(\varphi^H).$$

The variational formulation of the LOD method reads: find $u_{ms}^H \in V_{ms,k}^p(\Omega, \mathcal{T}_H)$ such that

$$B^\varepsilon(u_{ms}^H, w^H) = \int_{\Omega} f w^H dx, \quad \forall w^H \in V_{ms,k}^p(\Omega, \mathcal{T}_H),$$

where B^ε is given in (2.4). An a priori error estimate for the LOD is available [88, 73]

$$\|u^\varepsilon - u_{ms}^H\|_{H^1(\Omega)} \leq C(H + k^{\frac{d}{2}} \theta^k) + \inf_{v^h \in V^p(\Omega, \mathcal{T}_h)} \|u^\varepsilon - v^h\|_{H^1(\Omega)},$$

where $0 < \theta < 1$ represents the exponential decay of $Q_h^K(\varphi^H)$ to zero in $\Omega \setminus K$.

2.3.2 Numerical methods based on homogenization

Here, we discuss methods that are directly related to homogenization theory. Their nature differ from the methods given in 2.3.1, in the sense that the methods presented here are based on local computations in subdomains of diameter $\delta \ll H$. One can then guess the homogenized tensor a^0 from these computations and obtain, on a coarse grid, the homogenized solution u^0 as a solution of an effective equation. The complexity of such methods reduces drastically, since the fine scales are resolved only in small subdomains. However, they rely on homogenization theory, which requires assumptions on the heterogeneous tensor a^ε , such as scale separation, periodicity, or statistical distributions. Once the homogenized solution is computed, post-processing procedures can be applied and fine scale information can be recovered. From that strategy, we turn the L^2 approximation of the fine scale solution u^ε into an H^1 approximation. Global to local post-processing (or global to local downscaling) methods such as the L^2 global to local projection method [27] or the goal-oriented adaptive method [95] are the subject of Chapter 5.

The thesis focuses on multiscale problems with two scales and uses the heterogeneous multiscale method. Although it relies on homogenization theory as well, the heterogeneous multiscale method (HMM) is detailed in Section 2.4.

It is well-known, that the fine scale information in the solution of problem (2.1) has been lost in the homogenization process. Correctors can be added to the homogenized solution to gain the lost oscillations, but they are as costly as a fine scale solver like the FEM. For two-scales problems, the classical procedure to recover the fine scale is to do a multiscale asymptotic expansion and then prove the convergence using Tartar energy method, see [30, 44]. However, the generalization to multiple scales is not trivial, and we chose here to review one method that addresses this issue. We note that a fully discrete analysis of the FE-HMM for problems with N well separated scales is given by Abdulle and Bai in [6].

High-dimensional FEM for multiple scales elliptic problems

Nguetseng [91] and later Allaire [21] developed a new method for the two-scale convergence, where the asymptotic extension and the energy method are put together in one method. A generalization of this method to multiple scale problems can be found in [23] for periodic homogenization. Allaire and Briane [23] obtain a multidimensional limiting equation that contains the full fine scale information, i.e., recovery of the physical oscillations on all length scales in the exact solution.

Hoang and Schwab in [77] introduced a multiscale FEM for elliptic problems with scale separation based on [23], which allows us to fully resolve the fine scales with a degree of freedom that is of log-linear complexity in $1/H$ and obtain robust convergence rates in the H^1 norm. Let $\Omega \in \mathbb{R}^d$ be a bounded domain and consider Problem (2.1) with $n+1$ -scales described by ε , i.e., $a^\varepsilon(x) = a(x, x/\varepsilon_1, \dots, x/\varepsilon_n)$ is a symmetric and bounded matrix function of $n+1$ variables taking values in $\mathbb{R}^{d \times d}$, satisfies the assumption (2.2) and is Y -periodic with respect to $y_i = x/\varepsilon_i$, for all $i = 1, \dots, n$. For $n = 1$, we recover the classical two-scale homogenization problem. The limit of u^ε when ε converges to zero is posed on a tensorized domain $\mathbb{R}^{(n+1)d}$ and in order to obtain robust convergence, the FE spaces are chosen to be sparse tensor products FE spaces. Scale separation is assumed in the following sense: for $\varepsilon_1, \dots, \varepsilon_n$, positive functions of ε all converging to zero when ε tends to zero, we assume

$$\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon_{i+1}}{\varepsilon_i} = 0, \quad \forall i = 1, \dots, n.$$

Let Y_1, \dots, Y_n denote n unit cells for the n fast scales. The sequence $\{u^\varepsilon\}$ is said to $(n+1)$ -scale converge to $u^0(x, y_1, \dots, y_n) \in L^2(\Omega \times Y_1 \times \dots \times Y_n)$ if

$$\lim_{\varepsilon \rightarrow 0} \int_{\Omega} u^\varepsilon \varphi(x, \frac{x}{\varepsilon_1}, \dots, \frac{x}{\varepsilon_n}) dx = \int_{\Omega} \int_{Y_1} \dots \int_{Y_n} u^0(x, y_1, \dots, y_n) \varphi(x, y_1, \dots, y_n) dx dy_1 \dots dy_n,$$

for $\varphi \in L^2(\Omega, \mathcal{C}_{\text{per}}^1(Y_1 \times \dots \times Y_n))$. A compactness result allows us to extract from any bounded sequence $\{u^\varepsilon\}$, a subsequence that $(n+1)$ -scale converge to u^0 . We denote the converging subsequence by $\{u^\varepsilon\}$.

The variational formulation of problem (2.1) is: find $(u, u_1, \dots, u_n) \in \mathbf{V}$ such that

$$\begin{aligned} & B((u, \{u_i\}), (\varphi, \{\varphi_i\})) \\ &= \int_{\Omega} \int_{Y_1} \dots \int_{Y_n} a \left(\nabla_x u + \sum_{i=1}^n \nabla_{y_i} u_i \right) \cdot \left(\nabla_x \varphi + \sum_{i=1}^n \nabla_{y_i} \varphi_i \right) dx dy_1 \dots dy_n \\ &= \int_{\Omega} f \varphi dx, \quad \forall (\varphi, \{\varphi_i\}) \in \mathbf{V}, \end{aligned} \tag{2.19}$$

where \mathbf{V} is given by

$$\mathbf{V} = \{(\varphi, \{\varphi_i\}) \mid \varphi \in H_0^1(\Omega), \varphi_i \in L^2(\Omega \times Y_1 \dots \times Y_{i-1}, H_{\text{per}}^1(Y_i)/\mathbb{R}), i = 1, \dots, n\}.$$

2.3. A brief review of numerical homogenization methods

They look for numerical solution in a sparse tensor product FE space [77]. A sparse tensor product FE space in $\Omega \times Y_1 \times \dots \times Y_n$ is constructed with hierarchical sequences of FE spaces in the macro and micro domains. For the macro variable, we consider a sequence $\{V^l\}_{l=0}^\infty$ of subspaces of $H^1(\Omega)$ and for the micro variables, $\{V_{\text{per}}^l\}_{l=0}^\infty$ subspaces of $H_{\text{per}}^1(Y)$ such that $V^l \subset V^{l+1}$ and $V_{\text{per}}^l \subset V_{\text{per}}^{l+1}$, for all $l = 0, 1, \dots$. In order to define the sparse tensor product, we state some orthogonal projections defined in [77]

$$\begin{aligned} P_{\text{per}}^{l1} &: H_{\text{per}}^1(Y) \rightarrow V_{\text{per}}^l, \\ P_{\text{per}}^{l0} &: L^2(Y) \rightarrow V_{\text{per}}^l, \\ P^{l0} &: L^2(\Omega) \rightarrow V^l. \end{aligned}$$

The sparse tensor product FE spaces are

$$\tilde{V}_i^L = \bigoplus_{0 \leq j_0 + \dots + j_i \leq L} W^{j_0} \otimes W_{\text{per}}^{j_1 0} \otimes \dots \otimes W_{\text{per}}^{j_{i-1} 0} \otimes W_{\text{per}}^{j_i 1}, \quad \forall i = 1, \dots, n,$$

where the W -spaces are defined with the orthogonal projection given previously as

$$\begin{aligned} W^l &= (P^{l0} - P^{(l-1)0})V^l, \\ W_{\text{per}}^{l0} &= (P_{\text{per}}^{l0} - P_{\text{per}}^{(l-1)0})V_{\text{per}}^l, \\ W_{\text{per}}^{l1} &= (P_{\text{per}}^{l1} - P_{\text{per}}^{(l-1)1})V_{\text{per}}^l. \end{aligned}$$

The sparse FE space for the approximation of the variational problem (2.19) is

$$\tilde{\mathbf{V}}^L = \{(\tilde{u}, \{\tilde{u}_i^L\}) \mid \tilde{u}^L \in V_0^L, \tilde{u}_i^L \in \tilde{V}_i^L, i = 1, \dots, n\},$$

where the space V_0^L is the full tensor product spaces defined similarly to \tilde{V}_i^L , where one sums on $0 \leq j_k \leq L, k = 0, 1, \dots, i$ instead of $0 \leq j_0 + \dots + j_i \leq L$. The numerical approximation of $(u, \{u_i\})$, solution of (2.19) is $(\tilde{u}^L, \{\tilde{u}_i^L\}) \in \tilde{\mathbf{V}}^L$ solution of

$$B(\tilde{u}^L, \{\tilde{u}_i^L\}; \tilde{\varphi}^L, \{\tilde{\varphi}_i^L\}) = \int_{\Omega} f \tilde{\varphi}_i^L dx, \quad \forall (\tilde{\varphi}^L, \{\tilde{\varphi}_i^L\}) \in \tilde{\mathbf{V}}^L.$$

Robust convergence to the physical solution u^ε is stated in [77] for a tensor product of h -FE spaces. Ideally, one would like a convergence similar to (2.18) using the FE functions u_1^L and \tilde{u}_1^L . However, as the norms of u_1^L and \tilde{u}_1^L might be unbounded, a post-processing of the function u^1 is done, using a folding operator U^ε defined in [43].

For the multiple scale case, error estimates is harder to obtain and might not even exist. However, one is able to construct a numerical corrector for the special case where $\varepsilon_{i+1}/\varepsilon_i$ is an integer and established convergence in H^1 -norm[77].

Definition 2.3.1. Let $\psi \in L^1(\Omega \times Y_1 \times \dots \times Y_n)$, the folding operator U_n^ε is defined by

$$U_n^\varepsilon(\psi)(x) = \int_{Y_1} \dots \int_{Y_n} \psi \left(\varepsilon_1 \left\lfloor \frac{x}{\varepsilon} \right\rfloor + \varepsilon_1 z_1, \frac{\varepsilon_2}{\varepsilon_1} \left\lfloor \frac{\varepsilon_1}{\varepsilon_2} \left\{ \frac{x}{\varepsilon} \right\} \right\rfloor + \frac{\varepsilon_2}{\varepsilon_1} z_2, \dots, \right. \\ \left. \frac{\varepsilon_n}{\varepsilon_{n-1}} \left\lfloor \frac{\varepsilon_{n-1}}{\varepsilon_n} \left\{ \frac{x}{\varepsilon} \right\} \right\rfloor + \frac{\varepsilon_n}{\varepsilon_{n-1}} z_n, \left\{ \frac{x}{\varepsilon} \right\} \right) dz_n \dots dz_1,$$

where $\{x/\varepsilon\} := x/\varepsilon - \lfloor x/\varepsilon \rfloor$

Using the folding operator U_n^ε , the functions u_1, \dots, u_n approximate the oscillations of $\{u^\varepsilon\}$ as ε goes to 0.

Lemma 2.3.2. *For the problem (2.1) with scale separation, it holds*

$$\lim_{\varepsilon \rightarrow 0} \|\nabla u^\varepsilon - U_n^\varepsilon(\nabla_x u + \nabla_{y_1} u_1 + \dots + \nabla_{y_n} u_n)\|_{L^2(\Omega)} = 0$$

Schwab and Hoang [77] obtain the following error estimate for the multiple scale analysis, where the dependencies on ε and the FE spaces are dropped.

Theorem 2.3.3. *With the sparse tensor product FE approximation, it holds*

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ L \rightarrow \infty}} \|\nabla_x u^\varepsilon(x) - U_n^\varepsilon(\nabla_x \tilde{u}^L + \nabla_{y_1} \tilde{u}_1^L + \dots + \nabla_{y_n} \tilde{u}_n^L)\|_{L^2(\Omega)} = 0.$$

2.4 Finite Element Heterogeneous Multiscale Method

The finite element heterogeneous multiscale method (FE-HMM or often purely HMM) was initially introduced by E and Engquist [54]. The original idea of the method is to guess the unknown homogenized equation (2.16) by solving local sample problems around macroscopic quadrature points. These local problems are often formulated in analogy to the cell problems in classical homogenization theory for locally periodic structures.

The attractivity of such methods is the possibility to obtain numerical approximations that correctly describe the macroscopic behavior of the multiscale problem at a cost that however is independent of the smallest scale. This can be achieved when the small scales can be localized, i.e., when the problem features scale separation. In Figure 2.1, we plot a tensor with scale separation for different values of ε with a sampling domain of size ε ; the sampling domains are represented with a black square. In Figure 2.2, we plot a tensor which does not have explicit separation of scales, the tensor is oscillatory at various small scales that are not well-separated. Therefore, we cannot explicitly identify a value ε or δ , for the sampling domains. In Figure 2.2, we zoomed in a portion of the tensor of size $1/10$.

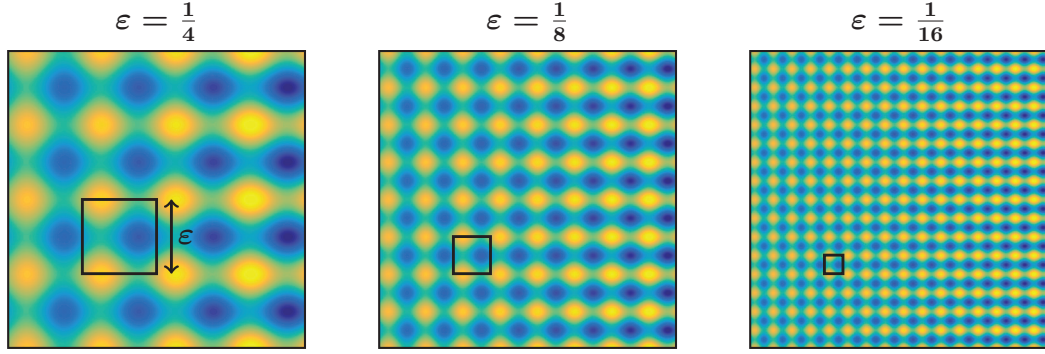


Figure 2.1 – Locally periodic tensor with explicit scale separation and with sampling domain (black squares) of size ε , for $\varepsilon = 1/4, 1/8$, and $1/16$.

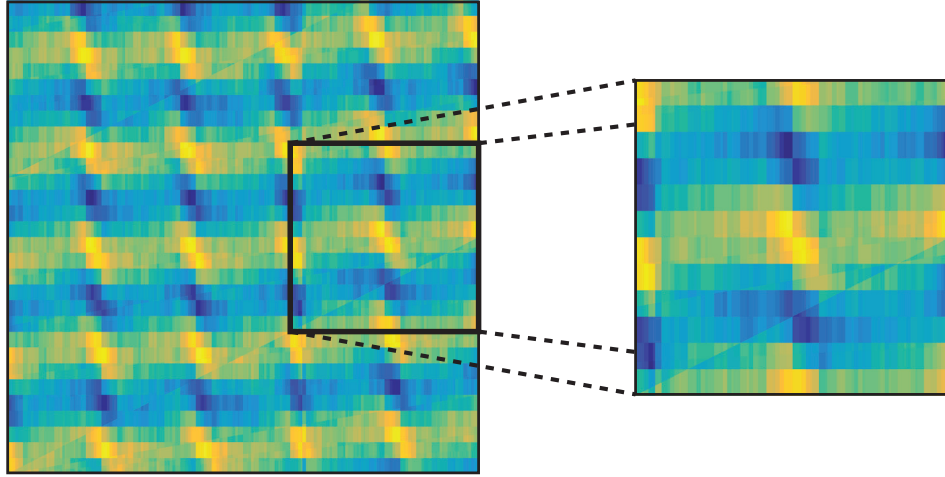


Figure 2.2 – Highly oscillatory tensor with non separated micro scales.

The framework that makes this family of multiscale methods efficient is that of a simultaneous coupling of a macro and a micro method. Rigorous convergence analysis has been established for locally periodic coefficients or random stationary coefficients [1, 56, 3]. We mention as well, the FE-HMM for the Stokes multiscale problems [7], the advection-diffusion multiscale problems [11] and non-linear multiscale parabolic problems [12].

From the homogenization theory (reviewed in Section 2.2) the exact heterogeneous solution u^ε of problem (2.1) can be decomposed into a coarse (homogenized), oscillations-free, part u^0 , and a fine part $u_1 := (u^\varepsilon - u^0)/\varepsilon$. The fine part u_1 is rapidly oscillating at a frequency of order ε , but remains bounded in L^∞ ; it holds $\|u_1\|_{L^\infty(\Omega)} = \mathcal{O}(1)$ and $\|\nabla u_1\|_{L^2(\Omega)} = \mathcal{O}(\varepsilon^{-1})$. In particular, provided sufficient regularity (see (2.18)), this implies

$$\|u^\varepsilon - u^0\|_{L^2(\Omega)} = \|\varepsilon u_1\|_{L^2(\Omega)} \leq C\varepsilon.$$

Chapter 2. Multiscale methods for elliptic problems

The heterogeneous multiscale method can be now divided in two main components: a macroscopic scheme for the approximation of the macro variables on a coarse grid (i.e. approximating the unknown homogenized coefficient a^0 and consequently also the homogenized solution u^0) and a recovery of the fine scale data (i.e. computing local approximations of the fine scale part u_1).

Let $\{\mathcal{T}_H\}$ be a family of coarse, admissible (T1), and shape-regular (T2) partitions over the domain Ω . We allow the mesh size to be larger than the fine scales, i.e., $H \gg \varepsilon$. For each partition \mathcal{T}_H in $\{\mathcal{T}_H\}$, we can define a macro FE space on Ω of degree $p \geq 1$ by $V_0^p(\Omega, \mathcal{T}_H)$ following (2.6).

Let \hat{K} be a reference triangle and consider a quadrature formula on \hat{K} defined as in 2.1.1. In each macro element $K \in \mathcal{T}_H$, we consider a quadrature formula $\{x_{j,K}, \omega_{j,K}\}$, $j = 1, \dots, J$ composed of nodes $x_{j,K}$ and weights $\omega_{j,K}$, and derived from the quadrature formula on \hat{K} . The quadrature formula satisfies (Q1) and (Q2). In each macro element $K \in \mathcal{T}_H$, we construct sampling domains $K_{\delta_j} = x_{j,K} + \delta(-1/2, 1/2)^d$, i.e. a cube of diameter $\delta \ll H$, that is centered around the quadrature point $x_{j,K}$. The sampling domains K_{δ_j} are used to sample the effective macroscopic coefficient a^0 at the quadrature points $x_{j,K}$ and to use $a^0(x_{j,K})$ as representative values for the whole coarse element K . The notation K_{δ_j} should be understood as $K_\delta(x_{j,K})$, but we drop the dependency on $x_{j,K}$ to simplify the heavy notations.

Let $\{\mathcal{T}_h(K_{\delta_j})\}$ be a family of admissible (T1) and shape-regular (T2) partitions, of size $h \leq \varepsilon$, over the sampling domain K_{δ_j} . We make an abuse of notations and use \mathcal{T}_h to denote $\mathcal{T}_h(K_{\delta_j})$. For each partition \mathcal{T}_h in $\{\mathcal{T}_h\}$, we define a micro FE space on K_{δ_j} of degree $q \geq 1$ by

$$V^q(K_{\delta_j}, \mathcal{T}_h) = \{v^h \in W(K_{\delta_j}) \mid v^h|_K \in \mathcal{R}^q(K), \forall K \in \mathcal{T}_h\}, \quad (2.20)$$

where the Sobolev space $W(K_{\delta_j}) \subset H^1(K_{\delta_j})$ incorporates the boundary condition imposed on the sampling domain. The space $W(K_{\delta_j})$ sets the coupling conditions between the macro and micro problems via boundary conditions; we use either periodic boundary conditions or homogeneous Dirichlet boundary conditions, i.e.,

$$W(K_{\delta_j}) = H_{per}^1(K_{\delta_j}), \quad \text{for periodic coupling}, \quad (2.21)$$

$$W(K_{\delta_j}) = H_0^1(K_{\delta_j}), \quad \text{for Dirichlet coupling}. \quad (2.22)$$

To obtain uniqueness of the micro solutions with the periodic coupling, we consider as solutions the functions with zero mean over K_{δ_j} .

In each sampling domain K_{δ_j} , we find the micro contribution to the macro stiffness matrix, by solving elliptic problems with periodic or Dirichlet boundary conditions. The fine scales are then used only inside K_{δ_j} .

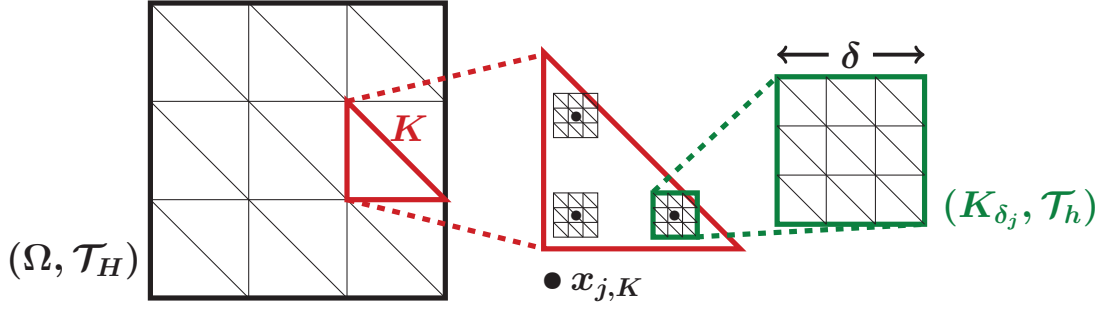


Figure 2.3 – Illustration of the FE-HMM procedure with a macro and micro partitions. The quadrature points (black bullets) are obtained with a QF of order 2.

The FE-HMM procedure is sketched in Figure 2.3. It is defined as follows: find u_j^h on the sampling domain K_{δ_j} such that $u_j^h - u_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$ and

$$\int_{K_{\delta_j}} a^\varepsilon(x) \nabla u_j^h \cdot \nabla z^h dx = 0, \quad \forall z^h \in V^q(K_{\delta_j}, \mathcal{T}_h), \quad (2.23)$$

where $u_{\text{lin},j}^H$ is a linearization of the macro function u^H at the quadrature node $x_{j,K}$; i.e.,

$$u_{\text{lin},j}^H(x) = u^H(x_{j,K}) + (x - x_{j,K}) \cdot \nabla u^H(x_{j,K}), \quad x \in K. \quad (2.24)$$

For piecewise linear functions, it holds that $u_{\text{lin},j}^H = u^H$. Here, we use the notation $u_{\text{lin},j}^H$ instead of $u_{\text{lin},x_{j,K}}^H$.

Problems (2.23) are well-posed as the tensor a^ε is bounded and uniformly elliptic (2.2). The variational formulation of the finite element heterogeneous multiscale method (FE-HMM) states [54, 3]: find $u_H \in V_0^p(\Omega, \mathcal{T}_H)$ such that

$$B_H(u^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (2.25)$$

where the bilinear form $B_H : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ is defined by

$$B_H(u^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla u_j^h \cdot \nabla w_j^h dx \quad (2.26)$$

and the right-hand side $F : V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ by

$$F(w^H) = \int_{\Omega} f w^H dx.$$

The function u_j^h (resp. w_j^h) is the solution of the micro problems (2.23) and satisfies $u_j^h - u_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$ (resp. $w_j^h - w_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$). The numerical homogenized tensor in a macro element K is denoted by a_K^0 (or sometimes $a_K^{0,h}$) and can be obtained during the assembly

Chapter 2. Multiscale methods for elliptic problems

process, using the solutions of the micro problems. Let $\{e_i\}$ be the canonical basis of \mathbb{R}^d . For each e_i and in each sampling domain K_{δ_j} , we seek $\psi_{K_{\delta_j}}^{i,h} \in V^q(K_{\delta_j}, \mathcal{T}_h)$ the solution of

$$\int_{K_{\delta_j}} a^\varepsilon(x) \nabla \psi_{K_{\delta_j}}^{i,h} \cdot \nabla z^h dx = - \int_{K_{\delta_j}} a^\varepsilon(x) e_i \cdot \nabla z^h dx, \quad \forall z^h \in V^q(K_{\delta_j}, \mathcal{T}_h), \quad (2.27)$$

where the boundary conditions are given by the Sobolev space $W(K_{\delta_j})$. Without further assumptions on the heterogeneous tensor a^ε , we can define

$$a_K^0(x_{j,K}) = \frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \left(I + J_{\psi_{K_{\delta_j}}^h}^T \right) dx, \quad (2.28)$$

where $J_{\psi_{K_{\delta_j}}^h}^T$ is a $d \times d$ matrix given by $(J_{\psi_{K_{\delta_j}}^h}^T)_{ik} = \partial \psi_{K_{\delta_j}}^{i,h}(x) / \partial x_k$.

Remark 2.4.1. With further assumptions on the tensor a^ε , such as periodicity, a_K^0 from (2.28) is a good approximation of the homogenized tensor a^0 . The error between a^0 and a_K^0 at a quadrature point $x_{j,K}$ is decomposed into a micro and a modeling error. Details are given in [3, 4, 5], and are recalled in Section 2.5.

A standard FEM based on the quadrature formulas can be applied on the homogenized problem (2.16), this leads to a variational formulation: find $u^{0,H} \in V_0^p(\Omega, \mathcal{T}_H)$ such that

$$B_{0,H}(u^{0,H}, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (2.29)$$

where the bilinear form $B_{0,H} : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ is given by

$$B_{0,H}(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} a^0(x_{j,K}) \nabla v^H(x_{j,K}) \cdot \nabla w^H(x_{j,K}), \quad (2.30)$$

provided that a^0 is well-defined at the quadrature points $x_{j,K}$.

Reformulation of the FE-HMM

We can write the FE-HMM bilinear form B_H (2.26) in terms of the macro functions $u^H, v^H \in V_0^p(\Omega, \mathcal{T}_H)$ and the numerical homogenized tensor a_K^0 . As we have $u_j^h - u_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$, it can be used as a test function in a scaled version of (2.23); i.e.,

$$\frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla u_j^h \cdot \nabla (u_j^h - u_{\text{lin},j}^H) dx = 0.$$

Then, we have the following Lemma (see [4, Lemma 12] and [5, Lemma 5.4] for a proof).

Lemma 2.4.2. *Let $v^H, w^H \in V_0^p(\Omega, \mathcal{T}_H)$ and v_j^h (resp. w_j^h) be such that $v_j^h - v_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$ (resp. $w_j^h - w_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$) and obtained from the micro problems (2.23), and where $v_{\text{lin},j}^H$*

2.4. Finite Element Heterogeneous Multiscale Method

(resp. $w_{lin,j}^H$) are defined by (2.24). Further, let a_K^0 be defined by (2.28). Then, we have

$$\frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla v_j^h \cdot \nabla w_j^h dx = \frac{1}{|K|} \int_K a_K^0(x_{j,K}) \nabla v_{lin,j}^H \cdot \nabla w_{lin,j}^H dx.$$

The bilinear form B_H (2.26) of the FE-HMM can be reformulated as

$$B_H(u^H, v^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} a_K^0(x_{j,K}) \nabla u^H(x_{j,K}) \cdot \nabla v^H(x_{j,K}), \quad (2.31)$$

where we used $\nabla u_{lin,j}^H(x) \equiv \nabla u^H(x_{j,K})$ (resp. $\nabla v_{lin,j}^H(x) \equiv \nabla v^H(x_{j,K})$) for $x \in K$. Notice that the reformulation (2.31) looks like a FEM bilinear form with quadrature points and weights for a coarse problem.

By considering the micro problems (2.27) in the exact Sobolev spaces $W(K_{\delta_j})$, one obtains a semi-discrete FE-HMM problem, which will be useful for the a priori error analysis in Section 2.5. Let $\psi_{K_{\delta_j}}^i \in W(K_{\delta_j})$ be the solution of

$$\int_{K_{\delta_j}} a^\varepsilon(x) \nabla \psi_{K_{\delta_j}}^i \cdot \nabla z dx = - \int_{K_{\delta_j}} a^\varepsilon(x) e_i \cdot \nabla z dx, \quad \forall z \in W(K_{\delta_j}), \quad (2.32)$$

where $W(K_{\delta_j})$ is defined by (2.21) or (2.22). Let \bar{a}_K^0 be given by

$$\bar{a}_K^0(x_{j,K}) = \frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \left(I + J_{\psi_{K_{\delta_j}}}^T \right) dx, \quad (2.33)$$

where $J_{\psi_{K_{\delta_j}}}^T$ is a $d \times d$ matrix given by $\left(J_{\psi_{K_{\delta_j}}}^T \right)_{ik} = \partial \psi_{K_{\delta_j}}^i(x) / \partial x_k$. Similarly to Lemma 2.4.2, it holds.

Lemma 2.4.3. *Let $v^H, w^H \in V_0^p(\Omega, \mathcal{T}_H)$ and v_j^h (resp. w_j^h) be such that $v_j^h - v_{lin,j}^H \in W(K_{\delta_j})$ (reps. $w_j^h - w_{lin,j}^H \in W(K_{\delta_j})$) and obtained from the micro problems (2.23) in $W(K_{\delta_j})$, and where $v_{lin,j}^H$ (resp. $w_{lin,j}^H$) are defined by (2.24). Further, let \bar{a}_K^0 be defined by (2.33). Then, we have*

$$\frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla v_j \cdot \nabla w_j dx = \frac{1}{|K|} \int_K \bar{a}_K^0(x_{j,K}) \nabla v_{lin,j}^H \cdot \nabla w_{lin,j}^H dx.$$

Then, the semi-discrete FE-HMM reads: find $\bar{u}^H \in V_0^p(\Omega, \mathcal{T}_H)$ such that

$$\bar{B}_H(\bar{u}^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (2.34)$$

where the bilinear form $\bar{B}_H : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ is given by

$$\bar{B}_H(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} \bar{a}_K^0(x_{j,K}) \nabla v^H(x_{j,K}) \cdot \nabla w^H(x_{j,K}).$$

2.4.1 Discontinuous Galerkin FE-HMM

In this section, the heterogeneous multiscale method is combined with the interior penalty discontinuous Galerkin method. The obtained method is denoted here by DG-FE-HMM [5]. As in 2.1.2 about discontinuous Galerkin FEM, we restrict the method to simplicial finite elements.

Let $\{\mathcal{T}_H\}$ be a family of shape-regular (T2) triangulations over the computational domain Ω , with mesh size $H = \max_{K \in \mathcal{T}_H} h_K$ greater than the fine scales ε . Consider a quadrature formula $\{\hat{x}_{j,K}, \hat{\omega}_{j,K}\}$ on a reference element \hat{K} , and assume that the quadrature formula verifies the assumption (Q2). The quadrature formula is exact for polynomials of degree $2p - 2$. In order to reduce the cost of the method, we take the minimal J such that the quadrature formula is still exact.

For each triangulation \mathcal{T}_H in $\{\mathcal{T}_H\}$, we define a FE space of degree p as

$$V^p(\Omega, \mathcal{T}_H) = \{v^H \in L^2(\Omega) \mid v^H|_K \in \mathcal{P}^p(K), \forall K \in \mathcal{T}_H\},$$

where $\mathcal{P}^p(K)$ is the space of polynomials on K of total degree at most p . In order to reduce the cost of the method, we take the minimal J such that the quadrature formula is still exact, i.e., $J = \frac{1}{2}p(p+1)$ if $d = 2$, or $J = \frac{1}{6}p(p+1)(p+2)$ if $d = 3$. Following the FE-HMM strategy, the bilinear form B_{DG} (2.13) is modified to allow large mesh size H : it is defined as $B_{H,DG} : V^p(\Omega, \mathcal{T}_H) \times V^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$

$$\begin{aligned} B_{H,DG}(v^H, w^H) = & \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla v_j^h \cdot \nabla w_j^h dx + \sum_{e \in E} \int_e \mu_e \llbracket v^H \rrbracket \llbracket w^H \rrbracket ds \\ & - \sum_{e \in E} \int_e \left(\{\Pi_{a^\varepsilon \nabla v_j^h}\} \llbracket w^H \rrbracket + \{\Pi_{a^\varepsilon \nabla w_j^h}\} \llbracket v^H \rrbracket \right) ds, \end{aligned}$$

where v_j^h (resp. w_j^h) is the solution of the micro problems (2.23) in the micro FE space $V^q(K_{\delta_j}, \mathcal{T}_h)$ defined in (2.20). Further for each macro element $K \in \mathcal{T}_H$, we define a quantity $\Pi_{a^\varepsilon \nabla v_j^h} \in \mathcal{P}^{p-1}(K)^d$ as

$$\Pi_{a^\varepsilon \nabla v_j^h}(x_{j,K}) = \frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla v_j^h dx, \quad j = 1, \dots, J.$$

Then we can define the average $\{\Pi_{a^\varepsilon \nabla v_j^h}(x_{j,K})\}$ as in (2.11). The DG-FE-HMM reads: find $u^H \in V^p(\Omega, \mathcal{T}_H)$ such that

$$B_{H,DG}(u^H, w^H) = F(w^H), \quad \forall w^H \in V^p(\Omega, \mathcal{T}_H), \quad (2.35)$$

where the right-hand side F is given by (2.8).

2.5 A priori error analysis for the (DG-)FE-HMM

In this section, we are interested in the a priori error estimates for the FE-HMM and DG-FE-HMM derived in [3, 4, 5]. We focus on the error between u^0 , the exact homogenized solution obtained by solving the homogenized problem (2.16), and u^H , the numerical homogenized approximation obtained from problem (2.25), if the FE-HMM is used, or by (2.35), if the DG-FE-HMM is used instead. The difference between u^0 and u^H is decomposed into

$$\|u^0 - u^H\| \leq e_{\text{MAC}} + e_{\text{MOD}} + e_{\text{MIC}},$$

respectively, the macro, modeling, and micro errors. The norm $\|\cdot\|$ stands for the norms $\|\cdot\|_{H^1}$, $\|\cdot\|_{L^2}$, or for the DG norm $\|\cdot\|_{\text{DG}}$ defined in equation (2.14). Using the numerical approximations $u^{0,H}$ and \bar{u}^H , given by (2.29) and (2.34), respectively, one obtains that the macro, modeling, and micro errors are

$$\begin{aligned} e_{\text{MAC}} &= \|u^0 - u^{0,H}\|, \\ e_{\text{MOD}} &= \|u^{0,H} - \bar{u}^H\|, \\ e_{\text{MIC}} &= \|\bar{u}^H - u^H\|. \end{aligned}$$

The modeling and micro errors can also be given in terms of the homogenized tensor a^0 and the two tensors \bar{a}_K^0 , and a_K^0 , defined in (2.28) and (2.33), respectively,

$$\begin{aligned} e_{\text{MOD}} &= \sup_{K \in \mathcal{T}_H, x_{j,K} \in K} \|a^0(x_{j,K}) - \bar{a}_K^0(x_{j,K})\|_F, \\ e_{\text{MIC}} &= \sup_{K \in \mathcal{T}_H, x_{j,K} \in K} \|\bar{a}_K^0(x_{j,K}) - a_K^0(x_{j,K})\|_F. \end{aligned}$$

We start by giving the estimates for the FE-HMM and then for the DG-FE-HMM.

A priori error estimates for the FE-HMM

One can obtain an estimate for the macro and micro errors with no assumptions on the tensor a^ε besides the ellipticity (2.2). However, specific knowledge about a^ε is needed for an estimate of the modeling error.

Macro error. Let u^0 be the homogenized solution of problem (2.16) and $u^{0,H}$ be its FEM approximation given by problem (2.29). Assume that the bilinear form (2.30) is based on a quadrature formula satisfying (Q1) and (Q2), and that (2.10) holds. Assume that $u^0 \in H^{p+1}(\Omega)$, and that $a^0 \in (W^{p+m,\infty}(\Omega))^{d \times d}$, with $m = 0, 1$ then

$$\begin{aligned} e_{\text{MAC}, H^1} &= \|u^0 - u^{0,H}\|_{H^1(\Omega)} \leq CH^p, \quad \text{for } m = 0 \\ e_{\text{MAC}, L^2} &= \|u^0 - u^{0,H}\|_{L^2(\Omega)} \leq CH^{p+1}, \quad \text{for } m = 1. \end{aligned}$$

Micro error. For the micro error, some assumptions on the micro functions $\psi_{K_{\delta_j}}^i$, the solutions of (2.32), are needed; i.e.,

$$(H1) \quad |\psi_{K_{\delta_j}}^i|_{H^{q+1}(K_{\delta_j})} \leq C\varepsilon^{-q}|K_{\delta_j}|^{1/2}, \text{ for all } i = 1, \dots, d, q \in \mathbb{N}, \text{ and where } C \text{ is independent of } \varepsilon, x_{j,K}, \text{ and } K_{\delta_j}.$$

Let \bar{u}^H and u^H be the solutions of problems (2.34) and (2.25), respectively, with the same coupling conditions (either periodic (2.21) or Dirichlet (2.22)). Then, assuming (H1), it holds

$$e_{\text{MIC}} = \|\bar{u}^H - u^H\|_{H^1(\Omega)} \leq \left(\frac{h}{\varepsilon}\right)^{2q},$$

where q is the degree of the micro FE space $V^q(K_{\delta_j}, \mathcal{T}_h)$.

Modeling error. First, we assume that a^ε is locally periodic in the fast variable y ; i.e.,

$$(H2) \quad a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y) \text{ is } Y\text{-periodic in } y.$$

Let $u^{0,H}$ and \bar{u}^H be given by problems (2.29) and (2.34), respectively. Then, assuming (H2), the modeling error $e_{\text{MOD}} = \|u^{0,H} - \bar{u}^H\|_{H^1(\Omega)}$ can be bounded by

$$\begin{aligned} e_{\text{MOD}} &\leq C\varepsilon, & \text{if } W(K_{\delta_j}) = H_{\text{per}}^1(K_{\delta_j}), \quad \delta/\varepsilon \in \mathbb{N}, \\ e_{\text{MOD}} &\leq C(\delta + \frac{\varepsilon}{\delta}), & \text{if } W(K_{\delta_j}) = H_0^1(K_{\delta_j}), \quad \delta/\varepsilon \notin \mathbb{N}, \delta > \varepsilon. \end{aligned} \quad (2.36)$$

Further, if we collocate the slow variable x in a^ε to the quadrature points $x_{j,K}$, i.e., we replace $a(x, x/\varepsilon)$ by $a(x_{j,K}, x/\varepsilon)$ in the macro (2.26) and micro (2.27) bilinear forms, we obtain

$$\begin{aligned} e_{\text{MOD}} &= 0, & \text{if } W(K_{\delta_j}) = H_{\text{per}}^1(K_{\delta_j}), \quad \delta/\varepsilon \in \mathbb{N}, \\ e_{\text{MOD}} &\leq C\frac{\varepsilon}{\delta}, & \text{if } W(K_{\delta_j}) = H_0^1(K_{\delta_j}), \quad \delta/\varepsilon \notin \mathbb{N}, \delta > \varepsilon. \end{aligned} \quad (2.37)$$

We can give a fully discrete analysis of the FE-HMM, and refer to [3, 4] and the references therein for complete analysis and proofs.

Theorem 2.5.1. *Let u^0, u^H be the solutions of (2.16) and (2.25) respectively. Further suppose that $u^0 \in H^{p+1}(\Omega)$. Assume that the bilinear form (2.30) is based on a QF satisfying (Q1) and (Q2), and that (2.10) and (H1) hold. Let $a^0 \in (W^{p+m,\infty}(\Omega))^{d \times d}$, for $m = 0, 1$. Further, assume that the same coupling conditions are used in the semi-discrete problem (2.34) and in the FE-HMM (2.25), we have*

$$\begin{aligned} \|u^0 - u^H\|_{H^1(\Omega)} &\leq C \left(H^p + \left(\frac{h}{\varepsilon}\right)^{2q} + e_{\text{MOD}} \right), & \text{for } m = 0, \\ \|u^0 - u^H\|_{L^2(\Omega)} &\leq C \left(H^{p+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + e_{\text{MOD}} \right), & \text{for } m = 1, \end{aligned}$$

where C is independent of ε, H , and h . If (H2) holds, the modeling error e_{MOD} is given by (2.36) or (2.37), depending on whether we consider collocation or not.

A priori error estimates for the DG-FE-HMM

Similar convergence rates to the FE-HMM can be derived for the DG-FE-HMM. As the effect of numerical quadrature in DG-FEM for elliptic problems with variable coefficients has not been treated yet (to the best of our knowledge), we will assume that the tensor a^0 is piecewise constant in any $K \in \mathcal{T}_H$; i.e.,

(H3) $a^\varepsilon(x) = a(x, y)$ is Y -periodic in y , and for any $K \in \mathcal{T}_H$, $a(\cdot, y)$ is constant in K .

Then, if the assumptions of Theorem 2.5.1 and hypothesis (H3) hold, we have

$$\begin{aligned} |||u^0 - u^H||| &\leq C \left(H^p + \left(\frac{h}{\varepsilon} \right)^{2q} + e_{\text{MOD}} \right), \quad \text{for } m = 0 \\ \|u^0 - u^H\|_{L^2(\Omega)} &\leq C \left(H^{p+1} + \left(\frac{h}{\varepsilon} \right)^{2q} + e_{\text{MOD}} \right), \quad \text{for } m = 1, \end{aligned}$$

where C is independent of ε, H , and h , and where the modeling error e_{MOD} is given by (2.36) or (2.37), depending on whether we consider collocation or not.

Notice that when the tensor is locally periodic, the optimal choice for the boundary condition in the micro problem (2.32) (for both FE-HMM and DG-FE-HMM) is to use periodic boundary conditions with $\delta/\varepsilon \in \mathbb{N}$.

Post-processing procedure

It is known that the homogenized solution u^0 is a good L^2 -approximation of the heterogeneous solution u^ε , but that u^0 fails to approximate u^ε in H^1 due to the lack of the fine scale information. A post-processing procedure can be done by using the micro functions u_j^h , computed during the FE-HMM procedure, to correct the numerical homogenized solution u^H . In each sampling domain $K_{\delta_j} \subset K$, we can access $u_j^h - u^H$ and periodically extend it to the whole macro element K ,

$$u_H^{rec}(x) = u^H(x) + (u_j^h - u^H)(x - [x]_{K_{\delta_j}}), \quad x \in K,$$

where $[x]_{K_{\delta_j}}$ is the unique combination $\delta \sum_{i=1}^d b_i e_i$, $b_i \in \mathbb{Z}$, such that $(x - [x]_{K_{\delta_j}}) \in K_{\delta_j}$. In order to derive a priori error estimates between the reconstructed solution u_H^{rec} and u^ε , we assume that the tensor a^ε is Y -periodic in y and verifies $a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y)$. If one assume P^1 macro and micro FE spaces and periodic coupling with $\delta = \varepsilon$, the reconstructed solution can be written as

$$u_H^{rec}(x) = u^H(x) + \sum_{i=1}^d \varepsilon \chi^{i,h}(x - [x]_{K_\varepsilon}, x/\varepsilon) \frac{\partial u^H(x)}{\partial x_i}, \quad x \in K.$$

Further details about the reconstruction, as well as the proof of convergence, can be found in [3]. Suppose that the assumptions of Theorem 2.5.1 holds for $p = 1$ and $q = 1$, and that the

Macro FE	u^0	Micro FE	H^1 norm	L^2 norm
P^1	$H^2(\Omega)$	P^1	$\sqrt{N_{mac}} = N_{mic}$	$N_{mac} = N_{mic}$
		P^2	$N_{mac}^{1/4} = N_{mic}$	$\sqrt{N_{mac}} = N_{mic}$
P^2	$H^3(\Omega)$	P^1	$N_{mac} = N_{mic}$	$N_{mac}^{3/2} = N_{mic}$
		P^2	$\sqrt{N_{mac}} = N_{mic}$	$N_{mac}^{3/4} = N_{mic}$

Table 2.1 – Best refinement strategies for optimal convergence rates.

map $x \in \tilde{\Omega} \rightarrow D^\alpha \chi(x, \cdot)$ is Lipschitz continuous for $|\alpha| = 1$, it holds

$$\|u^\varepsilon - u_H^{rec}\|_{\tilde{H}^1(\Omega)} \leq C \left(H + \frac{h}{\varepsilon} + \sqrt{\varepsilon} \right),$$

where the constant $C > 0$ is independent of H, h , and ε .

Costs of the (DG-)FE-HMM

To obtain the optimal convergence rates, one balances the macro and micro errors in the following way [3],

$$H^p \sim \left(\frac{h}{\varepsilon} \right)^{2q}, \quad \text{for the } H^1 \text{ norm}, \quad H^{p+1} \sim \left(\frac{h}{\varepsilon} \right)^{2q}, \quad \text{for the } L^2 \text{ norm},$$

where p and q denote the macro and micro degrees of the macro and micro FE spaces, respectively. Let $\mathcal{T}_H \in \{\mathcal{T}_H\}$ be a coarse partition of Ω and $\mathcal{T}_h \in \{\mathcal{T}_h\}$ be a fine partition of the sampling domains K_{δ_j} , and use N_{mac} (resp. N_{mic}) to denote the total number of degrees of freedom used for the macro (resp. micro) solver. Further, assume that $\delta/\varepsilon \in \mathbb{N}$, and that $\frac{h}{\varepsilon} \sim N_{mic}^{-1}$. To have optimal rates, we set $\frac{h}{\varepsilon} \sim H^{p/(2q)}$ for the H^1 norm, and $\frac{h}{\varepsilon} \sim H^{(p+1)/(2q)}$ for the L^2 norm. Then,

$$\begin{aligned} N_{mic}^d &\sim \left(\frac{h}{\varepsilon} \right)^{-d} \sim H^{-\frac{dp}{2q}} \sim N_{mac}^{\frac{dp}{2q}}, & \text{for the } H^1 \text{ norm}, \\ N_{mic}^d &\sim \left(\frac{h}{\varepsilon} \right)^{-d} \sim H^{-\frac{d(p+1)}{2q}} \sim N_{mac}^{\frac{d(p+1)}{2q}}, & \text{for the } L^2 \text{ norm}, \end{aligned}$$

in particular for $p = q = 1$; $N_{mic} = \sqrt{N_{mac}}$ for the H^1 norm, and $N_{mic} = N_{mac}$, for the L^2 norm. Table 2.1 sums up the optimal convergence rates.

2.6 Numerical results

In here, we test the FE-HMM on a heterogeneous elliptic problem in order to assess the convergence rates given in Theorem 2.5.1. We consider a two-dimensional elliptic problem in $\Omega = [0, 1]^2$,

$$-\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) = f,$$

with zero Dirichlet boundary condition on Γ , $f \equiv 1$, and a tensor $a^\varepsilon(x) = (\cos(2\pi x_1/\varepsilon) + 2)I_2$, with $x = (x_1, x_2) \in \Omega$. The tensor is Y -periodic in the fast-variable, i.e., $a^\varepsilon(x) = a(x/\varepsilon) = a(y)$, $y = (y_1, y_2) \in Y$. Explicit equations are available to compute the homogenized tensor a^0 ,

$$a^0 = \begin{pmatrix} \left(\int_Y \frac{1}{a(y)} dy_1\right)^{-1} & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 2 \end{pmatrix}.$$

A reference homogenized solution u^0 is computed on a very fine mesh with the tensor a^0 . From Theorem 2.5.1, if $u^0 \in H^{p+1}(\Omega)$, the a priori error estimates between u^0 and its FE-HMM approximation u^H are

$$\begin{aligned} \|u^0 - u^H\|_{H^1(\Omega)} &\leq C \left(H^p + \left(\frac{h}{\varepsilon}\right)^{2q} \right), \\ \|u^0 - u^H\|_{L^2(\Omega)} &\leq C \left(H^{p+1} + \left(\frac{h}{\varepsilon}\right)^{2q} \right), \end{aligned}$$

where one notices that the modeling error is zero, i.e., $e_{\text{MOD}} = 0$, as the tensor only depends on the fast variable x/ε . When we refine the macro mesh, the convergence rates are expected to reach a threshold value depending on the micro mesh size. An optimal convergence rate can be obtained if the macro and micro meshes are balanced following the refinement strategies given in Table 2.1. In Figure 2.4a, we plot the H^1 (in red, bullets) and L^2 (in blue, diamonds) errors between u^0 and u^H obtained with P^1 macro and micro FE, with $H = 1/4, 1/8, 1/16, \dots, 1/512$, for micro mesh size of $h/\varepsilon = 1/4, 1/8, 1/16, 1/32, 1/64, 1/128$. In Figure 2.4b, we conduct the same experiment with P^2 macro FE and P^1 micro FE. We see that in both figures the convergence rate follow the a priori results. Further, we see that the optimal refinements follow the ratio $h^{2q} = H^p$ for the H^1 norm and $h^{2q} = H^{p+1}$ for the L^2 norm; in Figures 2.4a and 2.4b, the optimal refinements for the H^1 and L^2 convergence rates are plotted in black (dotted) and we see that the H^1 and L^2 errors follow the optimal refinement.

2.7 Summary

In this chapter, we consider a highly heterogeneous two-scales multiscale elliptic problem and give a review of numerical multiscale methods used to solve such problems. We discuss two important classes of multiscale methods; the methods based on a global extraction of the small scales and the methods based on homogenization theory. In the first class, we reviewed

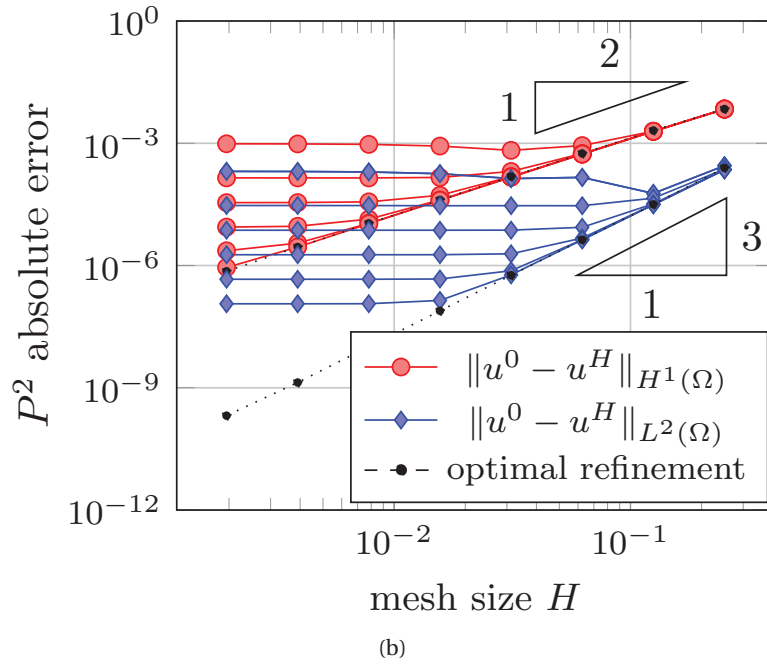
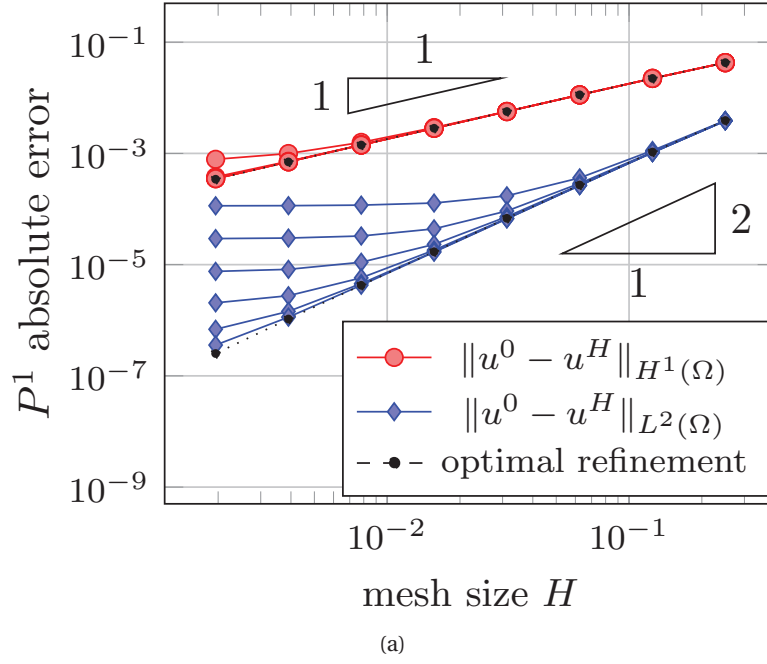


Figure 2.4 – Periodic homogenization, H^1 (bullets, red) and L^2 (diamonds, blue) errors between u^0 and u^H obtained with (a) P^1 macro and micro FE and (b) P^2 macro and P^1 micro FE.

the MsFEM and the LOD methods, and in the second class of methods, we gave a review of the high-dimensional FEM and the FE-HMM. The FE-HMM is given in details with its fully discrete a priori error analysis and a numerical example to verify the convergence rates. The strength of the FE-HMM is that the method has a cost independent of the smallest scale present in the problem. However, it relies on a clear separation of scales. The first class of methods, does not requires that the small scales should be well-separated, but the methods use and process all the fine scale information. This leads to computationally expensive methods.

3 Multiscale method for the wave equation in linear elastic heterogeneous media

In this chapter, we give the finite element heterogeneous multiscale method applied to the wave equation in linear elastic heterogeneous media for short times. This chapter is based on the article [14].

Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, be a domain filled with a linear elastic medium and seek u^ε the solution of

$$\partial_{tt} u^\varepsilon - \operatorname{div}(a^\varepsilon(x) : e(u^\varepsilon)) = f, \quad \text{in } \Omega, \quad (3.1)$$

where $f \in L^2(0, T; L^2(\Omega))^d$ and e is a linearized strain tensor. Further, we consider homogeneous Dirichlet boundary conditions on Γ and initial conditions at the time $t = 0$. The solution $u^\varepsilon = (u_1^\varepsilon, \dots, u_d^\varepsilon)$ corresponds to the displacement of the wave and $u_i^\varepsilon = u_i^\varepsilon(t)$, for all $i = 1, \dots, d$. The superscript $\varepsilon > 0$ in the tensor a^ε denotes the small scales present in the medium. The tensor a^ε is a fourth-order tensor verifying $a_{ijkl}^\varepsilon(x) \in L^\infty(\Omega)$, for $i, j, k, l = 1, \dots, d$, and, for some constants $0 < \lambda \leq \Lambda < +\infty$,

$$a_{ijkl}^\varepsilon = a_{jikl}^\varepsilon = a_{klij}^\varepsilon, \quad (3.2)$$

$$\lambda |M|^2 \leq a^\varepsilon M : M, \quad (3.3)$$

$$|a^\varepsilon M| \leq \Lambda |M|, \quad \text{for any symmetric matrix } M. \quad (3.4)$$

In addition, for any square matrices M, M_1 , and M_2 , we set

$$a^\varepsilon M_1 : M_2 = \sum_{i,j,k,l=1}^d a_{ijkl}^\varepsilon M_{1ij} M_{2kl},$$

$$|M| = (M : M)^{1/2} = \left(\sum_{i,j=1}^d M_{ij}^2 \right)^{1/2}.$$

A weak solution u^ε of problem (3.1) is obtained from the variational formulation

$$B^\varepsilon(u^\varepsilon, w) = F(w), \quad \forall w \in H_0^1(\Omega)^d.$$

Chapter 3. Multiscale method for the wave equation in linear elastic heterogeneous media

The bilinear form $B^\varepsilon : H^1(\Omega)^d \times H^1(\Omega)^d \rightarrow \mathbb{R}$ is given by

$$B^\varepsilon(v, w) = \langle \partial_{tt} v, w \rangle + \int_{\Omega} a^\varepsilon(x) e(v) : e(w) dx,$$

where $\langle \cdot, \cdot \rangle$ denotes the dual product $\langle \cdot, \cdot \rangle_{H^{-1}, H^1}$. The right hand side $F : H^1(\Omega)^d \rightarrow \mathbb{R}$ is defined by

$$F(w) = \int_{\Omega} f w dx.$$

When the boundary conditions are non-zero, additional terms are present in the weak formulation, involving the lifting of the Dirichlet data and the Neumann boundary conditions.

Outline. In the first part of this chapter, we drop the time dependency and focus on the linear elasticity problem with highly oscillatory coefficients. We give the FE-HMM and derive a priori error estimates. Numerical examples are given to assess the convergence rates. In the second part of this chapter, we consider the wave equation in a linear elastic medium. A priori error estimates are given and various numerical examples, inspired from seismic problems, are proposed.

Heterogeneous multiscale method for linear elasticity problems

In here, we introduce the heterogeneous multiscale method for linear elasticity problems, following the FE-HMM derived by Abdulle in [2]. This part is based on the article [14].

Outline. In Section 3.1 we briefly state the homogenization theory applied to multiscale linear elasticity problems. In Section 3.2 we give the finite element heterogeneous multiscale method and in Section 3.3, we derive the fully discrete a priori error analysis. This part ends with some numerical examples given in Section 3.4.

We consider the linear elasticity equation in a bounded domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, with homogeneous Dirichlet conditions on a Lipschitz continuous boundary Γ . The problem reads: find u^ε such that

$$\begin{aligned} -\frac{\partial}{\partial x_j} \left(a_{ijkl}^\varepsilon(x) \frac{\partial u_k^\varepsilon}{\partial x_l} \right) &= f_i, \quad \text{in } \Omega, \\ u_i^\varepsilon &= 0, \quad \text{on } \Gamma, \end{aligned} \quad (3.5)$$

for $i = 1, \dots, d$ and where the right-hand side f is in $L^2(\Omega)^d$. The superscript $\varepsilon > 0$ denotes, again, the multiscale nature of the problem, and the tensor a^ε is a fourth-order tensor, with $a_{ijkl}^\varepsilon(x) \in L^\infty(\Omega)$ for $i, j, k, l = 1, \dots, d$. The tensor verifies the assumptions (3.2), (3.3), and (3.4). We define the linearized strain tensor e , for $i, j = 1, \dots, d$, by

$$e(u) = (e_{ij}(u))_{1 \leq i, j \leq d}, \quad e_{ij}(u) = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right).$$

The weak formulation of problem (3.5) reads: find $u^\varepsilon \in H_0^1(\Omega)^d$ such that

$$B^\varepsilon(u^\varepsilon, w) = F(w), \quad \forall w \in H_0^1(\Omega)^d. \quad (3.6)$$

where the bilinear form $B^\varepsilon : H^1(\Omega)^d \times H^1(\Omega)^d \rightarrow \mathbb{R}$ and the right-hand side $f : H^1(\Omega)^d \rightarrow \mathbb{R}$ are given by

$$B^\varepsilon(v, w) = \int_{\Omega} a^\varepsilon(x) e(v) : e(w) dx, \quad \text{and } F(w) = \int_{\Omega} f w dx.$$

The weak formulation (3.6) is well-posed and therefore admits a family of unique solutions

$\{u^\varepsilon\}$, indexed by the subscript ε . Further it holds

$$\|u^\varepsilon\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)},$$

where the norm over $H^1(\Omega)^d$ is given by

$$\|u\|_{H^1(\Omega)} = \left(\sum_{i,j=1}^d \int_{\Omega} \left(\frac{\partial u_i}{\partial x_j} \right)^2 dx + \sum_{i=1}^d \int_{\Omega} u_i^2 dx \right)^{1/2}, \quad u \in H^1(\Omega)^d.$$

The existence and uniqueness of u^ε follows from the first Korn inequality; i.e.,

$$\|v\|_{H^1(\Omega)} \leq C \left(\int_{\Omega} |e(v)|^2 dx \right)^{1/2}. \quad (3.7)$$

Indeed, using the coercivity of the tensor a^ε and the first Korn inequality (3.7), one can show that the bilinear form B^ε is coercive; i.e.,

$$B^\varepsilon(v, v) = \int_{\Omega} a^\varepsilon(x) e(v) : e(v) dx \geq C \left(\int_{\Omega} |e(v)|^2 dx \right) \geq C \|v\|_{H^1(\Omega)}^2, \quad \forall v \in H_0^1(\Omega)^d.$$

We define a norm $\|\cdot\|$ over $H_0^1(\Omega)^d$ by

$$\|v\| = \left(\int_{\Omega} |e(v)|^2 dx \right)^{1/2}. \quad (3.8)$$

For the proof that $\|\cdot\|$ is indeed a norm over $H_0^1(\Omega)^d$, we refer to [96, Theorem 2.5].

As a consequence of the Korn inequalities, we have the following equivalence, see [96, 80]; there exist two constants $C_1, C_2 > 0$ such that

$$C_1 \|v\|_{H^1(\Omega)} \leq \|v\| \leq C_2 \|v\|_{H^1(\Omega)}, \quad v \in H_0^1(\Omega)^d.$$

Remark 3.0.1. Non-homogeneous Dirichlet conditions and Neumann boundary conditions can be considered instead of homogeneous Dirichlet boundary conditions. The existence and uniqueness of u^ε is proved using the second Korn inequality [80, Theorem 2.4]; i.e.,

$$\|v\|_{H^1(\Omega)} \leq C \left(\|v\|_{L^2(\Omega)}^2 + \left(\int_{\Omega} |e(v)|^2 dx \right) \right)^{1/2}.$$

3.1 Homogenization of linear elasticity problems and basic results

Solving (3.6) with standard FEM requires the mesh size to be smaller than the fine scales, which is prohibitive if ε is small. However, the effective dynamics of the problem can be described using the homogenization theory. The homogenization of an elliptic partial differential equation in a linear elastic medium is treated in [52, 96, 102, 99] and the references therein.

From the theory of H -convergence [90, 22], it can be established that a subsequence of the family of solutions $\{u^\varepsilon\}$ converges weakly to an effective solution u^0 , satisfying the homogenized formulation

$$B^0(u^0, w) = F(w), \quad \forall w \in H_0^1(\Omega)^d, \quad (3.9)$$

where $B^0 : H^1(\Omega)^d \times H^1(\Omega)^d \rightarrow \mathbb{R}$ is defined by

$$B^0(v, w) = \int_{\Omega} a^0(x) e(v) : e(w) dx.$$

The homogenized tensor a^0 verifies the properties (3.2), (3.3), and (3.4) for some constants $0 < \lambda^0 \leq \Lambda^0 \leq \infty$. As in Chapter 2, no explicit equations are available, in general, for the homogenized tensor a^0 ; however, under additional information on the small scale of the tensor a^ε , such as periodicity, it holds

$$a_{ijkl}^0(x) = \frac{1}{|Y|} \int_Y a_{ijkl}(x, y) + \sum_{h,m=1}^d a_{ijhm}(x, y) \frac{\partial \chi_h^{kl}(x, y)}{\partial y_m} dy, \quad x \in \Omega,$$

where $Y = [0, 1]^d$ is the d -dimensional hypercube. The functions $\chi^{kl} \in W_{per}(Y)$ are the solutions of the micro problems

$$-\frac{\partial}{\partial y_j} \left(a_{ijhm} \frac{\partial \chi_h^{kl}}{\partial y_m} \right) = \frac{\partial a_{ijkl}}{\partial y_j}, \text{ in } Y, \text{ for } i = 1, \dots, d, \quad (3.10)$$

with periodic boundary conditions. In weak form, the micro problem (3.10) reads: find $\chi^{kl} \in W_{per}(Y)$ solution of

$$\int_Y a(x, y) e(\chi^{kl}) : e(z) dy = \int_Y a(x, y) e(I^{kl}) : e(z) dy, \quad \forall z \in W_{per}(Y), \quad (3.11)$$

where $I^{kl} = (I_h^{kl})$ is given by

$$I_h^{kl} = y_l \delta_{hk}.$$

The micro problems (3.11) are well-posed thanks to the first Korn inequality for the periodic case [96]; i.e.,

$$\|v\|_{H^1(Y)} \leq C \left(\int_Y |e(v)|^2 dy \right)^{1/2}.$$

3.2 FE-HMM for linear elasticity

In this section, we derive the multiscale FEM for the problem (3.5). The FE-HMM gives us a macroscopic solution, approximation of u^0 , based on a macro to micro procedure without knowing the homogenized tensor a^0 . The FE-HMM is explained in details in Section 2.4 for highly heterogeneous multiscale elliptic PDEs.

Chapter 3. Multiscale method for the wave equation in linear elastic heterogeneous media

Macro problem. Let $\{\mathcal{T}_H\}$ be a family of admissible (T1) and shape-regular (T2) partitions over Ω with mesh size $H \gg \varepsilon$ given by $H = \max_{K \in \mathcal{T}_H} h_K$. In each macro element K , we consider integration nodes $x_{j,K}$ and weights $\omega_{j,K}$, for $j = 1, \dots, J$, and construct sampling domains $K_{\delta_j} = x_{j,K} + \delta[-1/2, 1/2]^d$, with $0 < \delta \ll H$. To ensure that a FEM with numerical quadrature converges to the exact solution with the rates obtained from a FEM with exact integration, we assume that the QF over each K is induced by a QF over a reference element \hat{K} and that they verify the assumptions (Q1) and (Q2).

For each partition \mathcal{T}_H in $\{\mathcal{T}_H\}$, we define a macro FE space of degree $p \geq 1$ by

$$V_0^p(\Omega, \mathcal{T}_H) = \{v^H \in H_0^1(\Omega)^d \mid v^H|_K \in \mathcal{P}^p(K)^d, \quad \forall K \in \mathcal{T}_H\}, \quad (3.12)$$

where $\mathcal{P}^p(K)$ is the space $\mathcal{P}^p(K)$ of polynomials on K of degree at most p if K is simplicial, or the space $\mathcal{Q}^p(K)$ of polynomials on K of degree at most p in each variables if K is rectangular. We construct a macro bilinear form $B_H : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ by

$$B_H(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(v_j^h) : e(w_j^h) dx, \quad (3.13)$$

where v_j^h (resp. w_j^h) is the solution of the micro problem (3.16) on the sampling domain K_{δ_j} . The FE-HMM reads: find u^H in $V_0^p(\Omega, \mathcal{T}_H)$ such that

$$B_H(u^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H). \quad (3.14)$$

Micro problem. Let $\{\mathcal{T}_h\}$ be a family of admissible (T1) and shape-regular (T2) partitions over K_{δ_j} , for $j = 1, \dots, J$, of mesh size $h \leq \varepsilon$, with $h = \max_{K \in \mathcal{T}_h} h_K$. For each micro partition \mathcal{T}_h , we define a micro FE space of degree $q \geq 1$, on the sampling domain K_{δ_j} , as

$$V^q(K_{\delta_j}, \mathcal{T}_h) = \{v^h \in W(K_{\delta_j}) \mid v^h|_K \in \mathcal{P}^q(K)^d, \quad \forall K \in \mathcal{T}_h\}. \quad (3.15)$$

The micro problems read: find u_j^h such that $(u_j^h - u_{\text{lin},j}^H) \in V^q(K_{\delta_j}, \mathcal{T}_h)$ and

$$\int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(z^h) dx = 0, \quad \forall z^h \in V^q(K_{\delta_j}, \mathcal{T}_h), \quad (3.16)$$

where $u_{\text{lin},j}^H(x) = u^H(x_{j,K}) + (x - x_{j,K}) e(u^H(x_{j,K}))$ is a linearization of u^H taken at the quadrature node $x_{j,K}$. Here, we use again the notation $u_{\text{lin},j}^H$ to denote $u_{\text{lin},K_{\delta_j}}^H$. We recall that the space $W(K_{\delta_j})$ sets the coupling between the micro and macro solvers and depends on the choice of boundary conditions in problem (3.16); i.e.,

$$W(K_{\delta_j}) = H_{\text{per}}^1(K_{\delta_j})^d, \quad \text{for periodic coupling}, \quad (3.17)$$

$$W(K_{\delta_j}) = H_0^1(K_{\delta_j})^d, \quad \text{for Dirichlet coupling}. \quad (3.18)$$

The FE-HMM problem (3.14) is well-posed and admits a unique solution. The following Proposition has been proved in [2] for piecewise linear FE, and we give here the proof for FE of order $p \geq 1$.

Proposition 3.2.1. *The problem (3.14) has a unique solution $u^H \in V_0^p(\Omega, \mathcal{T}_H)$ which verifies*

$$\|u^H\|_{H^1(\Omega)} \leq C \|f\|_{L^2(\Omega)},$$

where C is independent of ε , H , and h .

Proof. We prove the well-posedness by showing that the bilinear form B_H (3.13) is coercive and bounded. Let $u^H \in V_0^p(\Omega, \mathcal{T}_H)$, then $u_j^h - u_{\text{lin},j}^H$ is in the space $V^q(K_{\delta_j}, \mathcal{T}_h)$ and can be used as a test function in the micro problems (3.16); i.e.,

$$\int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(u_j^h - u_{\text{lin},j}^H) dx = 0.$$

By linearity, it holds

$$\int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(u_j^h) dx = \int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(u_{\text{lin},j}^H) dx.$$

Then, using the assumptions (3.2), (3.3), and (3.4) made on the tensor a^ε , we obtain the following bound

$$\int_{K_{\delta_j}} |e(u_j^h)|^2 dx \leq C \int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(u_j^h) dx \leq C \int_{K_{\delta_j}} e(u_j^h) : e(u_{\text{lin},j}^H) dx.$$

Using the Cauchy–Schwarz inequality, we obtain

$$\|e(u_j^h)\|_{L^2(K_{\delta_j})} \leq C \|e(u_{\text{lin},j}^H)\|_{L^2(K_{\delta_j})}.$$

The bilinear form B_H defined in (3.13) is bounded; i.e.,

$$\begin{aligned} B_H(u^H, v^H) &\leq C \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \|e(u_j^h)\|_{L^2(K_{\delta_j})} \|e(v_j^h)\|_{L^2(K_{\delta_j})} \\ &\leq C \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \|e(u_{\text{lin},j}^H)\|_{L^2(K_{\delta_j})} \|e(v_{\text{lin},j}^H)\|_{L^2(K_{\delta_j})} \\ &\leq C \|u^H\|_{H^1(\Omega)} \|v^H\|_{H^1(\Omega)}, \end{aligned}$$

where we used that $e(u_{\text{lin},j}^H(x)) = e(u^H(x_{j,K}))$, for all $x \in K$. We now show that the bilinear form B_H is coercive. Following [2, Lemma 4.3] using that $e(u_{\text{lin},j}^H)$ is constant in K , that the

difference $u_j^h - u_{\text{lin},j}^H$ is in $V^q(K_{\delta_j}, \mathcal{T}_h)$, and that $u_j^h - u_{\text{lin},j}^H = 0$ on ∂K_{δ_j} , it holds

$$\begin{aligned} \int_{K_{\delta_j}} e(u_j^h) : e(u_j^h) dx &= \int_{K_{\delta_j}} (e(u_j^h) - e(u_{\text{lin},j}^H)) : (e(u_j^h) - e(u_{\text{lin},j}^H)) dx + \int_{K_{\delta_j}} e(u_{\text{lin},j}^H) : e(u_{\text{lin},j}^H) dx \\ &\geq \int_{K_{\delta_j}} e(u_{\text{lin},j}^H) : e(u_{\text{lin},j}^H) dx. \end{aligned}$$

Then, B_H is coercive; i.e.,

$$\begin{aligned} B_H(u^H, u^H) &\geq C \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \|e(u_{\text{lin},j}^H)\|_{L^2(K_{\delta_j})}^2 \\ &\geq C \sum_{K \in \mathcal{T}_H} \|e(u^H)\|_{L^2(K)}^2 \\ &\geq C \|u^H\|_{H^1(\Omega)}^2. \end{aligned}$$

The existence and uniqueness of a solution u^H of problem (3.14) is a consequence of the Lax–Milgram lemma. \square

3.3 A priori error analysis

In this section we give a priori error estimates for the FE-HMM method applied to linear elasticity problems; note that such results have been first derived in [2] for piecewise linear FE. The error is decomposed into a macro, modeling, and micro error,

$$\|u^0 - u^H\| \leq e_{\text{MAC}} + e_{\text{MOD}} + e_{\text{MIC}},$$

where the norm $\|\cdot\|$ stands for the L^2 norm or the H^1 norm.

Macro error. Let $u^{0,H} \in V_0^p(\Omega, \mathcal{T}_H)$ be an approximation of the exact solution u^0 , obtained by solving the homogenized problem (3.9) using the FEM with a numerical quadrature verifying the assumptions (Q1) and (Q2). Assuming that $a_{ijkl}^0(x) \in W^{1,\infty}(\Omega)$, for $i, j, k, l = 1, \dots, d$, the problem reads: find $u^{0,H} \in V_0^p(\Omega, \mathcal{T}_H)$ such that

$$B_H^0(u^{0,H}, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (3.19)$$

where the bilinear form $B_H^0 : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ is defined by

$$B_H^0(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} a^0(x_{j,K}) e(v^H(x_{j,K})) : e(w^H(x_{j,K})).$$

The macro error is given by difference between u^0 and $u^{0,H}$. Let u^0 be the homogenized solution of problem (3.9) and let $u^{0,H} \in V_0^p(\Omega, \mathcal{T}_H)$ be the solution of (3.19). Further, assuming

that $u^0 \in H^{p+1}(\Omega)$, and that $a_{ijkl}^0 \in (W^{p+m,\infty}(\Omega))$, for $i, j, k, l = 1, \dots, d$ and $m = 0, 1$, it holds

$$\begin{aligned} e_{\text{MAC}, H^1} &= \|u^0 - u^{0,H}\|_{H^1(\Omega)} \leq CH^p, & \text{for } m = 0, \\ e_{\text{MAC}, L^2} &= \|u^0 - u^{0,H}\|_{L^2(\Omega)} \leq CH^{p+1}, & \text{for } m = 1, \end{aligned}$$

where the constant is independent of H, h , and ε .

Further, we can derive an estimate for the difference between the bilinear forms B^0 and B_H^0 .

Proposition 3.3.1. *Let $v^H, w^H \in V_0^p(\Omega, \mathcal{T}_H)$ and $a_{ijkl}^0 \in W^{p+m,\infty}(\Omega)$ for all $i, j, k, l = 1, \dots, d$ and $m = 0, 1$, it holds*

$$|B^0(v^H, w^H) - B_H^0(v^H, w^H)| \leq CH^{p+m} \max_{i,j,k,l} \|a_{ijkl}^0\|_{W^{p+m,\infty}(\Omega)} \|v^H\|_{\tilde{H}^{p+m}(\Omega)} \|w^H\|_{\tilde{H}^{1+m}(\Omega)},$$

where C is independent of H, h , and ε .

Proof. see [41, Chapter 4] and [42]. □

Micro Error. We focus now on the error made in the discretization of the micro problems. Let us consider a semi-discrete problem, where the micro solutions are taken in the exact Sobolev spaces, and look for the solution $\bar{u}^H \in V_0^p(\Omega, \mathcal{T}_H)$ of

$$\bar{B}_H(\bar{u}^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (3.20)$$

where the bilinear form \bar{B}_H is given by

$$\bar{B}_H(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(v_j) : e(w_j) dx, \quad (3.21)$$

where $v_j, w_j \in W(K_{\delta_j})$ are the solutions of (3.16). We assume that the micro solutions χ^{lm} (the solutions of equation (3.10)) are smooth enough; i.e.,

$$(H1) \quad \varepsilon \chi^{lm} \in H^{q+1}(K_{\delta_j})^d \text{ with } \|D^\alpha(\varepsilon \chi^{lm})\|_{L^\infty(K_{\delta_j})} \leq C\varepsilon^{-|\alpha|+1}, \text{ for } \alpha \leq q+1, l, m = 1, \dots, d.$$

Proposition 3.3.2. *Let $v^H, w^H \in V_0^p(\Omega, \mathcal{T}_H)$, and consider the same coupling condition (either (3.17) or (3.18)) for the micro problem in the discrete and semi-discrete problems. Suppose that $\varepsilon \chi^{lm} \in H^{q+1}(K_{\delta_j})^d$ and that assumption (H1) holds for $\alpha = q+1$, with $q > 1$. Then,*

$$|\bar{B}_H(v^H, w^H) - B_H(v^H, w^H)| \leq C \left(\frac{h}{\varepsilon} \right)^{2q} \|v^H\|_{H^1(\Omega)} \|w^H\|_{H^1(\Omega)}, \quad \forall v^H, w^H \in V_0^p(\Omega, \mathcal{T}_H)$$

where C is independent of H, h , and ε . The bilinear forms \bar{B}_H and B_H are defined in (3.21) and (3.13), respectively.

Chapter 3. Multiscale method for the wave equation in linear elastic heterogeneous media

Proof. Using the Cauchy–Schwarz inequality and the ellipticity of a^ε , we obtain

$$\begin{aligned} |\bar{B}_H(v^H, w^H) - B_H(v^H, w^H)| &= \left| \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) (e(v_j) : e(w_j) - e(v_j^h) : e(w_j^h)) dx \right| \\ &\leq C \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \|e(v_j^h) - e(v_j)\|_{L^2(K_{\delta_j})} \|e(w_j^h) - e(w_j)\|_{L^2(K_{\delta_j})}, \end{aligned}$$

and where we used that v_j, w_j are the solutions of (3.16). We bound each term following [2, Lemma 4.3] and [3, Lemma 10]. We obtain

$$\begin{aligned} \int_{K_{\delta_j}} |e(v_j^h) - e(v_j)|^2 dx &\leq C |e(v_{\text{lin},j}^H)|^2 h^{2q} |K_{\delta_j}| \max_{k,l,m} \|\varepsilon \chi_k^{lm}\|_{W^{q+1,\infty}(K_{\delta_j})}^2 \\ &\leq C \left(\frac{h}{\varepsilon} \right)^{2q} |K_{\delta_j}| |e(v_{\text{lin},j}^H)|^2. \end{aligned}$$

Using the fact that $e(v_{\text{lin},j}^H)$ is constant in K and that the norm $\|\cdot\|$ is equivalent to $\|\cdot\|_{H^1}$, we can conclude that

$$|\bar{B}_H(v^H, w^H) - B_H(v^H, w^H)| \leq C \left(\frac{h}{\varepsilon} \right)^{2q} \|v^H\|_{H^1(\Omega)} \|w^H\|_{H^1(\Omega)}.$$

□

The micro error is given by the error between the FE-HMM solution u^H and the semi-discrete solution \bar{u}^H , see [2, 3] for details. Let u^H and \bar{u}^H be given by (3.14) and (3.20), respectively, and assume the same coupling conditions (either (3.17) or (3.18)) for the micro problems. Further, suppose that the assumptions made in Proposition 3.3.2 hold. Then,

$$e_{\text{MIC}} = \|u^H - \bar{u}^H\|_{H^1(\Omega)} \leq C \left(\frac{h}{\varepsilon} \right)^{2q}, \quad (3.22)$$

where the constant C is independent of H, h , and ε .

Modeling Error. We gave bounds for the macro error $\|u^0 - u^{0,H}\|$ and the micro error $\|\bar{u}^H - u^H\|$, thus it remains to bound the error between $u^{0,H}$ and \bar{u}^H , which corresponds to the so-called modeling error. Notice that, for the macro and micro errors, no assumptions were made on the micro scale in the tensor a^ε . To derive explicit bounds on the micro error, we assume that the tensor is (locally) periodic,

$$(H2) \quad a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y) \text{ is } Y\text{-periodic in } y, \text{ where } Y = (0, 1)^d.$$

If an explicit separation of scale is present in the tensor, we can collocate the slow variable x at the quadrature nodes $x_{j,K}$ in the tensor a^ε . By doing so, we define a semi-discrete bilinear

form $\tilde{B}_H : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ by

$$\tilde{B}_H(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a(x_{j,K}, x/\varepsilon) e(v_j) : e(w_j) dx,$$

where $v_j, w_j \in W(K_{\delta_j})$ are the solutions of the cell problems (3.16) with $a^\varepsilon(x)$ replaced by $a(x_{j,K}, x/\varepsilon)$. We define \tilde{u}^H to be the solution of

$$\tilde{B}_H(\tilde{u}^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H). \quad (3.23)$$

The modeling error can then be given. Let \tilde{u}^H and \tilde{u}^H be the solutions of problems (3.20) and (3.23) respectively, with periodic coupling conditions. Suppose that (H2) holds and that $\delta/\varepsilon \in \mathbb{N}$. If $a_{ijkl}(x, y) \in W^{1,\infty}(\Omega, L^\infty(Y))$, for all $i, j, k, l = 1, \dots, d$, then

$$u^{0,H} = \tilde{u}^H, \text{ and } e_{\text{MOD}} = \|u^{0,H} - \tilde{u}^H\|_{H^1(\Omega)} \leq C\varepsilon,$$

where the constant C is independent of H, h , and ε , and where $u^{0,H}$ is the solution of problem (3.19).

As in the a priori error analysis for elliptic problem (2.5), it holds

$$\begin{aligned} e_{\text{MOD}} &\leq C\varepsilon, & \text{if } W(K_{\delta_j}) &= H_{\text{per}}^1(K_{\delta_j})^d, \quad \delta/\varepsilon \in \mathbb{N}, \\ e_{\text{MOD}} &\leq C(\delta + \frac{\varepsilon}{\delta}), & \text{if } W(K_{\delta_j}) &= H_0^1(K_{\delta_j})^d, \quad \delta/\varepsilon \notin \mathbb{N}, \delta > \varepsilon. \end{aligned} \quad (3.24)$$

Further, if we collocate the slow variable x in the tensor a^ε to the quadrature points $x_{j,K}$ in the macro (3.13) and micro (3.16) bilinear forms, we obtain

$$\begin{aligned} e_{\text{MOD}} &= 0, & \text{if } W(K_{\delta_j}) &= H_{\text{per}}^1(K_{\delta_j})^d, \quad \delta/\varepsilon \in \mathbb{N}, \\ e_{\text{MOD}} &\leq C\frac{\varepsilon}{\delta}, & \text{if } W(K_{\delta_j}) &= H_0^1(K_{\delta_j})^d, \quad \delta/\varepsilon \notin \mathbb{N}, \delta > \varepsilon. \end{aligned} \quad (3.25)$$

Considering $\delta/\varepsilon \notin \mathbb{N}$ leads to boundary layers and a deterioration of the modeling error.

Collecting all the previous results leads to the following theorem.

Theorem 3.3.3. *Let u^0 and u^H be the solutions of (3.9) and (3.14), respectively. Assume that $u^0 \in H^{p+1}(\Omega)^d$, and that the hypothesis (H1) holds. Further, assume that the hypothesis for the macro error and proposition 3.3.2 hold. Then,*

$$\begin{aligned} \|u^0 - u^H\|_{H^1(\Omega)} &\leq C \left(H^p + \left(\frac{h}{\varepsilon} \right)^{2q} + e_{\text{MOD}} \right), \\ \|u^0 - u^H\|_{L^2(\Omega)} &\leq C \left(H^{p+1} + \left(\frac{h}{\varepsilon} \right)^{2q} + e_{\text{MOD}} \right). \end{aligned}$$

If in addition, the hypothesis (H2) holds, the modeling error e_{MOD} is given by (3.24) or (3.25).

Recovery of the homogenized tensor

The homogenized tensor can be approximated during the assembling process of the FE-HMM, and one obtains error estimates between the exact homogenized tensor and its numerical approximation. This is done by following the lines of [3, Section 3.3.2.] and [2, Section 5.].

For general symmetric tensors and sampling domains, we can define, at each quadrature point $x_{j,K}$, two tensors $a_K^{0,h} = (a_{iklm}^{0,h}(x_{j,K}))$ and $\bar{a}_K^0 = (\bar{a}_{iklm}^0(x_{j,K}))$, for $i, k, l, m = 1, \dots, d$; i.e.,

$$a_{iklm}^{0,h}(x_{j,K}) = \frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(\varphi_{jik}^h) : e(\varphi_{jlm}^h) dx, \quad (3.26)$$

where the functions $\varphi_{jik}^h \in V^q(K_{\delta_j}, \mathcal{T}_h)$ are the solutions of (3.16) for $i, k \in \{1, \dots, d\}$, and

$$\bar{a}_{iklm}^0(x_{j,K}) = \frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(\varphi_{jik}) : e(\varphi_{jlm}) dx,$$

where φ_{jik} are the solutions of (3.16) in the exact Sobolev space $W(K_{\delta_j})$.

Theorem 3.3.4. *Let $a_K^{0,h}$ be the numerical tensor (3.26) computed with the FE-HMM using micro FE of order $q \geq 1$ and assume that (H2) holds. Then*

$$|a_{iklm}^{0,h}(x_{j,K}) - \bar{a}_{iklm}^0(x_{j,K})| \leq C \left(\left(\frac{h}{\varepsilon} \right)^{2q} + e_{\text{MOD}} \right),$$

where the modeling error is given by (3.24) or (3.25).

Proof. As $a_K^{0,h}$ and $e(v^H(x_{j,K}))$ are constant in the macro element K , it holds

$$\frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(v_j^h) : e(w_j^h) dx = \frac{1}{|K|} \int_K a_K^{0,h}(x) e(v^H(x_{j,K})) : e(w^H(x_{j,K})) dx$$

Then, similarly,

$$\frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(\varphi_{jik}^h) : e(\varphi_{jlm}^h) dx = \frac{1}{|K|} \int_K a_K^{0,h}(x) e(\varphi_{i,k}^H(x_{j,K})) : e(\varphi_{l,m}^H(x_{j,K})) dx,$$

for $i, k, l, m = 1, \dots, d$. The result follows by a triangular inequality

$$\begin{aligned} |a_{iklm}^{0,h}(x_{j,K}) - \bar{a}_{iklm}^0(x_{j,K})| &\leq |a_{iklm}^{0,h}(x_{j,K}) - \bar{a}_{iklm}^0(x_{j,K})| \\ &\quad + |\bar{a}_{iklm}^0(x_{j,K}) - a_{iklm}^0(x_{j,K})|, \end{aligned}$$

and by noting that

$$|a_{iklm}^{0,h}(x_{j,K}) - \bar{a}_{iklm}^0(x_{j,K})| = \left| \frac{1}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) (e(\varphi_{jik}^h) : e(\varphi_{jlm}^h) - e(\varphi_{jik}) : e(\varphi_{jlm})) dx \right|.$$

3.4. Numerical results for the FE-HMM applied to linear elasticity problems

The difference between \bar{a}_K^0 and a^0 follows from the lines of [3, Theorem 12] and Proposition 3.3.2. \square

3.4 Numerical results for the FE-HMM applied to linear elasticity problems

In this section we present numerical examples to verify the sharpness of the bounds obtained in Theorems 3.3.3 and 3.3.4. The numerical results presented here can be found in [14].

In Table 3.1, we recall the best refinement strategies, derived in Section 2.5, for the optimal H^1 and L^2 convergence rates with minimal computational cost.

Outline of Section 3.4. At first, we show that the macro convergence rates in H are sharp when using piecewise and quadratic finite elements spaces. At second, we investigate the effect of the micro error on the convergence rates. Then, we look at the influence of the modeling error in the total error, and ends this section by showing that the micro convergence rates are sharp in the micro mesh size h .

In the experiments, we take a periodic tensor $a^\varepsilon(x) = a(x/\varepsilon) = a(y)$ given by

$$a(y) = \begin{pmatrix} \sin(2\pi y_1) + 2 & 0 & 0 \\ 0 & \sin(2\pi y_2) + 2 & 0 \\ 0 & 0 & 10 \end{pmatrix}.$$

In that case, explicit equations are available to compute a^0 (see [80, 44]), and one obtains

$$a^0 = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 10 \end{pmatrix}.$$

Let $u^0 \in H^{p+1}(\Omega)$, and $u^H \in V_0^p(\Omega, \mathcal{T}_H)$, it holds

$$\begin{aligned} \|u^0 - u^H\|_{H^1(\Omega)} &\leq C(H^p + \left(\frac{h}{\varepsilon}\right)^{2q} + e_{\text{MOD}}) \\ \|u^0 - u^H\|_{L^2(\Omega)} &\leq C(H^{p+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + e_{\text{MOD}}). \end{aligned}$$

Notice that the modeling error is either zero (periodic coupling) or ε/δ (Dirichlet coupling) due to collocation of the variable x to the quadrature points $x_{j,K}$ in the tensor a^ε used in the bilinear forms.

Macro FE	u^0	Micro FE	H^1 norm	L^2 norm
P^1	$H^2(\Omega)$	P^1	$\sqrt{N_{mac}} = N_{mic}$	$N_{mac} = N_{mic}$
		P^2	$N_{mac}^{1/4} = N_{mic}$	$\sqrt{N_{mac}} = N_{mic}$
P^2	$H^3(\Omega)$	P^1	$N_{mac} = N_{mic}$	$N_{mac}^{3/2} = N_{mic}$
		P^2	$\sqrt{N_{mac}} = N_{mic}$	$N_{mac}^{3/4} = N_{mic}$

Table 3.1 – Best refinement strategies for optimal convergence rates.

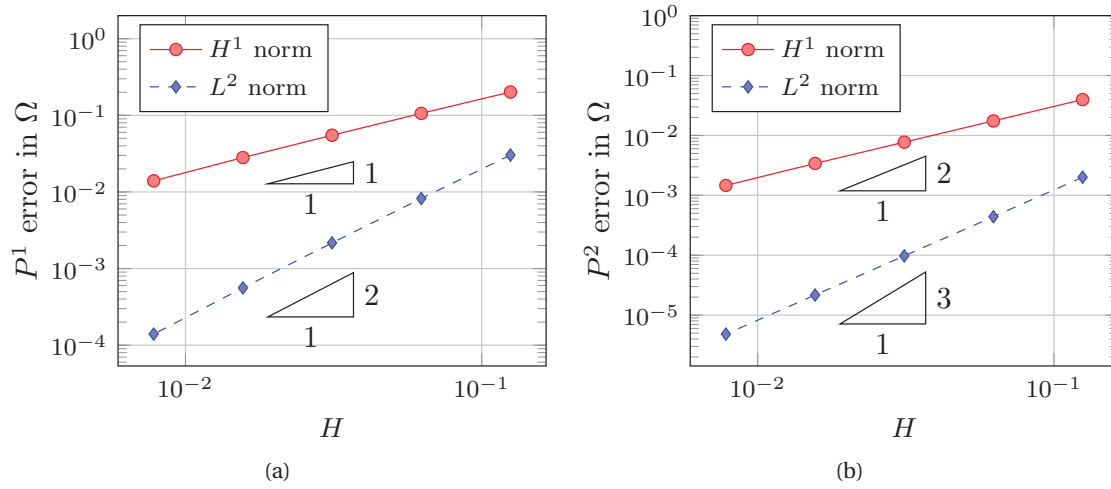


Figure 3.1 – H^1 (bullets, full) and L^2 (diamonds, dashed) errors between u^0 (the solution of (3.9)) and u^H (the solution of (3.14)) in Ω for (a) P^1 macro and micro FE spaces and (b) P^2 macro and micro FE spaces.

Experiment 1. We start by showing that the macro convergence rates in H are sharp. Let $\varepsilon = 1/10$, and consider equation (3.5) in $\Omega = [0, 1]^2$ with homogeneous Dirichlet boundary condition, a right-hand side $f \equiv 1$. A reference solution for u^0 is computed on a very fine mesh obtained from the initial mesh by uniform refinement. We use periodic coupling with $\delta = \varepsilon$ in order to have zero modeling error. Further, the micro degrees of freedom are chosen such that the micro error can be neglected, and take $H = 1/8, 1/16, 1/32, 1/64$, and $1/128$. In Figure 3.1a, we monitor the H^1 and L^2 errors to the homogenized solution u^0 for the piecewise macro and micro FE-HMM. The solution u^0 is in $H^2(\Omega)$ and one can see the linear and quadratic rates for the piecewise H^1 and L^2 errors, respectively. However, as one can see in Figure 3.1b, u^0 is not smooth enough to observe the H^2 and H^3 convergence rates for the quadratic H^1 and L^2 norms, respectively. The optimal rates can be seen in Figure 3.2 where we restrict the errors to a subdomain $\omega \subset \Omega$ to avoid corner singularities.

3.4. Numerical results for the FE-HMM applied to linear elasticity problems

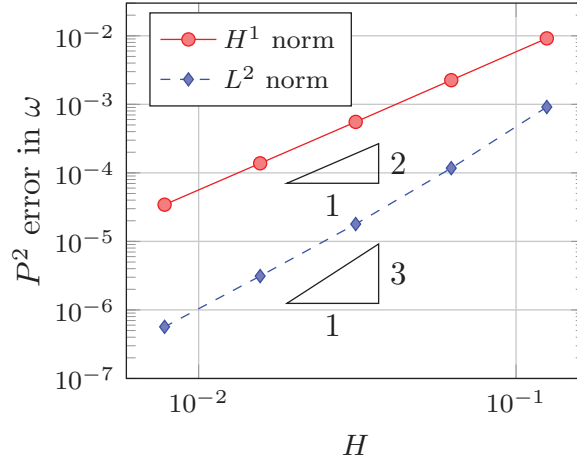


Figure 3.2 – H^1 (bullets, full) and L^2 (diamonds, dashed) errors between u^0 (the solution of (3.9)) and u^H (the solution of (3.14)) in $\omega \subset \Omega$ for P^2 macro and micro FE.

Experiment 2. Consider now problem (3.5) with $f \equiv 1$, on an L -shaped domain centered around $(0,0)$ with width 2. We impose free Neumann boundary conditions on the sets $\{x = 0, y \in [-1, 0]\}$ and $\{y = 0, x \in [0, 1]\}$, and homogeneous Dirichlet boundary conditions elsewhere.

In Figure 3.3a one can see the reference displacement in comparison to the initial coarse mesh with $H = 1/4$. Using periodic coupling and $\delta = \varepsilon$, we compute the FE-HMM solutions for P^1 macro and micro FE and for P^2 macro and micro FE; they are shown in Figures 3.3b and 3.3c, respectively.

In Figures 3.4a and 3.4b, we plot the H^1 and L^2 convergence rates for P^1 macro and micro FE spaces. We take $H = 1/8, 1/16, \dots, 1/512$. The optimal refinement follows the ratio given in Table 3.1.

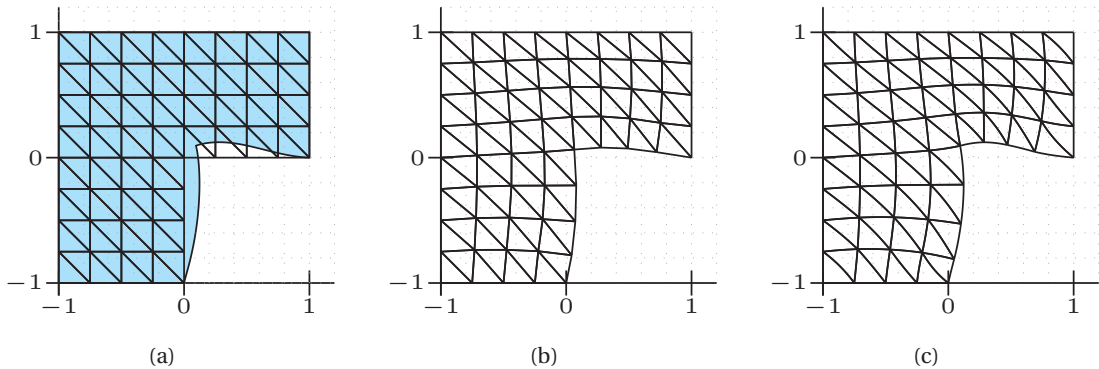


Figure 3.3 – (a) Reference solution. Finite element solution u^H for P^1 macro and micro FE (b), and P^2 macro and micro FE (c).

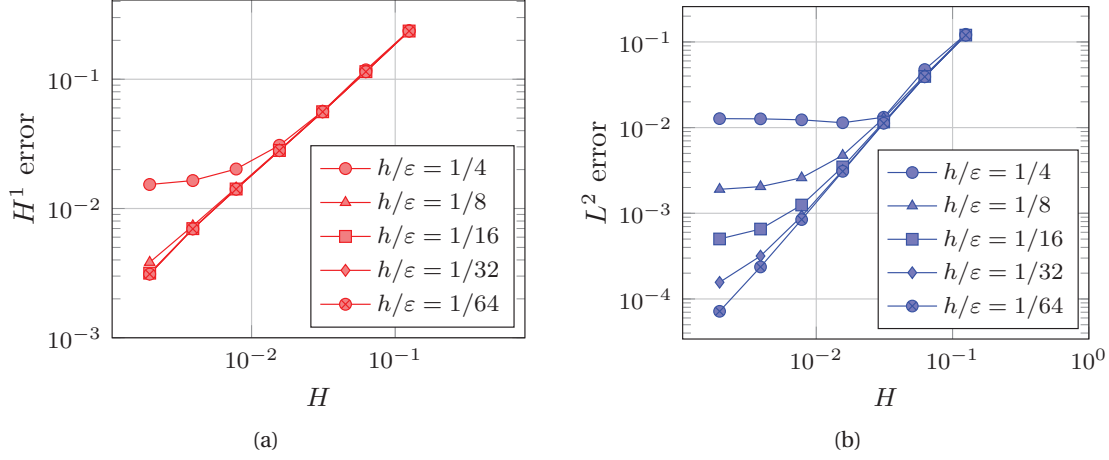


Figure 3.4 – H^1 (a) and L^2 (b) errors between u^0 and u^H for piecewise macro and micro FE spaces.

We show next the influence of the modeling error on the same problem with sampling domains K_δ with $\delta > \varepsilon$. We take $H = 1/8, 1/16, 1/32$, and $1/64$, with micro mesh size sufficiently small to eliminate the micro error. We use piecewise FE for the macro and micro mesh size. The size of the sampling domains K_δ are $\delta = 5/3\varepsilon$ and $\delta = 1.1\varepsilon$, and for those values we solve the micro problems (3.16) with homogeneous Dirichlet boundary conditions. In Figures 3.5a and 3.5b, we see that the choice of δ has an important influence on the error. Increasing the size of the sampling domain from $\delta = 1.1\varepsilon$ to $\delta = 5/3\varepsilon$ improves the quality of the error, as expected from Theorem 3.3.3. The periodic coupling with $\delta = \varepsilon$ gives the optimal convergence rate since the modeling error is zero, as predicted by Theorem 3.3.3.

Modeling error and random coefficients. The use of artificial boundary conditions for the micro problem (3.16) leads to a modeling (or resonance) error of size $\mathcal{O}(\varepsilon/\delta)$ for elliptic problems. Such error terms also appear for problems with random stationary fields, where (3.16) is usually defined in the whole \mathbb{R}^d [100]. Truncations using either Dirichlet or periodic boundary conditions can then be used for numerical approximation. In [70], a reduction of this resonance error is obtained by adding a zero-order term the cell problem (3.16) and using a suitable Richardson extrapolation of the modified cell problem. Such strategies could also be of interest for elastic problems.

Finally, we study the bound in Lemma 3.3.4. We use piecewise FE for the macro problem and compare the exact homogenized tensor with the numerical homogenized tensor. In Figure 3.6, we show the convergence rate

$$|a_{1111}^0 - a_{1111}^{0,h}| = |\sqrt{3} - a_{1111}^{0,h}|,$$

for piecewise (full) and quadratic (dashed) micro FE, and observe the expected rates.

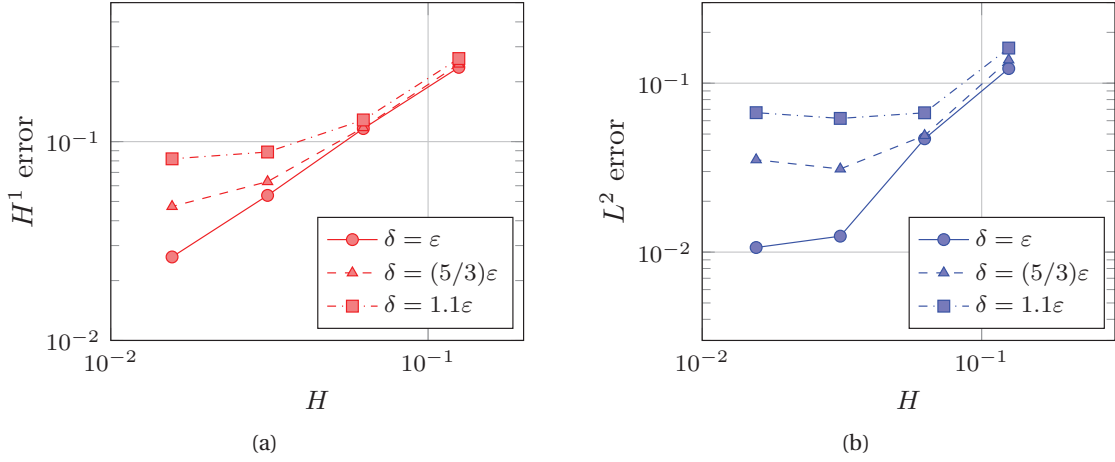


Figure 3.5 – H^1 error (a) and L^2 error (b) between the homogenized solution and the FE-HMM with Dirichlet coupling for $\delta = 5/3\epsilon$ (dashed) and $\delta = 1.1\epsilon$ (dash-dotted). The error $\delta = \epsilon$ (full) is obtained with periodic coupling.

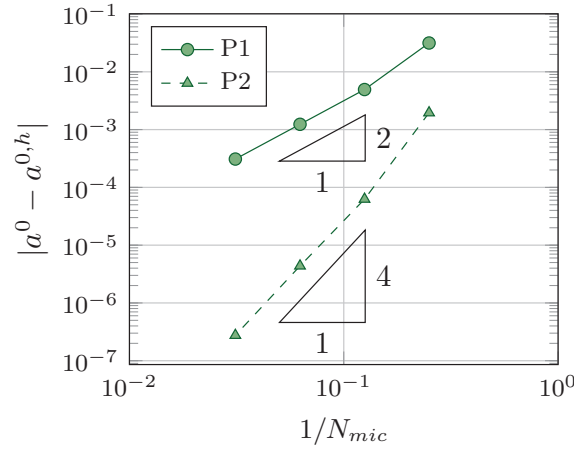


Figure 3.6 – Convergence rates $|a^0 - a_K^{0,h}|$ with respect to N_{mic}^{-1} for P^1 (full) and P^2 (dashed) micro FE spaces.

3.5 Summary

The finite element heterogeneous multiscale method applied to linear elastic problems is given here and established the basis for the wave equation in a linear elastic medium treated in the second part of Chapter 3. A fully discrete a priori error analysis is given and the sharpness of the error bounds are verified through numerical experiments.

Multiscale method for the wave equation in linear elastic heterogeneous media

In this part, we are concerned with the wave equation in a heterogeneous multiscale linear elastic medium and study the asymptotic behavior of the displacement u^ε when ε tends to zero.

Outline. In Section 3.6 we give the FE-HMM applied to the wave equation. In Section 3.7 we give the fully discrete a priori error analysis and, in Section 3.8, give numerical examples to verify the error estimates.

We consider the wave equation for short time $T > 0$, and seek the heterogeneous solution $u^\varepsilon : [0, T] \rightarrow H_0^1(\Omega)^d$ of

$$\begin{aligned} \partial_{tt} u^\varepsilon(t) - \operatorname{div}(a^\varepsilon(x) : e(u^\varepsilon(t))) &= f(t), & \text{in } \Omega \times (0, T] \\ u^\varepsilon(t) &= 0, & \text{on } \Gamma \times [0, T], \end{aligned} \quad (3.27)$$

with initial conditions at the time $t = 0$,

$$u^\varepsilon(0) = g_1, \quad \partial_t u^\varepsilon(0) = g_2, \quad \text{in } \Omega. \quad (3.28)$$

It holds $u^\varepsilon = (u_1^\varepsilon, \dots, u_d^\varepsilon)$, $f = (f_1, \dots, f_d)$, and $g_i = (g_{i_1}, \dots, g_{i_d})$ for $i = 1, 2$. We will sometimes drop the dependency in time and use u^ε to denote $u^\varepsilon(t)$.

In weak formulation, problem (3.27) reads: find $u^\varepsilon : [0, T] \rightarrow H_0^1(\Omega)^d$ such that

$$\langle \partial_{tt} u^\varepsilon(t), w \rangle + B^\varepsilon(u^\varepsilon(t), w) = F(w), \quad \forall w \in H_0^1(\Omega)^d, \quad (3.29)$$

with the initial conditions (3.28), and where $\langle \cdot, \cdot \rangle$ denotes $\langle \cdot, \cdot \rangle_{H^{-1}, H^1}$; if we assume that $\partial_{tt} u^\varepsilon(t)$ is in $L^2(\Omega)^d$, then we can use the standard L^2 inner product, which is denoted by (\cdot, \cdot) . The bilinear form $B^\varepsilon : H^1(\Omega)^d \times H^1(\Omega)^d \rightarrow \mathbb{R}$ is given by

$$B^\varepsilon(v, w) = \int_{\Omega} a^\varepsilon(x) e(v) : e(w) dx,$$

Chapter 3. Multiscale method for the wave equation in linear elastic heterogeneous media

and the right hand side $F : H^1(\Omega)^d \rightarrow \mathbb{R}$

$$F(w) = \int_{\Omega} f w dx.$$

The heterogeneous tensor a^ε is a fourth-order tensor with $a_{ijkl}^\varepsilon(x) \in L^\infty(\Omega)$, for $i, j, k, l = 1, \dots, d$ and verifying assumptions (3.2), (3.3), and (3.4). Using Korn's inequality (3.7), the bilinear form B^ε is symmetric, uniformly elliptic, and bounded in $H_0^1(\Omega)^d$. Further if we assume sufficient regularity on the data; i.e.,

$$f \in L^2(0, T; L^2(\Omega)^d), \quad g_1 \in H_0^1(\Omega)^d, \quad \text{and} \quad g_2 \in L^2(\Omega)^d,$$

we can prove that the weak formulation is well-posed; it holds that the wave equation (3.27) (or in weak form (3.29)) has a unique (weak) solution u^ε with

$$u^\varepsilon \in L^2(0, T; H_0^1(\Omega)^d) \quad \text{and} \quad \partial_t u^\varepsilon \in L^2(0, T; L^2(\Omega)^d).$$

The solutions u^ε are in fact more regular (see [86]) because $u^\varepsilon \in L^\infty(0, T; H_0^1(\Omega)^d)$ with time derivative $\partial_t u^\varepsilon \in L^\infty(0, T; L^2(\Omega)^d)$, and even,

$$u^\varepsilon \in \mathcal{C}([0, T]; H_0^1(\Omega)^d) \quad \text{and} \quad \partial_t u^\varepsilon \in \mathcal{C}([0, T]; L^2(\Omega)^d).$$

Homogenization of the wave equation

The effective dynamics at the macro scale can be approximated using homogenization theory [30, 96, 44]. By using the theory of H -convergence [90, 22], one can show that the effective behavior of the heterogeneous solution u^ε is well-described by the solution of the homogenized wave equation for short times $T > 0$ [30, 44]; i.e., u^ε converges weakly in H^1 to u^0 the solution of

$$\langle \partial_{tt} u^0(t), w \rangle + B^0(u^0(t), w) = F(w), \quad \forall w \in H_0^1(\Omega)^d. \quad (3.30)$$

with the initial conditions (3.28), and where $B^0 : H^1(\Omega)^d \times H^1(\Omega)^d \rightarrow \mathbb{R}$ is given by

$$B^0(v, w) = \int_{\Omega} a^0(x) e(v) : e(w) dx,$$

where the homogenized tensor a^0 verifies (3.2), (3.3), and (3.4) for some constants $0 < \lambda^0 \leq \Lambda^0 < \infty$.

When the time is increased, e.g. for time $T^\varepsilon = \varepsilon^{-2} T$, the heterogeneous solution deviates from the global behavior set from the homogenized equation and develops a dispersive behavior due to an interplay between the small scales. To capture this dispersive effect, the homogenized model requires some additional terms leading to a family of effective Boussinesq-type equations. We refer to [10, 19] for the treatment of the wave equation in heterogeneous media for long time with the FE-HMM, and focus on the propagation of a linear elastic wave propagation for short time.

We mention as well the multiscale method for the long time wave propagation in a heterogeneous medium introduced in [59, 60], where finite differences are used instead of finite elements, and where the authors consider dynamical micro problems.

3.6 FE-HMM for the wave equation in a linear elastic medium

The FE-HMM for the wave equation was introduced in [9] and later in [10] for long time. Following the macro to micro HMM approach (see Section 2.4), we consider a family of coarse partitions $\{\mathcal{T}_H\}$ over Ω of mesh size $H \gg \varepsilon$. Let \hat{K} be a reference element and $\{\hat{\omega}_j, \hat{x}_j\}$ be a quadrature formula on \hat{K} . We make the following assumptions; for $j = 1, \dots, J$, with $J \geq 1$, let $\hat{\omega}_j > 0$ and

$$(Q1) \quad \text{there exists } \lambda > 0 : \sum_{j=1}^J \hat{\omega}_j |\nabla \hat{p}(\hat{x}_j)|^2 \geq \lambda \|\nabla \hat{p}\|_{L^2(\hat{K})}^2, \text{ for all } \hat{p} \in \mathcal{R}^p(\hat{K})^d,$$

$$(Q2) \quad \int_{\hat{K}} \hat{p}(\hat{x}) d\hat{x} = \sum_{j=1}^J \hat{\omega}_j \hat{p}(\hat{x}_j), \text{ for all } \hat{p} \in \mathcal{R}^\sigma(\hat{K})^d, \text{ where } \sigma = \max(2p-2, p) \text{ if } \mathcal{R}^\sigma = \mathcal{P}^\sigma, \text{ and } \sigma = \max(2p-1, p+1).$$

Further, for quadrature formula used in the discrete form of the product (\cdot, \cdot) we assume

$$(Q3) \quad \sum_{j=1}^J \hat{\omega}_j |\hat{p}(\hat{x}_j)|^2 \geq \lambda \|\hat{p}\|_{L^2(\hat{K})}^2, \text{ for all } \hat{p} \in \mathcal{R}^p(\hat{K})^d.$$

Inside each element K , we define two different QF $\{\omega_{j,K}, x_{j,K}\}$ and $\{\tilde{\omega}_{l,K}, \tilde{x}_{l,K}\}$ with $j = 1, \dots, J$, $l = 1, \dots, L$, to evaluate the bilinear form B_H , and the discrete inner product $(\cdot, \cdot)_H$, respectively. We assume that both QF satisfy assumptions (Q1) and (Q2), and further that (Q3) holds for the QF used in the inner product $(\cdot, \cdot)_H$.

Around each quadrature node $x_{j,K}$ we construct a sampling domain K_{δ_j} of size $\delta \ll H$ and consider a family of fine partitions $\{\mathcal{T}_h\}$ over K_{δ_j} , of mesh size $h \leq \varepsilon$. Let $V_0^p(\Omega, \mathcal{T}_H)$ be a macro FE over Ω and $V^q(K_{\delta_j}, \mathcal{T}_h)$ be a micro FE over K_{δ_j} defined as in equations (3.12) and (3.15), respectively. The FE-HMM reads: find $u^H \in V_0^p(\Omega, \mathcal{T}_H)$ the solution of

$$(\partial_{tt} u^H, w^H)_H + B_H(u^H, w^H) = F(w^H), \quad \forall w^H \in V_0^p(\Omega, \mathcal{T}_H), \quad (3.31)$$

with the initial conditions obtained by nodal interpolation of the data in (3.28), and where

$$\begin{aligned} (\partial_{tt} v^H, w^H)_H &= \sum_{K \in \mathcal{T}_H} \sum_{l=1}^L \tilde{\omega}_{l,K} \partial_{tt} v^H(\tilde{x}_{l,K}) w^H(\tilde{x}_{l,K}), \\ F(w^H) &= \int_{\Omega} f w^H dx. \end{aligned}$$

and the bilinear form $B_H : V_0^p(\Omega, \mathcal{T}_H) \times V_0^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$ is given by

$$B_H(u^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{j,K}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(w_j^h) dx,$$

where u_j^h (resp. w_j^h) is such that $u_j^h - u_{\text{lin},j}^H \in V^q(K_{\delta_j}, \mathcal{T}_h)$ and

$$\int_{K_{\delta_j}} a^\varepsilon(x) e(u_j^h) : e(z^h) dx = 0, \quad \forall z^h \in V^q(K_{\delta_j}, \mathcal{T}_h).$$

The term $u_{\text{lin},j}^H(x)$ corresponds to a linearization of u^H at the integration nodes $x_{j,K}$, i.e., $u_{\text{lin},j}^H(x) = u^H(x_{j,K}) + (x - x_{j,K}) e(u^H(x_{j,K}))$. Note that the micro solutions do not depend on time and that the micro problems are well-posed; this follows from the Lax–Milgram lemma together with the Korn's inequalities.

Following [10], there exist constants $C_1, C_2 > 0$ such that

$$C_1 \|v^H\|_{L^2(\Omega)} \leq \|v^H\|_H \leq C_2 \|v^H\|_{L^2(\Omega)}, \quad (3.32)$$

where $\|v^H\|_H^2 = (v^H, v^H)_H$.

The problem (3.31) is well-posed; i.e., it admits a unique solution $u^H \in L^\infty(0, T; V_0^p(\Omega, \mathcal{T}_H))$, for all ε, H, h , (see [10, 44]).

3.7 A priori error analysis

In this section, we give an a priori error analysis for the FE-HMM applied to the wave equation in a linear elastic medium; it follows [10, Section 4]. We consider an elliptic projection of the homogenized solution u^0 , that we denote by $\pi_H u^0$, with

$$B_H(\pi_H u^0, w^H) = B^0(u^0, w^H) + \langle \partial_{tt} u^0, w^H \rangle - (I_H \partial_{tt} u^0, w^H)_H, \quad (3.33)$$

where I_H is a nodal interpolant satisfying, for all integers m, k with $0 \leq m \leq 1$ and $2 \leq k \leq p+1$,

$$\|v - I_H v\|_{H^m(\Omega)} \leq C H^{k-m} \|v\|_{H^k(\Omega)}. \quad (3.34)$$

It holds $I_H u^0 \in V_0^p(\Omega, \mathcal{T}_H)$ and the projection $\pi_H u^0 \in V_0^p(\Omega, \mathcal{T}_H)$ is uniquely determined as the solution of a boundary value problem. Further, we can define the initial conditions of the FE-HMM problem (3.31) by

$$u^H(0) = I_H u^0(0) = I_H g_1, \quad \partial_t u^H(0) = I_H \partial_t u^0(0) = I_H g_2, \quad \text{in } \Omega.$$

The goal of the analysis is to find an upper bound for the error

$$\|\partial_t(u^0 - u^H)\|_{L^\infty(0, T; L^2(\Omega)^d)} + \|u^0 - u^H\|_{L^\infty(0, T; H^1(\Omega)^d)},$$

where u^0 and u^H are the solutions of (3.30) and (3.31) respectively. The key is to use a triangular inequality together with an estimate for the difference between $\pi_H \partial_t^k u^0$ and $\partial_t^k u^0$ for $k = 0, 1, 2$.

Assuming that the homogenized tensor $a_{ijkl}^0(x) \in W^{1,\infty}(\Omega)$, for $i, j, k, l = 1, \dots, d$, we can consider the FEM applied to the homogenized problem (3.30) and obtain the bilinear form $B_H^0 : V^p(\Omega, \mathcal{T}_H) \times V^p(\Omega, \mathcal{T}_H) \rightarrow \mathbb{R}$

$$B_H^0(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} a^0(x_{j,K}) e(v^H(x_{j,K})) : e(w^H(x_{j,K})).$$

Assuming that the quadrature formula used for the bilinear form B_H^0 satisfies (Q1) and (Q2) and that the one used for $(\cdot, \cdot)_H$ satisfies (Q1), (Q2), and (Q3), the following estimates hold for $v^H, w^H \in V^p(\Omega, \mathcal{T}_H)$ and $a_{ijkl}^0(x) \in W^{m+p,\infty}(\Omega)$, for $i, j, k, l = 1, \dots, d$, with $m = 0, 1$, (see [41, 42])

$$|B^0(v^H, w^H) - B_H^0(v^H, w^H)| \leq CH^{p+m} \max_{i,j,k,l} \|a_{ijkl}^0\|_{W^{p+m,\infty}(\Omega)} \|v^H\|_{\tilde{H}^{p+m}(\Omega)} \|w^H\|_{\tilde{H}^{1+m}(\Omega)}, \quad (3.35)$$

$$|(v^H, w^H) - (v^H, w^H)_H| \leq CH^{p+m} \|v^H\|_{\tilde{H}^{p+m}(\Omega)} \|w^H\|_{\tilde{H}^{1+m}(\Omega)}, \quad (3.36)$$

where $\|\cdot\|_{\tilde{H}^p(\Omega)}$ is a broken norm. Then, we have a first error bound. The proof follows the lines of [10, Lemma 4.6].

Lemma 3.7.1. *Let u^0 be the solution of (3.30) and suppose that (3.35) and (3.36) hold for $m = 0$. Further assume that (3.32) holds and*

$$\begin{aligned} \partial_t^k u^0 &\in L^2(0, T; H^{p+1}(\Omega)^d), & k = 0, 1, 2, \\ \partial_t^{2+k} u^0 &\in L^2(0, T; H^p(\Omega)^d), & k = 0, 1, 2, \\ a_{ijlm}^0 &\in W^{p,\infty}(\Omega), & i, j, l, m = 1, \dots, d. \end{aligned}$$

Then

$$\|\partial_t^k u^0 - \pi_H \partial_t^k u^0\|_{L^2(0,T;H^1(\Omega)^d)} \leq C(H^p + e_{\text{MIC}} + e_{\text{MOD}}),$$

where the constant C is independent of H, h , and ε . The micro error e_{MIC} is given by (3.22) and the modeling error e_{MOD} is given by (3.24) or (3.25).

Proof. We give the proof for $k = 0$. Using equation (3.33) and the linearity of the forms B_0 and B_H , we write

$$\begin{aligned} B_H(\pi_H u^0 - I_H u^0, v^H) &= B_0(u^0 - I_H u^0, v^H) + B_0(I_H u^0, v^H) - B_H(I_H u^0, v^H) + (\partial_{tt} u^0, v^H) \\ &\quad - (I_H \partial_{tt} u^0, v^H)_H + (I_H \partial_{tt} u^0, v^H) - (I_H \partial_{tt} u^0, v^H) \\ &= B^0(u^0 - I_H u^0, v^H) + B^0(I_H u^0, v^H) - B_H^0(I_H u^0, v^H) \\ &\quad + B_H^0(I_H u^0, v^H) - B_H(I_H u^0, v^H) \\ &\quad + (\partial_{tt} u^0 - I_H \partial_{tt} u^0, v^H) + (I_H \partial_{tt} u^0, v^H) - (I_H \partial_{tt} u^0, v^H)_H. \end{aligned} \quad (3.37)$$

We bound each term of the last equation, and use the short-hand notation $\|\cdot\|_{L^2(H^p)}$ to denote

Chapter 3. Multiscale method for the wave equation in linear elastic heterogeneous media

the norm $\|\cdot\|_{L^2(0,T;H^p(\Omega))}$, $p \geq 1$. Using the boundedness of B^0 , it holds

$$B^0(u^0 - I_H u^0, v^H) \leq CH^p \|u^0\|_{L^2(H^{p+1})} \|v^H\|_{L^2(H^1)},$$

where we use that I_H verifies (3.34). For the second and third terms of (3.37), we use equation (3.35), with $m = 0$, and obtain

$$B^0(I_H u^0, v^H) - B_H^0(I_H u^0, v^H) \leq CH^p \|I_H u^0\|_{L^2(\tilde{H}^p)} \|v^H\|_{L^2(H^1)} \leq CH^p \|u^0\|_{L^2(H^{p+1})} \|v^H\|_{L^2(H^1)}.$$

Then, using [2, Lemma 4.3] and [10, Lemma 4.1], it holds

$$\begin{aligned} B_H^0(I_H u^0, v^H) - B_H(I_H u^0, v^H) &\leq C(e_{\text{MIC}} + e_{\text{MOD}}) \|I_H u^0\|_{L^2(H^1)} \|v^H\|_{L^2(H^1)} \\ &\leq C(e_{\text{MIC}} + e_{\text{MOD}}) \|u^0\|_{L^2(H^1)} \|v^H\|_{L^2(H^1)}. \end{aligned}$$

We bound the first inner product of (3.37) by

$$(\partial_{tt} u^0 - I_H \partial_{tt} u^0, v^H) \leq CH^p \|\partial_{tt} u^0\|_{L^2(H^p)} \|v^H\|_{L^2(H^1)},$$

where we use equation (3.34). The last term of (3.37) is bounded by

$$(I_H \partial_{tt} u^0, v^H) - (I_H \partial_{tt} u^0, v^H)_H \leq CH^p \|u^0\|_{L^2(H^p)} \|v^H\|_{L^2(H^1)}.$$

Using the coercivity of B_H ,

$$\begin{aligned} \|\pi_H u^0 - I_H u^0\|_{L^2(H^1)}^2 &\leq CB_H(\pi_H u^0 - I_H u^0, \pi_H u^0 - I_H u^0) \\ &\leq C(H^p \|u^0\|_{L^2(H^p)} + (e_{\text{MIC}} + e_{\text{MOD}}) \|u^0\|_{L^2(H^1)} \\ &\quad + H^p \|\partial_{tt} u^0\|_{L^2(H^p)}) \|\pi_H u^0 - I_H u^0\|_{L^2(H^1)}. \end{aligned}$$

We can conclude with an integration, a triangle inequality, and equation (3.34). \square

A similar error bound holds for the L^2 norm.

Lemma 3.7.2. *Let u^0 be the solution of (3.30) and suppose that (3.35) and (3.36) hold for $m = 1$. Further assume that (3.32) holds and*

$$\begin{aligned} \partial_t^k u^0 &\in L^2(0, T; H^{p+1}(\Omega)^d), \quad k = 0, 1, \\ \partial_t^{2+k} u^0 &\in L^2(0, T; H^p(\Omega)^d), \quad k = 0, 1, \\ a_{ijlm}^0 &\in W^{p+1,\infty}(\Omega), \quad i, j, l, m = 1, \dots, d. \end{aligned}$$

Then

$$\|\partial_t^k u^0 - \pi_H \partial_t^k u^0\|_{L^2(0,T;L^2(\Omega)^d)} \leq C(H^{p+1} + e_{\text{MIC}} + e_{\text{MOD}}),$$

where the constant C is independent of H , h , and ε . The micro error e_{MIC} is given by (3.22) and the modeling error e_{MOD} is given by (3.24) or (3.25).

We can now express a bound for the difference between u^0 and u^H .

Theorem 3.7.3. *Let u^0 and u^H be the solution of (3.30) and (3.31) respectively. Suppose that (3.35) and (3.36) hold for $m = 0$. Further assume that (3.32) and (3.34) hold, and that*

$$\begin{aligned} \partial_t^k u^0 &\in L^2(0, T; H^{p+1}(\Omega)^d), & k = 0, 1, 2, \\ \partial_t^{k+2} u^0 &\in L^2(0, T; H^p(\Omega)^d), & k = 0, 1, 2, \\ a_{ijlm}^0 &\in W^{p,\infty}(\Omega), & i, j, l, m = 1, \dots, d, \\ g_1 &\in H^{p+1}(\Omega)^d, \quad g_2 \in H^{\max(2,p)}(\Omega)^d, \\ \partial_t^k u^H &\in L^2(0, T; H^1(\Omega)^d), & k = 0, 1, 2. \end{aligned}$$

Then

$$\|\partial_t(u^0 - u^H)\|_{L^\infty(0,T;L^2(\Omega)^d)} + \|u^0 - u^H\|_{L^\infty(0,T;H^1(\Omega)^d)} \leq C(H^p + e_{\text{MIC}} + e_{\text{MOD}}),$$

where the micro error e_{MIC} is given by (3.22) and the modeling error e_{MOD} is given by (3.24) or (3.25).

Proof. We use the short-hand notation $\|\cdot\|_{L^2(H^p)}$ to denote the norm $\|\cdot\|_{L^2(0,T;H^p(\Omega))}$, $p \geq 1$. We decompose the H^1 error of $u^0 - u^H$ into

$$\|u^0 - u^H\|_{L^2(H^1)} \leq \|u^0 - \pi_H u^0\|_{L^2(H^1)} + \|\pi_H u^0 - u^H\|_{L^2(H^1)}.$$

The first error is bounded by Lemma 3.7.1, thus

$$\|u^0 - \pi_H u^0\|_{H^1(\Omega)} \leq C(H^p + e_{\text{MIC}} + e_{\text{MOD}}).$$

We write

$$\begin{aligned} &(\partial_{tt}(u^H - \pi_H u^0), v^H)_H + B_H(u^H - \pi_H u^0, v^H) \\ &= (\partial_{tt} u^H, v^H)_H - (\partial_{tt} \pi_H u^0, v^H)_H + B_H(u^H, v^H) - B_H(\pi_H u^0, v^H) \\ &= F(v^H) - (\partial_{tt} \pi_H u^0, v^H)_H - B_H(\pi_H u^0, v^H) \\ &= (\partial_{tt} u^0, v^H)_H + B^0(u^0, v^H) - (\partial_{tt} \pi_H u^0, v^H)_H - B_H(\pi_H u^0, v^H) \\ &= (\partial_{tt} u^0, v^H)_H + B^0(u^0, v^H) - (\partial_{tt} \pi_H u^0, v^H)_H - B^0(u^0, v^H) - (\partial_{tt} u^0, v^H)_H + (I_H \partial_{tt} u^0, v^H)_H \\ &= (I_H \partial_{tt} u^0 - \partial_{tt} \pi_H u^0, v^H)_H, \end{aligned}$$

We then follow [10], and obtain

$$\frac{1}{2} \frac{d}{dt} ((\partial_t \eta_H, \partial_t \eta_H)_H + B_H(\eta_H, \eta_H)) = (I_H \partial_{tt} u^0 - \pi_H \partial_{tt} u^0, \partial_t \eta_H)_H,$$

where $\eta_H = (u^0 - \pi_H u^0)$. We then call $\xi(t) = (\partial_t \eta_H, \partial_t \eta_H)_H + B_H(\eta_H, \eta_H)$, and using Young's

Chapter 3. Multiscale method for the wave equation in linear elastic heterogeneous media

inequality, it holds

$$\frac{1}{2} \frac{d}{dt} \xi(t) \leq C(\|I_H \partial_{tt} u^0 - \pi_H \partial_{tt} u^0\|_{L^2(L^2)}^2 + \|\partial_t \eta_H\|_{L^2(L^2)}^2).$$

We can bound

$$\|\partial_t \eta_H\|_{L^2(L^2)}^2 \leq (\partial_t \eta_H, \partial_t \eta_H)_H \leq (\partial_t \eta_H, \partial_t \eta_H)_H + B_H(\eta_H, \eta_H) = \xi(t).$$

Then, it holds

$$\frac{1}{2} \frac{d}{dt} \xi(t) \leq C(\|I_H \partial_{tt} u^0 - \pi_H \partial_{tt} u^0\|_{L^2(L^2)}^2 + \xi(t)),$$

and, from Gronwall's inequality, we obtain a bound on $\sup \xi(t)$

$$\sup_{0 \leq t \leq T} \xi(t) \leq C(\xi(0) + \|I_H \partial_{tt} u^0 - \pi_H \partial_{tt} u^0\|_{L^2(L^2)}^2),$$

and using Lemma 3.7.2, it holds

$$\sup_{0 \leq t \leq T} \xi(t) \leq C(\xi(0) + H^{2(p+1)} + (e_{\text{MIC}} + e_{\text{MOD}})^2).$$

It remains to bound $\xi(0)$, by definition,

$$\xi(0) = ((\partial_t \eta_H, \partial_t \eta_H)_H + B_H(\eta_H, \eta_H))|_{t=0} \leq C(\|\partial_t \eta_H(0)\|_{L^2(\Omega)}^2 + B_H(\eta_H(0), \eta_H(0)).$$

We then obtain

$$\begin{aligned} |B_H(\eta_H(0), \eta_H(0))| &\leq C\|\eta_H(0)\|_{H^1(\Omega)}^2 = C\|u^H(0) - \pi_H u^0(0)\|_{H^1(\Omega)}^2 \\ &= C\|I_H g_1 - \pi_H u^0(0)\|_{H^1(\Omega)}^2 \\ &\leq C(\|I_H g_1 - g_1\|_{H^1(\Omega)}^2 + \|u^0(0) - \pi_H u^0(0)\|_{H^1(\Omega)}^2) \\ &\leq C(H^{2p}\|g_1\|_{H^{p+1}(\Omega)}^2 + \|u^0 - \pi_H u^0\|_{L^2(H^1)}^2 + \|\partial_t u^0 - \pi_H \partial_t u^0\|_{L^2(H^1)}^2) \\ &\leq C(H^{2p}\|g_1\|_{H^{p+1}(\Omega)}^2 + H^{2p} + (e_{\text{MIC}} + e_{\text{MOD}})^2), \end{aligned}$$

using the continuous embedding of $H^1(H^1)$ into $\mathcal{C}(H^1)$, Lemma 3.7.1 together with equation (3.34) with $m = 1, k = p + 1$, and assuming that $g_1 \in H^{p+1}(\Omega)$. Similarly,

$$\begin{aligned} \|\partial_t \eta_H(0)\|_{L^2(\Omega)} &= \|I_H g_2 - \pi_H \partial_t u^0(0)\|_{L^2(\Omega)} \\ &\leq C(\|I_H g_2 - g_2\|_{L^2(\Omega)} + \|\partial_t u^0(0) - \pi_H \partial_t u^0(0)\|_{L^2(\Omega)}) \\ &\leq C(H^p\|g_2\|_{H^p(\Omega)} + \|\partial_t u^0 - \pi_H \partial_t u^0\|_{L^2(L^2)} + \|\partial_{tt} u^0 - \pi_H \partial_{tt} u^0\|_{L^2(L^2)}) \\ &\leq C(H^p\|g_2\|_{H^p(\Omega)} + H^{p+1} + e_{\text{MIC}} + e_{\text{MOD}}). \end{aligned}$$

All together,

$$\sup_{0 \leq t \leq T} \xi(t) \leq C(H^{2p} + (e_{\text{MIC}} + e_{\text{MOD}})^2),$$

and

$$\|\partial_t \eta_H\|_{L^\infty(0,T;L^2(\Omega))}^2 + \|\eta_H\|_{L^\infty(0,T;H^1(\Omega))}^2 \leq C \sup_{0 \leq t \leq T} \xi(t) \leq C(H^{2p} + (e_{\text{MIC}} + e_{\text{MOD}})^2).$$

□

Following [10], we can also prove L^2 a priori error estimates. We state the result in the next theorem.

Theorem 3.7.4. *Let $u^0 \in H_0^1(\Omega)^d$ and u^H be the solutions of (3.30) and (3.31) respectively. Suppose that (3.35) and (3.36) hold for $m = 1$. Further assume that (3.32) and (3.34) hold, and that*

$$\begin{aligned} \partial_t^k u^0 &\in L^2(0, T; H^{p+1}(\Omega)^d), \quad k = 0, 1, 2, 3 \\ \partial_t^4 u^0 &\in L^2(0, T; H^p(\Omega)^d), \\ a_{ijlm}^0 &\in W^{p+1,\infty}(\Omega), \quad i, j, l, m = 1, \dots, d, \\ g_1 &\in H^{p+1}(\Omega)^d, \\ \partial_t^k u^H &\in L^2(0, T; H^1(\Omega)^d), \quad k = 0, 1. \end{aligned}$$

Then

$$\|u^0 - u^H\|_{L^\infty(0,T;L^2(\Omega)^d)} \leq C(H^{p+1} + e_{\text{MIC}} + e_{\text{MOD}}),$$

where the constant C is independent of H, h , and ε . The micro error e_{MIC} is given by (3.22) and the modeling error e_{MOD} is given by (3.24) or (3.25).

3.8 Numerical examples

In this section, we give numerical examples to test the FE-HMM applied to a wave propagation in a linear elastic medium for short time $T > 0$; we seek a numerical approximation of u^0 the solution of

$$\partial_{tt} u^0 - \text{div}(a^0(x) : e(u^0)) = f.$$

We recall the expected convergence rates

$$\begin{aligned} \|u^0(T) - u^H(T)\|_{H^1(\Omega)} &\leq C(H^p + e_{\text{MIC}} + e_{\text{MOD}}), \quad \text{and} \\ \|u^0(T) - u^H(T)\|_{L^2(\Omega)} &\leq C(H^{p+1} + e_{\text{MIC}} + e_{\text{MOD}}). \end{aligned}$$

Outline. The section is organized as follows. In 3.8.1, we verify the sharpness of the convergence rates of Theorem 3.7.3 and Theorem 3.7.4 through two experiments, one with a periodic tensor and one with a locally periodic tensor. Then, in 3.8.2, we consider a layered material and compare the FE-HMM with the homogenization method proposed by Schoenberg and Muir [104]. Finally, in 3.8.3 we construct an arbitrarily random layered medium, where the randomness is generated from a von-Karman correlation function, and test the FE-HMM.

3.8.1 Periodic and locally periodic tensor

In here we verify the sharpness of the error bounds of Theorem 3.7.3 and Theorem 3.7.4.

Experiment 1. For the first experiment, we choose a Y -periodic tensor. Let $\Omega = [-1, 1]^2$ and a^ε be given by

$$a^\varepsilon(x) = \begin{pmatrix} \sin(2\pi x_1/\varepsilon) + 2 & 0 & 0 \\ 0 & \sin(2\pi x_2/\varepsilon) + 2 & 0 \\ 0 & 0 & 10 \end{pmatrix}.$$

Explicit equations are available for the homogenized tensor a^0 , and one obtains

$$a^0 = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 10 \end{pmatrix}.$$

We choose Neumann boundary conditions, a Gaussian initial condition g_1 , and zero initial condition g_2 . In this example, the reference solution u^0 is computed on a fine mesh obtained from the initial mesh with uniform refinements. Collocation to the slow variable is used, setting the modeling error to zero. Further, we set the size of the sampling domain to $\delta = \varepsilon$.

We first show the convergence rate for piecewise linear macro and micro FE and for quadratic macro and micro FE. We chose a discretization in the micro FE such that the micro error can be neglected. We use a Leapfrog scheme and we impose a CFL condition for stability of

$$\Delta t \leq \frac{h_f}{50}, \quad (3.38)$$

where h_f is the mesh size of the fine mesh used to compute the reference solution. In order to neglect the error in time, we set the CFL condition to a small value. We chose $\varepsilon = 1/10$, $T = 0.2$, and choose an initial mesh size of $H = 1/12$. At first, we fix the micro mesh size and refine uniformly the macro mesh size. The CFL condition for the FE-HMM can be chosen much larger than the CFL condition (3.38) used for the computation of the reference solutions u^0 and u^ε . Indeed, the CFL condition for the FE-HMM depends on the macro mesh size H , e.g.,

$$(\Delta t)_{HMM} \leq \frac{H}{50}.$$

The H^1 and L^2 errors between the numerical and reference solutions are expected to reach a threshold value depending on the micro mesh size. In Figures 3.7a and 3.7b, one can see the H^1 and L^2 errors, respectively, at time $t = 0.2$, with piecewise macro and micro FE, with $H = 1/12$, and for the micro mesh sizes $h = 1/4, 1/8$, and $h = 1/16$. The macro and micro number of degrees of freedom can be chosen in order to obtain optimal convergence rates. Indeed, if $H \approx \frac{1}{N_{macro}}$ and $\frac{h}{\varepsilon} \approx \frac{1}{N_{micro}}$, we obtain, from the convergence rates given in Theorems

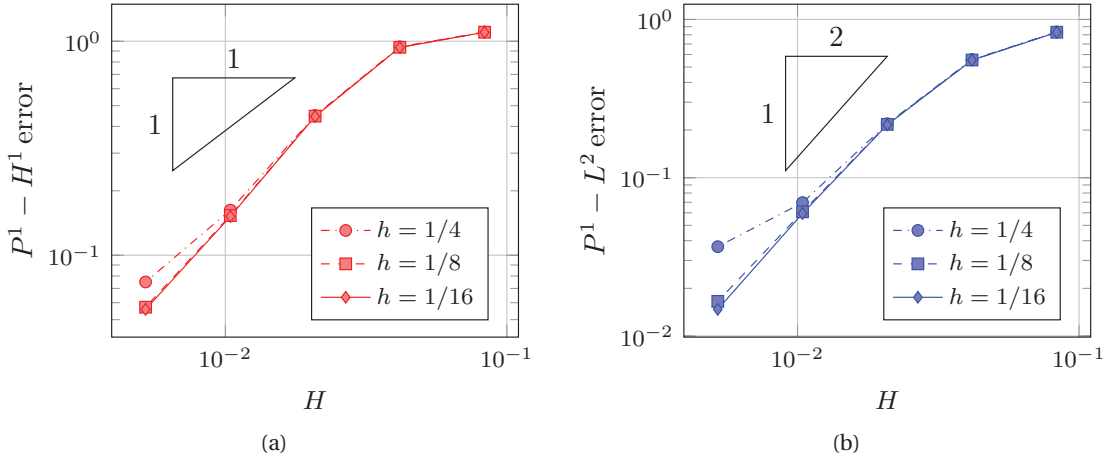


Figure 3.7 – Error between u^0 and u^H in Ω for (a) H^1 and (b) L^2 errors for P^1 macro and micro FE spaces for different micro mesh sizes.

3.7.3 and 3.7.4,

$$N_{micro} = N_{macro}^{1/2}, \quad (H^1 \text{ norm}), \quad N_{micro} = N_{macro}, \quad (L^2 \text{ norm}).$$

The effect of the micro error in the H^1 and L^2 norm can be better seen when P^2 macro FE are used with P^1 micro FE. In Figures 3.8a and 3.8b, we plot the H^1 and L^2 errors, respectively, at time $t = 0.2$, using P^2 macro and P^1 micro FE with $H = 1/6$, and micro mesh sizes $h = 1/4, 1/8, 1/16$, and $h = 1/32$. The optimal macro-micro refinements are given by

$$N_{micro} = N_{macro}, \quad (H^1 \text{ norm}), \quad N_{micro} = N_{macro}^{3/2}, \quad (L^2 \text{ norm}).$$

We plot horizontal snapshots at the depth $z = -0.5$ of the displacements in the x and z directions at the final time $T = 0.2$ second. In Figure 3.9a, we can see the amplitude of the displacement along the x direction of the homogenized solution in black and the heterogeneous solution for $\varepsilon = 1/50$ in red. The z displacement is represented in Figure 3.9b. When ε is made smaller, we see that the homogenized solution captures the global behavior of the heterogeneous solution; in Figures 3.10a and 3.10b we compare the two displacements along the x and z direction, respectively, for $\varepsilon = 1/100$. As the error between the two tensors a_K^0 and a^0 is small, the global displacement of the numerical homogenized solution u^H is similar to the displacement of the homogenized solution u^0 at the final time $T = 0.2$. Thus, we decided not to plot u^H in Figures 3.9 and 3.10.

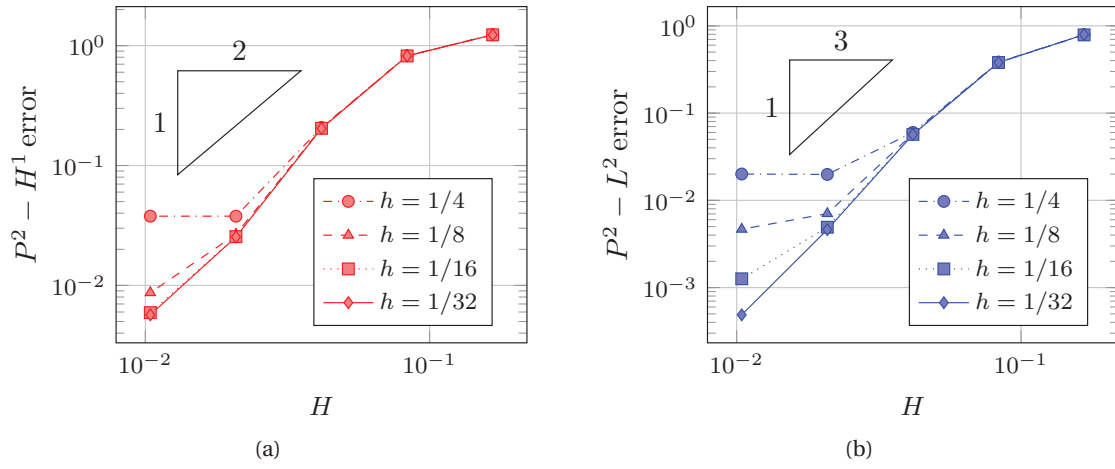


Figure 3.8 – Error between u^0 and u^H in Ω for (a) H^1 and (b) L^2 errors for P^2 macro and P^1 micro FE spaces for different micro mesh sizes.

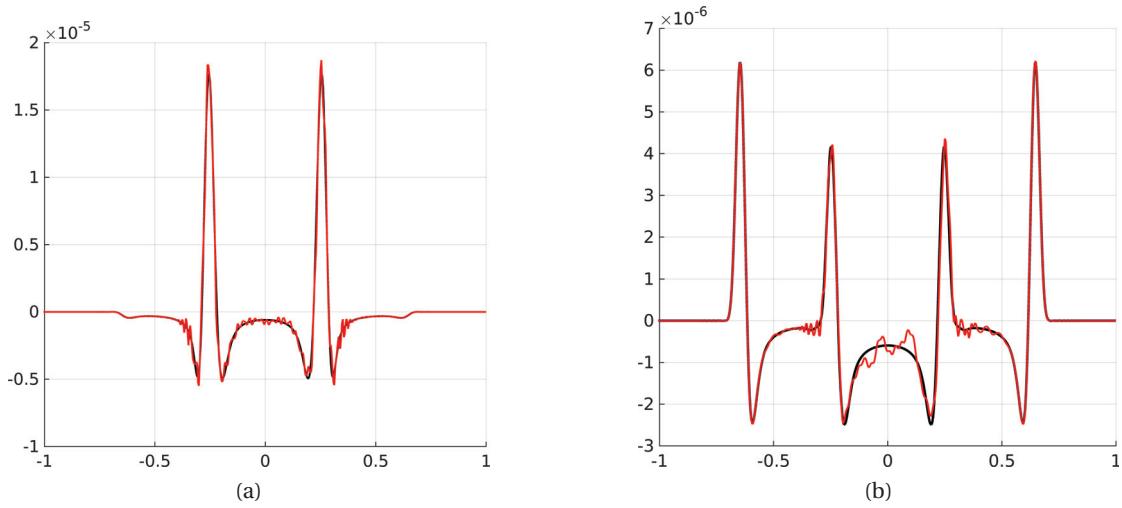


Figure 3.9 – Horizontal snapshots at the depth $z = -0.5$ of the displacements u^0 (in black) and u^ϵ (in red) in the x and z directions at the final time $T = 0.2$ second, with $\epsilon = 1/50$.

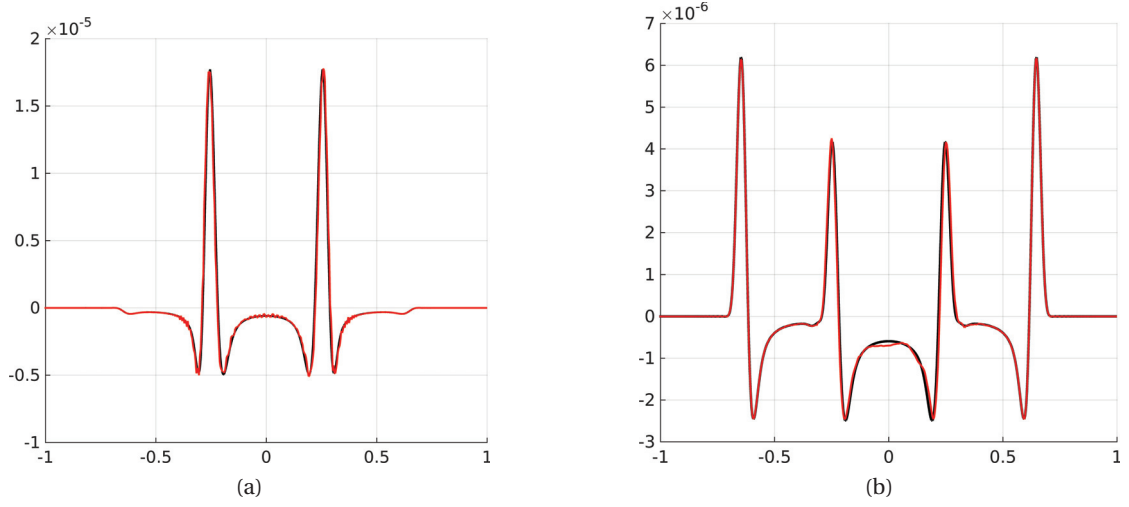


Figure 3.10 – Horizontal snapshots at the depth $z = -0.5$ of the displacements u^0 (in black) and u^ε (in red) in the x and z directions at the final time $T = 0.2$ second, with $\varepsilon = 1/100$.

Experiment 2. In this experiment, we take a locally periodic tensor a^ε . Let a^ε be defined by

$$a^\varepsilon(x) = \begin{pmatrix} \sin(2\pi x_1/\varepsilon) \sin(x_1^2 x_2^2) + 2 & 0 & 0 \\ 0 & \sin(2\pi x_2/\varepsilon) \sin(x_1^2 x_2^2) + 2 & 0 \\ 0 & 0 & 10 \end{pmatrix}.$$

In that case, the exact homogenized tensor is not known and the reference solution u^0 is computed with FE-HMM using a very fine mesh for the macro and micro problems. Collocation is still used, and we fix the micro mesh size to a very fine value to neglect the micro error. We set $T = 0.1$ and chose an initial mesh of size $H = 1/8$. In Figures 3.11a and 3.11b, we plot H^1 and L^2 errors at the final time T , using P^2 macro and P^1 micro FE with sampling domains of different size δ .

3.8.2 Horizontally layered material

Consider now an horizontally layered material split into horizontal isotropy and vertical isotropy, and where the tensor a^ε is given by

$$a^\varepsilon(x) = a_H(x) + a_V(x), \quad x \in \Omega.$$

The subscripts H and V stand for horizontal and vertical, respectively, and ε denotes the width of the layers. We assume that the domain $\Omega = [-2, 0]^2$, and that the distribution of the layers are horizontal-vertical.

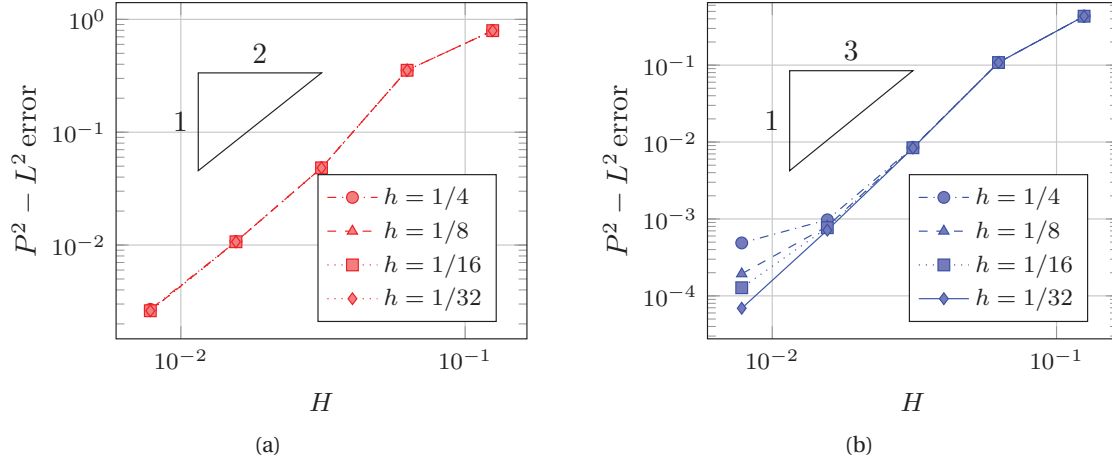


Figure 3.11 – (a) H^1 error and (b) L^2 error between u^0 and u^H in Ω for with a locally periodic tensor a^ϵ using P^2 macro and P^1 micro FE spaces.

We consider that the tensors a_H and a_V are given by

$$a_H = \begin{pmatrix} 46 & 18 & 0 \\ 18 & 30 & 0 \\ 0 & 0 & 7 \end{pmatrix}, \quad a_V = \begin{pmatrix} 30 & 18 & 0 \\ 18 & 46 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

For $\epsilon = 1/10$, the first component of a^ϵ is represented in Figure 3.12. We compute the homogenized tensor on a cell problem of size $\delta = 2\epsilon$ with a mesh size $h = \frac{1}{1024}$ to be in accordance with the numerical results of [62].

For horizontally layered anisotropic elastic media, an effective homogenized tensor can be derived using an averaging method proposed by Schoenberg and Muir [104]. Detailed equations for the computations of the homogenized tensor can be found in [37], and one obtain

$$a^0 = \begin{pmatrix} 38 & 18 & 0 \\ 18 & 36.30 & 0 \\ 0 & 0 & 7 \end{pmatrix},$$

and the numerical homogenized tensor $a_K^{0,h}$ can be computed during the assembly process of the FE-HMM using equation (3.26). As the medium is periodic, the value of the numerical homogenized tensor $a_K^{0,h}$, computed in the sampling domain K_{δ_j} , is the same at each quadrature node $x_{j,K}$. the numerical homogenized tensor $a_K^{0,h}$ obtained with the FE-HMM, using P^1 macro and micro FE, is

$$a_K^{0,h} = \begin{pmatrix} 38 & 18 & 0 \\ 18 & 36.3158 & 0 \\ 0 & 0 & 7 \end{pmatrix},$$

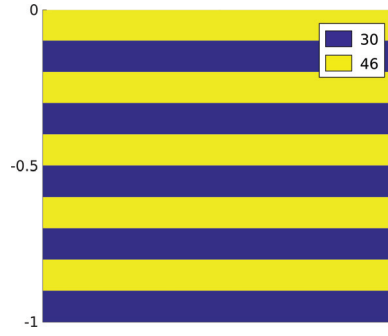


Figure 3.12 – Layered material component a_{1111}^ε for $\varepsilon = 1/10$.

and the error $e = |a_K^{0,h} - a^0|$ in percents is

$$e = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.04\% & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Using quadratic micro FE in the micro problem gives

$$a_K^{0,h} = \begin{pmatrix} 37.99 & 17.99 & 0 \\ 17.99 & 36.328 & 0 \\ 0 & 0 & 7 \end{pmatrix},$$

and an error $e = |a_K^{0,h} - a^0|$ in percents of

$$e = \begin{pmatrix} 0.026\% & 0.055\% & 0 \\ 0.055\% & 0.077\% & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In Figure 3.13a, we plot the reference solution u^ε at time $t = 0.1$ for $\varepsilon = 1/50$ and in Figure 3.13b, we plot the numerical homogenized solution. The global behavior of the two solutions is similar.

3.8.3 Arbitrarily heterogeneous media.

In this experiment, we consider a random layered medium generated by the von-Karman correlation function [72, 62, 93]

$$v_\kappa(x) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)} \left(\left\| \frac{x}{c} \right\| \right)^\kappa K_\kappa \left(\left\| \frac{x}{c} \right\| \right),$$

where κ is the hurst number, K_κ is a modified Bessel function of order κ , $c = (c_H, c_V)$ is the correlation distance of the heterogeneities in the medium; c_H (c_V) stands for the horizontal (vertical) size of the heterogeneities. The medium is obtained by filtering a white noise by a

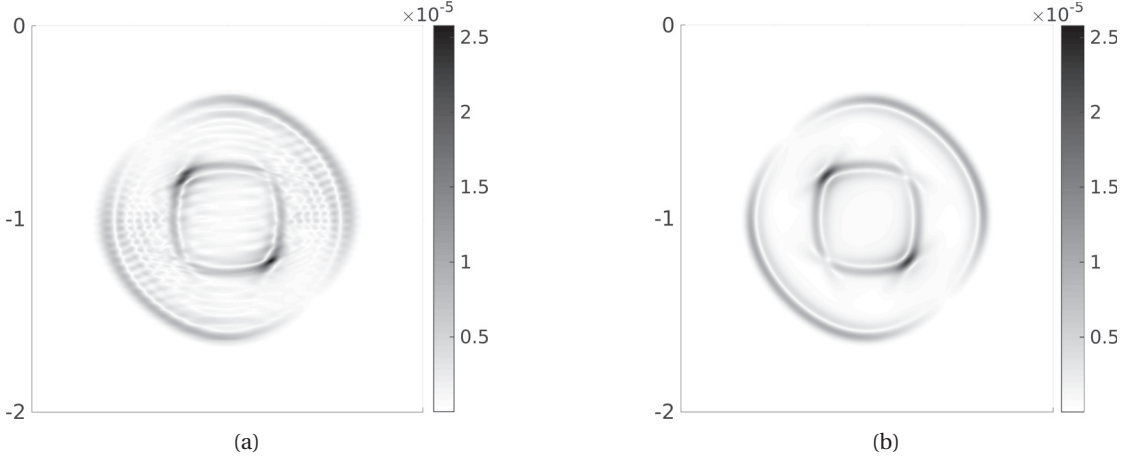


Figure 3.13 – Snapshots at time $t = 0.1$ second of (a) the reference solution u^ϵ , and (b) the homogenized solution for $\epsilon = 1/50$.

spectral filter, which is the square root of the power spectrum density function (the Fourier transform) of the von-Karman correlation function. We start by computing the von-Karman function on a fine grid over Ω , then we compute its Fourier transform and take the square root; this is the spectral filter. We take a uniform distribution between 0 and 1, compute its Fourier transform and multiply it by the filter. The random data are then obtained by going back to the spacial domain Ω .

Discussion in [81], leads us to take negative values of κ ; further, in two-dimensional experiments they give a lower bound $-\frac{3}{4} < \kappa$.

The tensor a^ϵ is given by summing the random von-Karman medium to a layered tensor with layer size of ϵ . The layered tensor a_L^ϵ is given by

$$a_L^\epsilon = \begin{pmatrix} a_{1111} & a_{1122} & 0 \\ a_{1122} & a_{2222} & 0 \\ 0 & 0 & a_{1212} \end{pmatrix},$$

where each entry is layered. Let K_κ be the first order modified Bessel function and set $\kappa = -0.2$, $c = (0.2, 0.01)$, and $\epsilon = 1/16$, the different components of the tensor a^ϵ are shown in Figures 3.14a to 3.14f.

As the tensor is four-layered, we take sampling domains of size $\delta = 4n\epsilon$, $n \in \mathbb{N}^*$ with a mesh size h small enough to capture the heterogeneities of the medium, i.e., set by the correlation distance c . The Schoenberg–Muir averaging method fails for layered media with arbitrary heterogeneities, and similarly we cannot use the exact formula for the homogenized tensor in a layered medium [80, 44]. Thus, no explicit equations are available for the homogenized tensor. However, from the theory of homogenization in random media, the homogenized tensor at a point $x \in \Omega$ reaches a stable value when δ increases. We take a micro number of

	$a_{1111}^{0,h}$	$a_{1122}^{0,h}$	$a_{1112}^{0,h}$	$a_{2222}^{0,h}$	$a_{2212}^{0,h}$	$a_{1212}^{0,h}$
$\delta = 8\epsilon$	13.5734	4.1497	0.6353	11.0198	1.4192	5.2596
$\delta = 16\epsilon$	13.5824	4.1541	0.6377	11.0261	1.4265	5.2665
$\delta = 24\epsilon$	13.5653	4.137	0.6277	11.0144	1.4028	5.2475
$\delta = 32\epsilon$	13.5822	4.1538	0.6328	11.0600	1.4137	5.2723
$\delta = 40\epsilon$	13.5759	4.1512	0.6354	11.0331	1.4216	5.2662
$\delta = 48\epsilon$	13.5751	4.1534	0.6353	11.0453	1.4218	5.2712

Table 3.2 – Components of the numerical homogenized tensor computed on sampling domains with increasing size δ .

degrees of freedom $N_{micro} = 1025$ and increase δ . We set $\epsilon = 1/50$, $c = (\epsilon/2, \epsilon/4)$ and take $n = 2, 3, \dots, 12$, leading to sampling domains of sizes $8\epsilon \leq \delta < 1$ — we omit $\delta = 4\epsilon$ as it gives results slightly off the grid. We take a quadrature point $x_{j,K}$, at the center of Ω , and compute the numerical homogenized tensor $a_K^{0,h}$ at the quadrature point $x_{j,K}$. With δ increasing, and by keeping a mesh size that fully resolve the fine scales, the numerical tensor tends to stabilize to

$$a_K^{0,h} = \begin{pmatrix} 13.575 & 4.153 & 0.635 \\ & 11.045 & 1.421 \\ & & 5.271 \end{pmatrix}, \quad (3.39)$$

where the blank entries can be filled by symmetry. Table 3.2 gives the values of the numerical homogenized tensor $a_K^{0,h}$ for different values of δ .

Take homogeneous Neumann boundary conditions and an initial condition given by a Gaussian pulse located at the center of Ω . We consider a CFL condition of

$$\Delta t \leq \frac{h_f}{30}.$$

Let u^H be the numerical homogenized solution computed with FE-HMM using P^2 macro FE and P^1 micro FE with $\delta = 8\epsilon$ with a number of micro degrees of freedom $N_{micro} = 128$, where the micro problems are solved around each macro quadrature points. Consider as well, a numerical solution $u^{0,H}$ computed with the tensor (3.39) obtained by taking a sampling domain of size $\delta = 48\epsilon$ with a number of micro degrees of freedom $N_{micro} = 1024$. In Figure 3.15, we plot snapshots at time $t = 0.025, 0.05, 0.075$, and 0.1 second of the numerical homogenized solutions u^H in the first column, the numerical solution $u^{0,H}$ in the second column, and the reference solution u^ϵ in the third column. We can see that the three solutions give similar behavior, albeit with oscillations for the reference solution. At last, in Figure 3.16, we plot the reference solution u^ϵ for $\epsilon = 1/100$ at time $t = 0.05$ and $t = 0.1$ second. When ϵ is made smaller we see that the heterogeneous and homogenized displacements are akin.

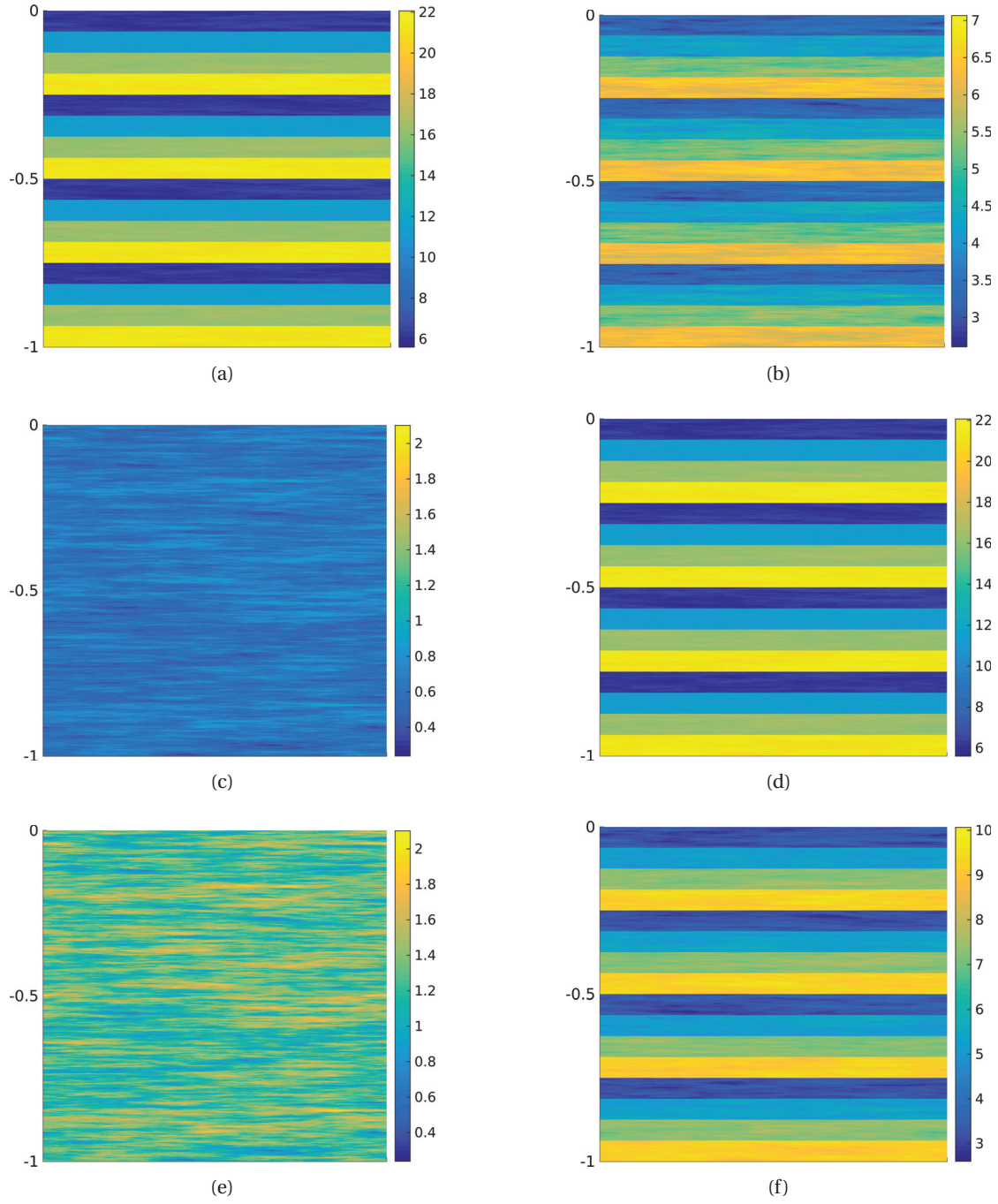


Figure 3.14 – Layered von-Karman random medium for $\varepsilon = 1/16$, (a) a_{1111}^ε , (b) a_{1122}^ε , (c) a_{1112}^ε , (d) a_{2222}^ε , (e) a_{2212}^ε , and (f) a_{1212}^ε .

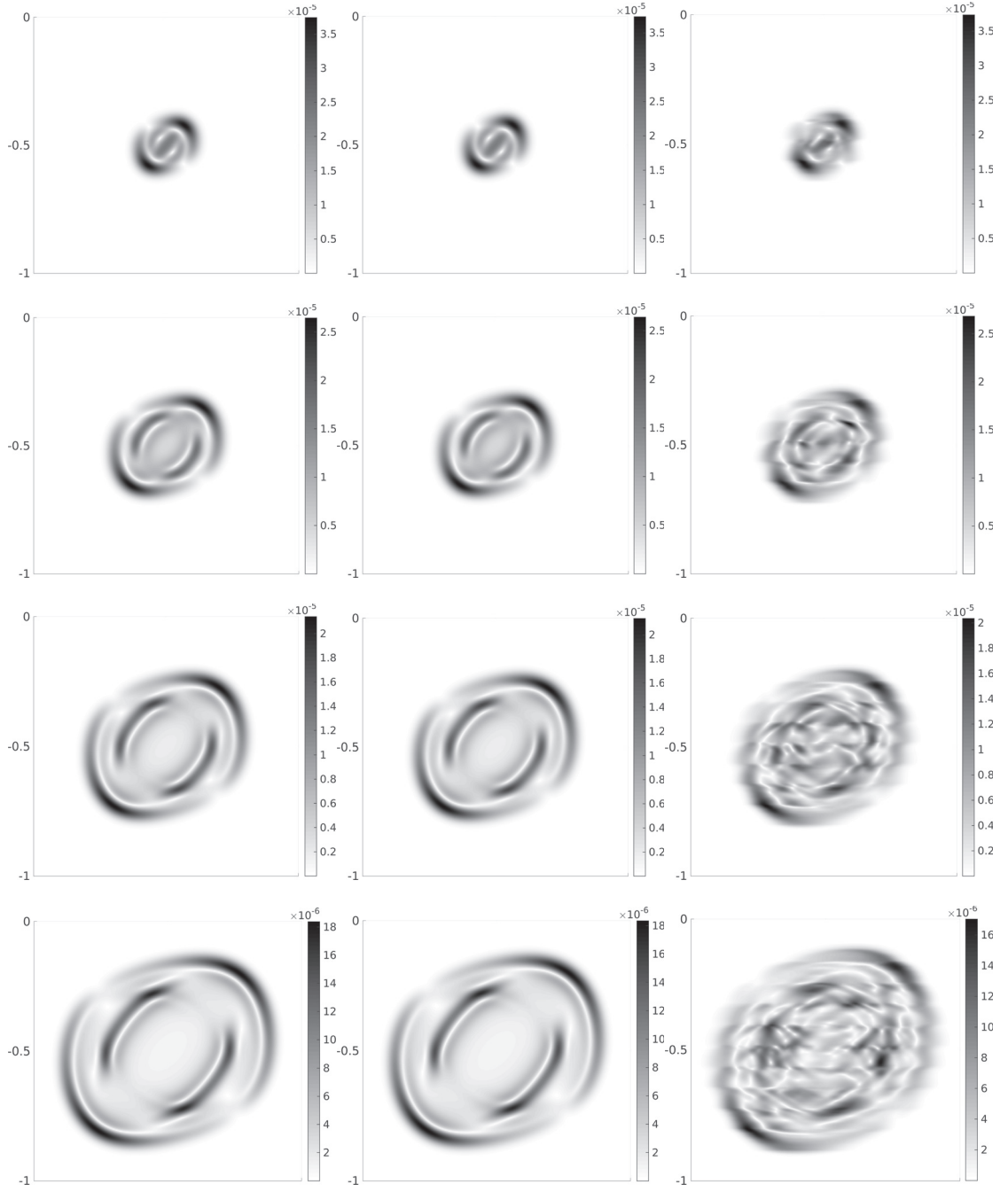


Figure 3.15 – Snapshots at time $t = 0.025, 0.05, 0.075$, and 0.1 second of the numerical homogenized solution (1^{st} column), the effective solution with the tensor (3.39) (2^{nd} column), and the reference heterogeneous solution u^ε (3^{rd} column).

In Figure 3.17 we compare the numerical homogenized solution computed with the FE-HMM using a numerical homogenized tensor computed once on a sampling domain of size $\delta = 8\varepsilon$ (left column) with the numerical homogenized solution computed with the FE-HMM using a numerical homogenized tensor computed in each sampling domains, whose sizes are set to $\delta = 8\varepsilon$ (right column). Both numerical solution are computed using P^2 -macro and P^1 -micro FE. The two solutions give similar displacement.

Further, we compare horizontal snapshots at the depth $z = -0.5$ in the x -direction and z -direction. In Figures 3.18a and 3.18b, we compare the x and z displacements of the numerical homogenized solutions with $a_K^{0,h}$ computed in one sampling domain K_{δ_j} with $\delta = 48\varepsilon$ (in blue) with numerical homogenized solution computed with $a_K^{0,h}$ in one sampling domain K_{δ_j} with $\delta = 8\varepsilon$ (in black). In Figures 3.18c and 3.18d, we compare the x and z displacement of the numerical homogenized solution computed with $a_K^{0,h}$ in one sampling domain K_{δ_j} with $\delta = 8\varepsilon$ (in black) with the numerical solution where the numerical homogenized tensor $a_K^{0,h}$ is computed in all sampling domains K_{δ_j} where $\delta = 8\varepsilon$ (in red). They all give the same behavior.

3.9 Summary

The wave equation in a highly heterogeneous linear elastic medium is considered for short time, and the finite element heterogeneous multiscale method is given. Fully discrete a priori error analysis is derived and numerical examples are proposed to verify the sharpness of the error bounds.

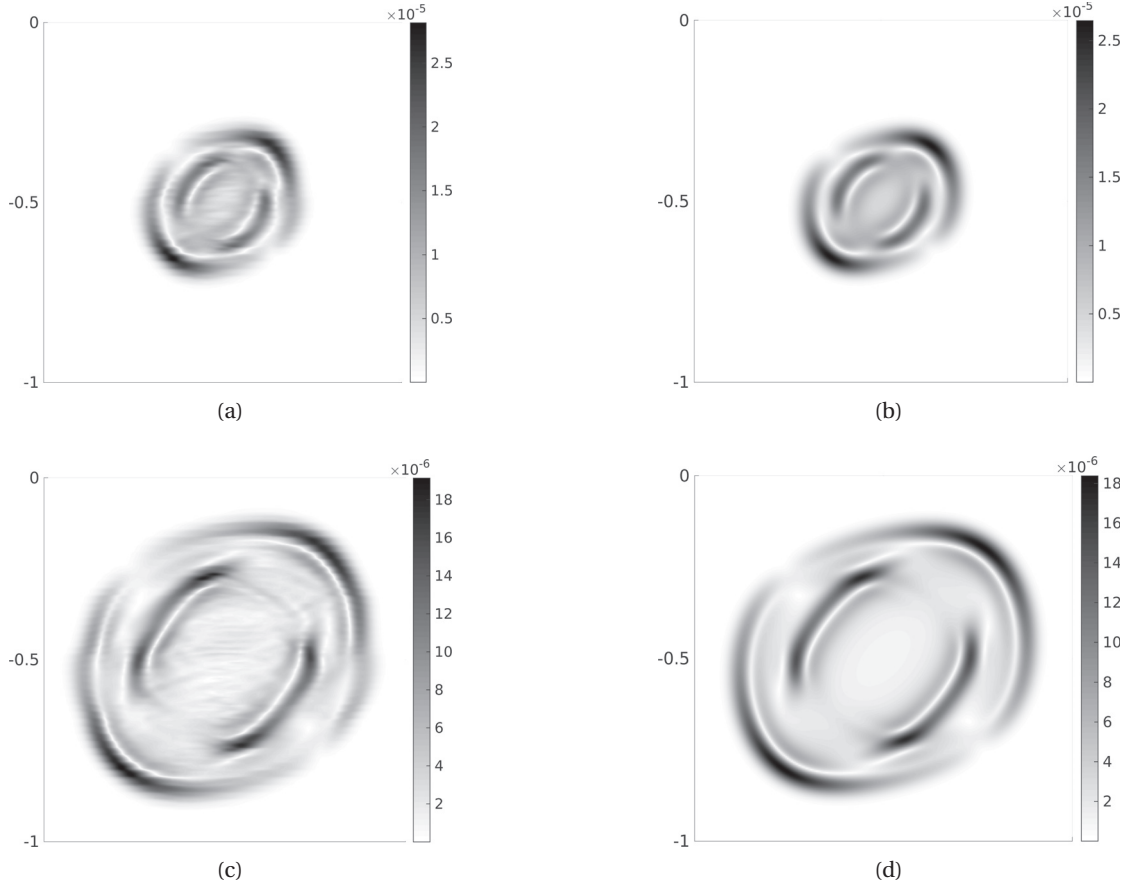


Figure 3.16 – (a) and (c) snapshots at time $t = 0.05$ and 0.1 second of the reference solution u^ϵ for $\epsilon = 1/100$, (b) and (d) snapshots at time $t = 0.05$ and 0.1 second of the numerical homogenized solution u^H .

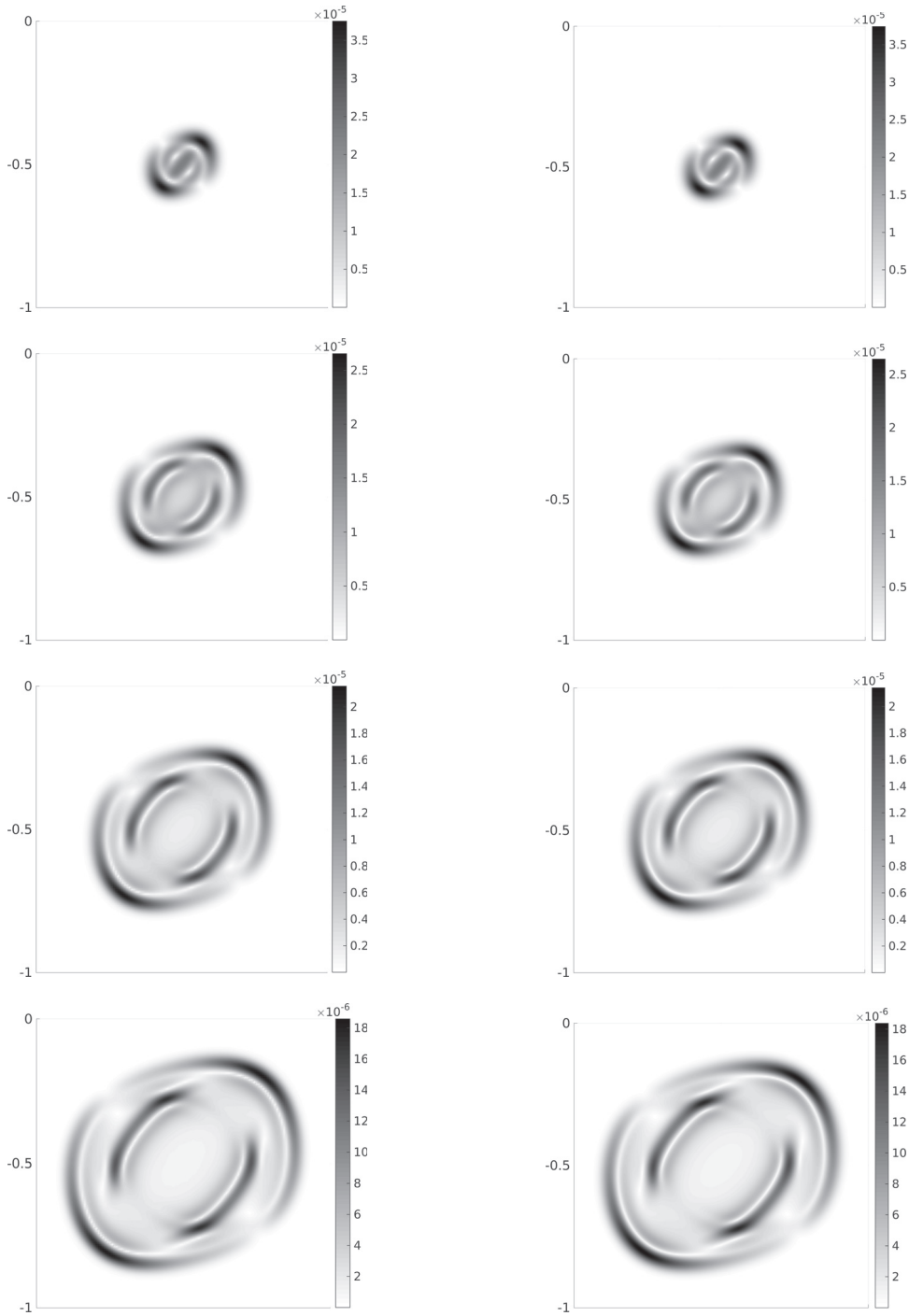


Figure 3.17 – Snapshots at time $t = 0.025, 0.05, 0.075$, and 0.1 second of the numerical homogenized solution with $a_K^{0,h}$ computed in one sampling domain K_{δ_j} with $\delta = 8\epsilon$, and whose value is used at each quadrature points (1st column), snapshots of the homogenized solution with $a_K^{0,h}$ computed in each sampling domain K_{δ_j} with $\delta = 8\epsilon$ (2nd column).

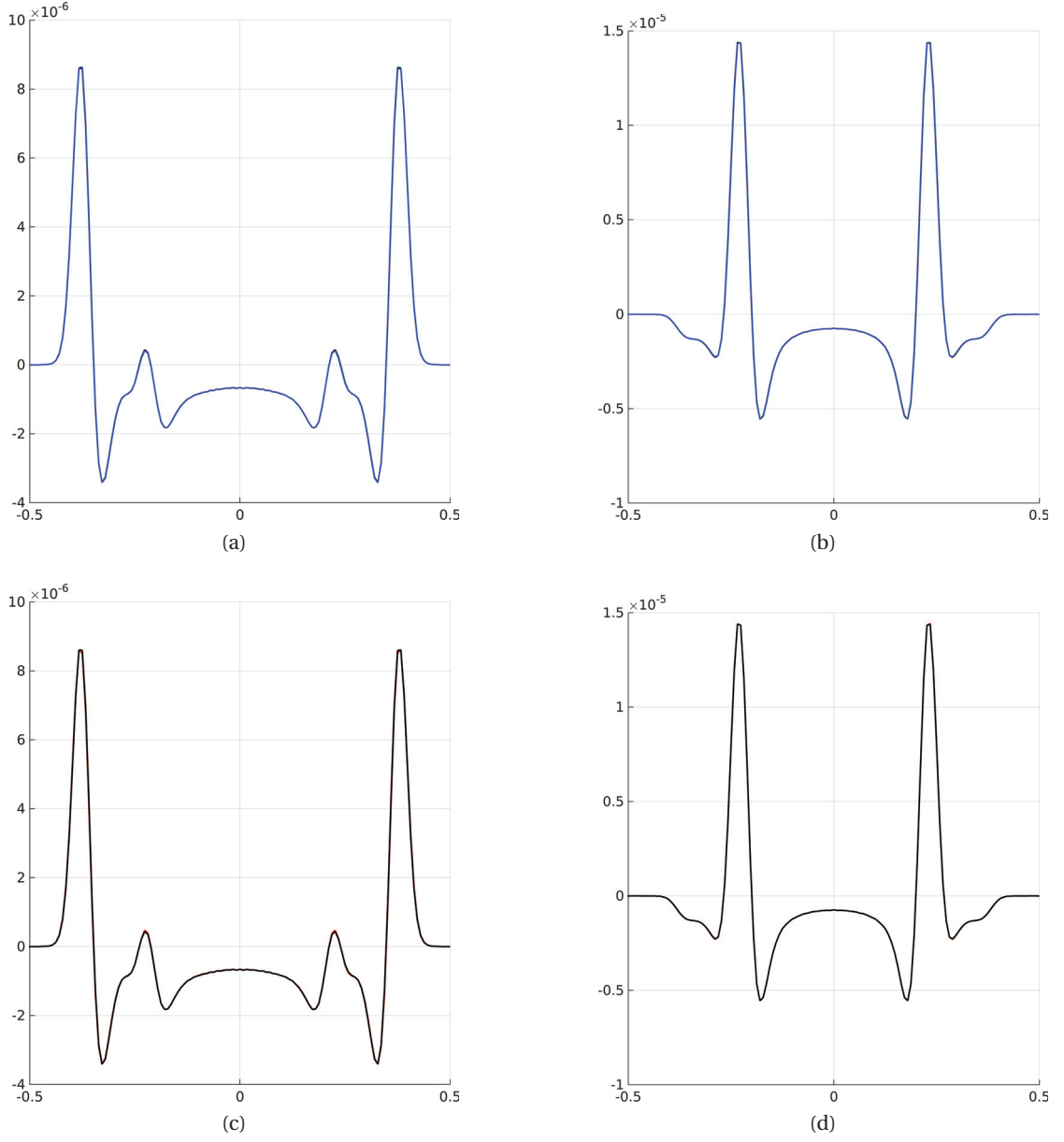


Figure 3.18 – Horizontal snapshots at depth $z = -0.5$ at time $t = 0.1$ of the (a) x -direction and (b) z -direction of the numerical homogenized solution with $a_K^{0,h}$ computed in one sampling domain K_{δ_j} with $\delta = 48\epsilon$ (in blue) and with $\delta = 8\epsilon$ (in black). Horizontal snapshots of the (c) x -direction and (d) z -direction of the numerical homogenized solution with $a_K^{0,h}$ computed in one sampling domain K_{δ_j} with $\delta = 8\epsilon$ (in black) and the numerical homogenized solution with $a_K^{0,h}$ computed in each sampling domain K_{δ_j} with $\delta = 8\epsilon$ (in red).

4 Conclusion and outlook of Part I

The Part I of the thesis, is separated in two chapters. At first, we focus on numerical methods for multiscale elliptic problems, and at second, we consider the wave equation in a linear elastic medium with highly heterogeneous coefficients.

The Chapter 2 gives the foundations of the thesis, as we recall there useful theory, different numerical methods and a priori error analysis that are used throughout the thesis. Advantages and limitations of the proposed numerical methods are given as well. The numerical multiscale method used in the thesis is the finite element heterogeneous multiscale method (FE-HMM) based on the HMM framework. The FE-HMM proposes a fast and robust method to approximate the homogenized solution of a heterogeneous problem, and states a post-processing procedure to recover the fine scales and gives a good approximation of the fine scale solution. However, the method relies on homogenization theory, which requires structural assumptions on a^ε such as a clear scale separation or certain statistical distributions. If the period ε is approximatively known, we construct sampling domains K_{δ_j} with size δ satisfying $\delta/\varepsilon \notin \mathbb{N}$, but this leads to a deterioration of the modeling error from $C\varepsilon$ to $C(\delta + \frac{\varepsilon}{\delta})$, see [56]. If the tensor has no scale separation, and thus no ε to quantify the size of the heterogeneities, the size δ is often set to the finite element size of the partition \mathcal{T}_H over Ω . The fine scales are then used almost everywhere in Ω and the cost of the FE-HMM is no less than that of a fine scale solver. Further, the modeling error has only been derived for periodic and random coefficients, and can therefore not be quantified in terms of explicit convergence rates.

Assuming that in local regions of the domain Ω , the tensor has no explicit separation of scales, for example near local defects in composite material, but present explicit separation of scales in the rest of the domain. In such situations, the FE-HMM can be used in regions with scale separation and a fine scale method (such as the LOD or FEM) should be used elsewhere. This leads to a coupling between two solvers. In Part II, we propose a new coupling method, between the FE-HMM and the FEM, based on optimization techniques and virtual control methods.

In Chapter 3, we derived the FE-HMM for the wave equation in a linear elastic medium. The

Chapter 4. Conclusion and outlook of Part I

derivation of the method followed the FE-HMM framework for elliptic equations derived in Chapter 2 with obvious changes. A priori error estimates are given for the linear elastic problem and for the wave propagation in a linear elastic medium. Explicit convergence rates in term of the macro and micro mesh sizes are derived and numerical examples, using piecewise linear and quadratic finite element spaces, are proposed to verify the sharpness of the bounds. We further test the FE-HMM on layered media and on arbitrarily heterogeneous layered media and compare the FE-HMM results with numerical method used in seismology.

In future work, we would like to test the FE-HMM on more complicated media, such as the Marmousi model [32], and test it with existing numerical methods. We only focus on short time effects, and thus, one future project can be to study the long time effects of a wave propagating through a heterogeneous linear elastic medium.

An optimization based coupling method for highly heterogeneous multiscale elliptic PDEs

Part II

In the second part of the thesis, we derive a new global to local method for a highly heterogeneous multiscale elliptic problem

$$-\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) = f, \quad \text{in } \Omega.$$

As opposed to the Part I, the heterogeneous tensor a^ε , is assumed to have an explicit separation of scales only in some subregions of the domain Ω . We propose a new coupling strategy inspired by the virtual control method and based on a decomposition of the physical domain into a region without scale separation, where the homogenized model is not valid, and a region where the homogenized solution describes adequately the physical problem.

Outline of Part II. The outline of Part II is as follows.

In **Chapter 5**, we give an overview of global to local methods and recall the L^2 global to local projection method and the goal-oriented method. The latter acts as a reference method for our method, and comparison between the two methods are conducted in Chapter 6.

In **Chapter 6**, we develop a new coupling method based on overlapping domains decompositions. Virtual (interface) controls are introduced as boundary conditions at the interface between overlapping subdomains, and the problem is reformulated as a minimization problem with state equations as constraints. In this chapter, the method is based on the minimization of the discrepancy between the models on the overlapping regions. Well-posedness and a priori error analysis are given, first for the optimization based method in the continuous case. Then, the fully discrete coupling method is derived, and existence and uniqueness of the solution are shown. The fully discrete a priori error analysis of the discrete method is proposed. The two analysis rely on the Caccioppoli inequalities and on a strong version of the well-known Cauchy–Schwarz inequality. Numerical examples compare the new method with other existing global to local methods.

In **Chapter 7**, we propose numerical improvements to the coupling methods. At first, we consider a minimization over the boundary of the overlapping regions instead of minimizing over the whole overlapping region, as proposed in Chapter 6. A second improvement is related to the meshing used in the coupling. In Chapter 6, we used identical FE in the overlapping regions, whereas in Chapter 7, we use an interpolation between a fine mesh (for the fine scale solver) and a coarse mesh (for the coarse scale solver). Both improvements reduce the number of degrees of freedom of the problem as well as the computational time, and, when compared to the coupling without these improvements, give similar convergence rates.

Publications. The Chapter 6 about the optimization based method and its a priori analysis is based on [A. Abdulle, O. Jecker, Commun. Math. Sci., 2015] and [A. Abdulle, O. Jecker, A. Shapeev, Multiscale Model. Simul., 2016]. The Chapter 7 about numerical improvements to the coupling method is based on [A. Abdulle, O. Jecker, submitted to publication 2016].

5 Homogenization based global to local methods

In this chapter, we review some homogenization based global to local methods. The methods presented here differ from the multiscale methods of Chapter 2, in the sense that global to local downscaling methods are rather a post-processing step that requires a precomputed homogenized solution. Such methods could be linked to the post-processing procedure of the FE-HMM, given in Section 2.4.

In many applications, the microscopic H^1 features are needed in small subdomains of the physical domain and global to local downscaling methods are a good tool to locally recover the fine scale information in an efficient way. Further, they exploit Caccioppoli inequalities to bound a local H^1 error by an L^2 error over a larger subdomain.

Consider a heterogeneous multiscale elliptic problem: find u^ε such that

$$\begin{aligned} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, & \text{in } \Omega, \\ u^\varepsilon &= g, & \text{on } \Gamma, \end{aligned} \quad (5.1)$$

where $g \in H^{1/2}(\Gamma)$ and $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$, is uniformly bounded and elliptic

$$\exists 0 < \lambda \leq \Lambda: \quad \lambda|\xi|^2 \leq a^\varepsilon(x)\xi \cdot \xi, \quad \text{and} \quad |a^\varepsilon(x)\xi| \leq \Lambda|\xi|, \quad \forall \xi \in \mathbb{R}^d, \quad \text{a.e. } x \in \Omega, \quad \varepsilon > 0. \quad (5.2)$$

The tensor is assumed to have scale separation in Ω , is of the form $a^\varepsilon(x) = a(x, x/\varepsilon)$, and is locally periodic in the fast variable $y = x/\varepsilon$. The FE-HMM (see Section 2.4) gives us a good L^2 approximation u^H of the heterogeneous solution u^ε . Let $\omega \subset \Omega$ be the regions where the fine scales need to be recovered; for example it could be around a defect or a crack in a material. The first step is to consider domains ω_1 , slightly larger than ω , such that $\omega \subset \omega_1 \subset \Omega$ and denote by ω_0 the overlapping regions, i.e., $\omega_0 = \omega_1 \setminus \omega$. The spaces ω, ω_1 , and ω_0 can also share boundaries with Ω . The idea behind the selection of larger domains ω_1 is that it allows to apply Caccioppoli inequalities. Such inequalities are the keys to the error analysis as they bound an H^1 norm on ω by an L^2 norm on ω_1 .

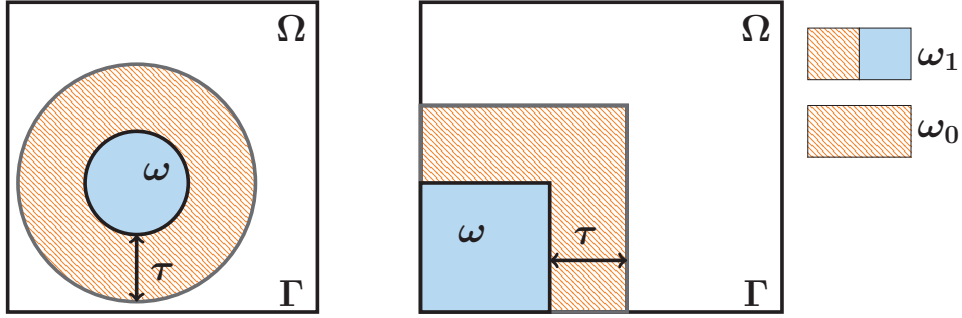


Figure 5.1 – Possible domain decompositions with ω (in blue), ω_0 (in orange, hatched), and $\omega_1 = \omega \cup \omega_0$.

Outline. The outline of this chapter is as follows. In Section 5.1, we recall the Caccioppoli inequalities. In Section 5.2, we give the L^2 global to local projection method, which projects a numerical homogenized solution onto a specific functions space. In Section 5.3, we recall the goal-oriented method which resolves, with a fine scale solver, a heterogeneous multiscale PDE with a numerical homogenized solution as boundary condition.

5.1 Caccioppoli inequalities

In this section, we recall the Caccioppoli inequalities [66] and give some Caccioppoli related inequalities. Let $\omega \subset \omega_1$ be subdomains of Ω with $\tau = \text{dist}(\partial\omega, \partial\omega_1)$ and set $\Gamma = \partial\Omega$; see Figure 5.1 for examples of domain decompositions. For a tensor $a \in (L^\infty(\Omega))^{d \times d}$ verifying (5.2), we define the set of a -harmonic functions by $\mathcal{H}(\omega_1)$, which consists of functions $u \in L^2(\omega_1) \cap H_{\text{loc}}^1(\omega_1)$ such that

$$B(u, v) = \int_{\omega_1} a(x) \nabla u \cdot \nabla v dx = 0, \quad \forall v \in \mathcal{C}_0^\infty(\omega_1).$$

If the domains have shared boundaries, i.e., $\partial\omega_1 \cap \Gamma \neq \emptyset$, we define the space of a -harmonic functions, denoted by $\mathcal{H}_0(\omega_1)$, which consists of functions $u \in \mathcal{H}(\omega_1)$ with zero boundary condition on $\partial\omega_1 \cap \Gamma$; and we recall that $\Gamma_1 = \partial\omega_1 \setminus \Gamma$. The following Caccioppoli inequality holds for interior domains and for domains with shared boundaries.

Theorem 5.1.1 (Caccioppoli inequality [66, 67]). *Let $u \in \mathcal{H}(\omega_1)$, then*

$$\|\nabla u\|_{L^2(\omega)} \leq \frac{C}{\tau} \|u\|_{L^2(\omega_1)},$$

where C depends on the coercivity constants λ and Λ of the tensor a given by (5.2), and where τ is the width of the overlapping domain ω_0 .

Remark 5.1.2. We note that elliptic problem with a non null right-hand side can also be considered and we refer to [66] for details

5.2 L^2 global to local projection method

We discuss here a global to local approach introduced by Babuška and Lipton [27], and based on the L^2 -projection of a numerical homogenized solution onto a specific functions space. To follow the derivation of the method given in [27], we assume that $f = 0$ in problem (5.1). In weak form it reads: find u^ε such that

$$B^\varepsilon(u^\varepsilon, w) = 0, \quad \forall w \in H_0^1(\Omega), \quad (5.3)$$

with $u^\varepsilon = g$ on Γ , and where the bilinear form $B^\varepsilon : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ is given by

$$B^\varepsilon(v, w) = \int_{\Omega} a^\varepsilon(x) \nabla v \cdot \nabla w dx. \quad (5.4)$$

The idea of the method is to project the homogenized solution u^0 , the solution of (5.3) with a^0 instead of a^ε in the bilinear form B^ε given by equation (5.4), onto a functions space spanned by solutions of local problems. The choice of the functions space is crucial and comes from the Caccioppoli inequalities.

Remark 5.2.1. One can generalize the method proposed in [27] to non null right-hand side f in (5.1) as done in [28]. To do so, we need to consider a corrected a^ε -harmonic space, defined as

$$\mathcal{H}^{\text{par}}(\omega_1) := \mathcal{H}(\omega_1) \oplus u_{\text{par}}^\varepsilon,$$

where $u_{\text{par}}^\varepsilon$ denotes a local particular solution of (5.1) in ω_1 subject to suitable Dirichlet boundary conditions on $\partial\omega_1$; for example it can be given by the homogenized solution u^0 .

We will state the result for interior domain decompositions, and refer to [27] for the generalization.

Theorem 5.2.2. *Let $\omega \subset \omega_1 \Subset \Omega$ with $\tau = \text{dist}(\partial\omega_1, \partial\omega) > 0$. Let u^ε be the solution of problem (5.3) and u^0 be the corresponding homogenized solution, and consider their restrictions on ω_1 . The global to local approximation $u_{L^2}^0 \in \mathcal{H}^{\text{par}}(\omega_1)$ is given by the L^2 -projection of u^0 onto $\mathcal{H}^{\text{par}}(\omega_1)$ and satisfies*

$$\|u^0 - u_{L^2}^0\|_{L^2(\omega_1)} = \inf_{w \in \mathcal{H}^{\text{par}}(\omega_1)} \|u^0 - w\|_{L^2(\omega_1)},$$

and further it holds

$$\|\nabla(u^\varepsilon - u_{L^2}^0)\|_{L^2(\omega)} \leq \frac{C}{\tau} \|u^\varepsilon - u^0\|_{L^2(\omega_1)},$$

where C depends on the coercivity constants λ and Λ of the tensor a given by (5.2).

As an immediate consequence of Theorem 5.2.2 we have a result for periodic homogenization.

Theorem 5.2.3. *Let $\omega, \omega_1, u^\varepsilon$, and u^0 be given as in Theorem 5.2.2. Let a^ε be Y -periodic in y , and assume sufficient regularity on the domain and the data to have $\|u^\varepsilon - u^0\|_{L^2(\Omega)} \leq C\varepsilon$ (see*

(2.18)). It holds

$$\|\nabla(u^\varepsilon - u_{L^2}^0)\|_{L^2(\omega)} \leq C\varepsilon,$$

where the constant $C > 0$ depends only on τ , Λ , and λ .

Proof. Follows from (2.18) together with Theorem 5.2.2. \square

5.3 Goal-oriented adaptive method

In this section, we recall the goal-oriented method given by Oden and Vernaganti in [95, 109] and later in [94]. The method can be seen as a model adaptation technique based on local a posteriori error estimates. The small scale information is needed only in small local domains and we assume that outside these small domains, the homogenized, effective property can be used. Let u^ε be the heterogeneous solution of an elliptic PDE

$$-\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) = f, \quad \text{in } \Omega,$$

and that a coarse scale solution \tilde{u} exists and is the solution of a reduced model

$$-\operatorname{div}(\tilde{a}(x)\nabla \tilde{u}) = f, \quad \text{in } \Omega. \quad (5.5)$$

For example, \tilde{a} can be the homogenized tensor a^0 , and in that case the function \tilde{u} is the homogenized solution u^0 . Then a quantity of interest, denoted by $L \in H^{-1}(\Omega)$, is used to measure the local physical features. To estimate the error $L(u^\varepsilon - \tilde{u})$, one needs to define a residual $R \in H^{-1}(\Omega)$ by

$$R(v, w) = \int_{\Omega} (a^\varepsilon - \tilde{a})(x) \nabla v \cdot \nabla w \, dx.$$

Let a^ε be symmetric, and define dual problems: find $v^\varepsilon \in H^1(\Omega)$ and $\tilde{v} \in H^1(\Omega)$ with $v^\varepsilon = g$, $\tilde{v} = g$ on Γ and such that

$$\begin{aligned} \int_{\Omega} a^\varepsilon(x) \nabla w \cdot \nabla v^\varepsilon \, dx &= L(w), \quad \forall w \in H_0^1(\Omega), \\ \int_{\Omega} \tilde{a}(x) \nabla w \cdot \nabla \tilde{v} \, dx &= L(w), \quad \forall w \in H_0^1(\Omega). \end{aligned}$$

Consider $I^\varepsilon = I - (a^\varepsilon)^{-1} \tilde{a}$. The following bounds hold [95, Theorem 3.1].

Theorem 5.3.1. *Let u^ε be the solution of (5.1) and \tilde{u} be the solution of (5.5). It holds*

$$\eta_{low} \leq L(u^\varepsilon - \tilde{u}) \leq \eta_{upp}.$$

The lower and upper bounds are

$$\eta_{low} = \frac{1}{4}(\eta_{low}^+)^2 - \frac{1}{4}(\eta_{upp}^-)^2 + R(\tilde{u}, \tilde{v}) \quad \text{and} \quad \eta_{upp} = \frac{1}{4}(\eta_{upp}^+)^2 - \frac{1}{4}(\eta_{low}^-)^2 + R(\tilde{u}, \tilde{v}),$$

where

$$\eta_{upp}^{\pm} = (a^{\varepsilon} I^{\varepsilon} \nabla(\tilde{u} \pm \tilde{v}), I^{\varepsilon} \nabla(\tilde{u} \pm \tilde{v}))_{L^2(\Omega)}^{1/2} \text{ and } \eta_{low}^{\pm} = \frac{|R(\tilde{u} \pm \tilde{v}, \tilde{u} \pm \tilde{v})|^2}{B^{\varepsilon}(\tilde{u} \pm \tilde{v}, \tilde{u} \pm \tilde{v})}.$$

The algorithm can be given in three steps;

1. solve the reduced model problem (5.5) and obtain \tilde{u} ;
2. estimate the modeling error $L(u^{\varepsilon} - \tilde{u})$. Check if it is below a given tolerance, and if not, determine regions ω where the modeling error is large;
3. enhance the reduced solution \tilde{u} by taking into account the fine-scale material into the regions ω . The problems are formulated as: find u such that

$$\int_{\omega} a^{\varepsilon}(x) \nabla u \cdot \nabla w dx = \int_{\omega} f w dx, \quad \forall w \in H_0^1(\omega),$$

with $u = \tilde{u}$ on $\partial\omega$.

If we consider $\tilde{u} = u^0$, the modeling error will be bounded but not small. To solve that issue, one could take a domain ω_1 such that $\omega \subset \omega_1$, and enhance the reduced solution \tilde{u} (the homogenized solution) using step 3¹. Let $\omega \subset \omega_1 \subset \Omega$, and let $\Gamma_1 = \partial\omega_1 \setminus \Gamma$. Consider \tilde{u} the solution of

$$-\operatorname{div}(a^{\varepsilon}(x) \nabla \tilde{u}) = f, \quad \text{in } \omega_1, \quad (5.6)$$

with $\tilde{u} = u^0$ on Γ_1 and $\tilde{u} = g$ on $\partial\omega_1 \cap \Gamma$. A priori error estimates can be derived using the Caccioppoli inequalities.

Theorem 5.3.2. *Let u^{ε} be the solution of (5.1) and \tilde{u} be the solution of (5.6)*

$$\|u^{\varepsilon} - \tilde{u}\|_{H^1(\omega)} \leq C\varepsilon + \|u^{\varepsilon} - u^0\|_{H^{1/2}(\Gamma_1)},$$

where the constant C depends on λ, Λ , and τ .

Remark 5.3.3. In Theorem 5.3.2, the difference between u^0 and u^{ε} is small only in the L^2 norm, and thus by the trace theorems, we have that the difference in the $H^{-1/2}$ norm is small; i.e.,

$$\|u^{\varepsilon} - u^0\|_{H^{-1/2}(\Gamma_1)} \leq C\|u^{\varepsilon} - u^0\|_{L^2(\omega_1)} \leq C\varepsilon.$$

However, the norm in the $H^{1/2}$ is bounded but not small.

¹An other solution is to add a condition on the reduced model in order to guarantee that the modeling error is small; i.e., one should consider \tilde{u} a solution of the reduced problem (5.5) with

$$\|\tilde{u} - u^{\varepsilon}\|_{L^2(\Omega)} \leq C\varepsilon, \quad \text{and} \quad \|\tilde{u} - u^{\varepsilon}\|_{H^1(\Omega \setminus \omega)} \leq C\varepsilon.$$

5.4 Summary

In this chapter, we reviewed two global to local downscaling methods, the L^2 global to local projection method and the goal-oriented adaptive method. The two methods allow for a recovery of the fine scales in region of interest but are based on the assumption that a precomputed homogenized solution is available. However, it might be expensive to compute the numerical homogenized solution with one multiscale method given in Section 2.3. Further the error estimates depend on the quality of the numerical homogenized solution.

6 Optimization based coupling method

In this chapter, we propose a new multiscale method for elliptic problems with highly oscillatory coefficients based on virtual controls and optimization techniques. This chapter is based on the articles [13] and [16].

Consider a heterogeneous multiscale elliptic problem in a convex, polygonal domain $\Omega \subset \mathbb{R}^d$, with Lipschitz continuous boundary Γ ; find u^ε such that

$$-\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) = f, \quad \text{in } \Omega, \quad (6.1)$$

with some boundary conditions on Γ , and let the tensor $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$ be uniformly bounded and elliptic

$$\exists 0 < \lambda \leq \Lambda : \lambda |\xi|^2 \leq a^\varepsilon(x) \xi \cdot \xi, \text{ and } |a^\varepsilon(x) \xi| \leq \Lambda |\xi|, \quad \forall \xi \in \mathbb{R}^d, \text{ a.e. } x \in \Omega, \varepsilon > 0. \quad (6.2)$$

As opposed to the Chapter 2, the heterogeneous tensor a^ε , considered in problem (6.1), is assumed to have an explicit separation of scales only in some subregions of the domain Ω . By explicit scale separation, we mean problems where the medium has a characteristic length that can be defined by a small parameter $\varepsilon > 0$, and where the solution has a scale larger than ε . We speak of micro and macro scales and are able to separate them explicitly. The assumption of scale separation allows to take full advantage of the numerical homogenization method FE-HMM given in Section 2.4, where the fine scales are needed only in sampling domains of size $\mathcal{O}(\varepsilon)$. The cost of the method is then reduced as the fine scales are needed only in small sampling domains and that the partition over Ω can be coarse.

When no separation of scales are considered, the size of the sampling domains remains an open question and one possible answer is to take sampling domains of size H , where H is the size of the mesh used to partition Ω . However, this is a costly choice as it results in a numerical method with a cost no less than that of the FEM. The FE-HMM provides us with a good tool in regions with scale separation, but in regions without scale separation, the approach should be different.

6.1 Literature overview

Numerous methods have been developed in the past decades, and we give here a non-exhaustive list.

At first, we mentioned the methods based on coarse oscillatory basis functions such as the MsFEM [58] and the LOD method [88, 74] (for a recap, see 2.3.1). These methods can be applied to problems with general coefficients — that is without assumptions on separability — but they come with a high computational cost. Indeed, to precompute the coarse basis functions, fine scale problems have to be solved on localized coarse elements whose union is a partition of the computational domain Ω .

At second, we regroup the global to local methods such as the L^2 projection method [27] (see Section 5.2) or the goal-oriented method [95] (see Section 5.3). The L^2 projection and goal-oriented methods couple a numerical homogenization method with a fine scale solver, with the advantage that the fine scale solver is used only in small local subdomains. However, these methods rely on the availability of a good numerical homogenized solution on Ω , which is computationally expensive to obtain; especially when no scale separation are involved and FE-HMM is used. Further, assuming that the high cost is not an issue, the error estimates of such methods are dependent on the accuracy of the numerical homogenized solution.

In this thesis, we derive a new coupling strategy inspired by the virtual control method pioneered in [71, 87, 65] (see also [49] for recent developments). We note that such problems have numerous applications in the sciences, we mention for example heterogeneous structures with defects [63, 31] or steady flow problems with singularities [64]. The coupling strategy derived here shares some similarities with the recent work on atomistic-to-continuum coupling [97, 98]. For a convergence analysis of quasi-continuum methods, we refer to [83, 84, 17] and the reference therein. We mention, further, the energy-based coupling method [36], where the strategy is to compose the energy of the problem from the homogenized and heterogeneous state equations. Then one steer the system to a stable equilibrium. Atomistic-to-continuum and energy-based coupling methods arise in the search of a coupling method between different description of material, such as continuous to atomistic coupling in the analysis of material with defects [107, 110, 105] and the references therein. We cite as-well the recent work in [46] on the coupling of local and nonlocal diffusions models.

The method is based on a decomposition of the domain Ω into a region without scale separation, denoted by ω , where the homogenized model is not valid, and a region ω_2 where the homogenized solution describes adequately the physical problem. We consider a neighbourhood of ω , denoted ω_0 , where both the fine scale and homogenization models are valid; the subdomain $\omega_0 \subset \omega_2$ plays the role of an overlapping region. Figure 6.1 illustrates two possible domains decompositions, with interior domains $\omega \Subset (\omega \cup \omega_0) \Subset \Omega$ (left) and with shared boundaries $\partial\omega \cap \Gamma \neq \emptyset$ (right). Note that the domains ω, ω_0 can be polyhedrons as well, and that their representation as circles is a mere choice.

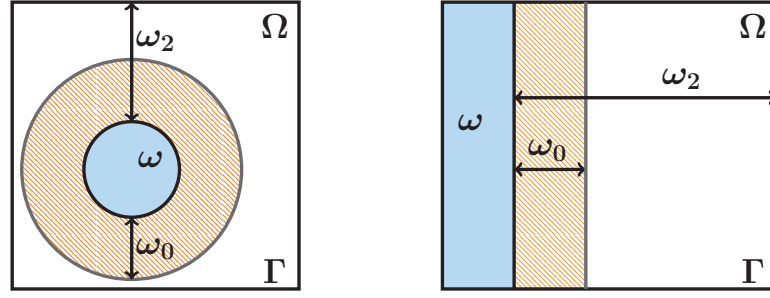


Figure 6.1 – Illustration of two possible domains decompositions with interior domains (left) and with shared boundaries (right).

Outline. The chapter is separated in two parts. In the first part, we derive the continuous optimization based method, and the second part give the fully discrete optimization based coupling method. Before giving the coupling methods, we recall in Section 6.2, some theoretical results about optimal and virtual controls, useful for the analysis of the method.

6.2 Optimal control of partial differential equations

In this section, we give basic results about the theory of optimal and virtual controls given by Lions [85]; we will use its notations. Let \mathcal{L} be an elliptic operator and seek u the solution of

$$\mathcal{L}(u) = f,$$

with some boundary conditions, and assume that u depends on some value θ ; for example θ could be some boundary conditions or some distributions on the right-hand side f . The solution u of the elliptic problem is called the state and θ the control.

The goal is then to find a control θ such that it minimizes a cost functional J that depends on the controls θ (as well as on the state variable u). The controls belong to a Hilbert space \mathcal{U} , which is a functional space describing the role of θ in the elliptic PDE. The admissible controls θ considered as potential minimizers are found in a space $\mathcal{U}_{ad} \subseteq \mathcal{U}$ called the space of admissible controls.

The optimal control problem is: find $\theta \in \mathcal{U}_{ad}$ such that

$$J(\theta) = \inf J(\mu), \quad \forall \mu \in \mathcal{U}_{ad}.$$

If $\mathcal{U} = \mathcal{U}_{ad}$ the optimal control problem is unconstrained, otherwise it is said to be constrained. We assume that the cost functional J can be written as

$$J(\mu) = \pi(\mu, \mu) - 2F(\mu), \quad \forall \mu \in \mathcal{U}_{ad},$$

where $\pi : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$ is a symmetric, bilinear form on \mathcal{U} , and $F : \mathcal{U} \rightarrow \mathbb{R}$ is a continuous linear

Chapter 6. Optimization based coupling method

form on \mathcal{U} . The following result given in [85] guarantees the existence and uniqueness of the optimal control θ . For the proof we refer to [85, Chapter 1] and [101, Theorem 16.1].

Theorem 6.2.1. *Let π be a continuous symmetric bilinear form on \mathcal{U} . If in addition π is coercive on \mathcal{U} ; i.e.,*

$$\pi(\mu, \mu) \geq C \|\mu\|_{\mathcal{U}}^2, \quad \forall \mu \in \mathcal{U}, \quad C > 0,$$

then there exists a unique element $\theta \in \mathcal{U}_{ad}$ such that

$$J(\theta) = \inf_{\mu \in \mathcal{U}_{ad}} J(\mu).$$

Further, if $\mathcal{U}_{ad} = \mathcal{U}$, the optimal control θ satisfies the Euler equation

$$\pi(\theta, \mu) = F(\mu), \quad \forall \mu \in \mathcal{U}.$$

Continuous optimization based method

In here we give a new coupling method for elliptic problem with highly oscillatory coefficients with non scale separation.

The domain Ω is decomposed into a family of overlapping domains and virtual (interface) controls are introduced as boundary conditions at the interface between overlapping subdomains. The interface controls act as unknown traces or fluxes and the problem is reformulated as a minimization problem with state equations as constraints. The optimal boundary controls of overlapping domains are found by an optimization problem that is based on minimizing the discrepancy between the models in the overlapping regions.

Let Ω be a convex, polygonal domain in \mathbb{R}^d , $d = 1, 2, 3$, with a boundary $\Gamma = \Gamma_D \cup \Gamma_N$; where Dirichlet conditions are imposed on Γ_D and Neumann conditions on Γ_N . We further assume that $\Gamma_D \cap \Gamma_N = \emptyset$ and that Γ_D has positive measure. Let $f \in L^2(\Omega)$, $g_D \in H^{1/2}(\Gamma_D)$, and $g_N \in L^2(\Gamma_N)$, and consider the following second-order elliptic problem: find u^ε such that

$$\begin{aligned} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, \quad \text{in } \Omega, \\ u^\varepsilon &= g_D, \quad \text{on } \Gamma_D, \\ n \cdot (a^\varepsilon(x)\nabla u^\varepsilon) &= g_N, \quad \text{on } \Gamma_N, \end{aligned} \tag{6.3}$$

where the coefficients of the tensor $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$ are highly oscillatory and bounded with scale separation only in some subregions of Ω . Further, a^ε is uniformly elliptic (6.2). Thanks to the Lax–Milgram lemma, problem (6.3) is well-posed, and admits a family of solutions $\{u^\varepsilon\}$ indexed by ε . The subscript ε denotes the length of the heterogeneities present in the region with explicit separation of scales. In general, the heterogeneities in the regions without separation of scales cannot be characterized by a quantity ε . However, it is convenient to keep the subscript ε to indicate the presence of heterogeneities.

Outline. We start by recalling the model problem and derive the optimization based coupling method applied to the model problem. In Section 6.3 we prove that the coupling in the continuous case is well-posed, using some Caccioppoli inequalities and a strong Cauchy–Schwarz inequality. In Section 6.4, we derive the optimality systems associated to the minimization problem. In Section 6.5, we derive a priori error estimates, where it is shown that, by using a

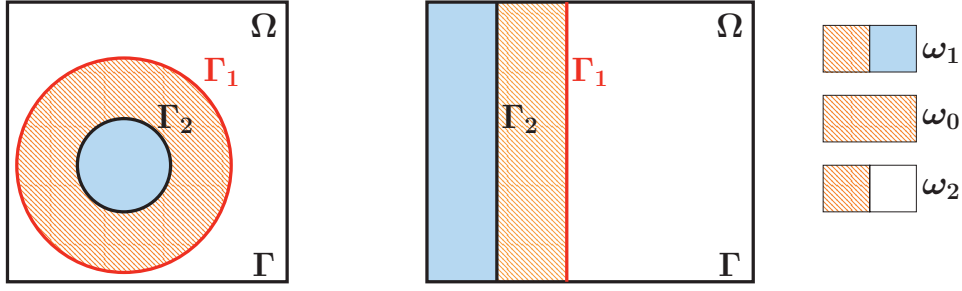


Figure 6.2 – Two possible domains decompositions with Γ_1 (in red) and Γ_2 (in black) for interior domains (left) and for domains with shared boundaries (right).

Caccioppoli inequality, the H^1 error estimate in the region without scale separation can be bounded by an L^2 norm. In the region with scale separation an energy approximation towards the fine scale problems can also be obtained through the use of a locally periodic corrector. Overall, we are able to obtain H^1 convergence rates towards the fine scale solution over the whole computational domain Ω .

Model problem. For simplicity, we decompose the domain Ω into two regions; one with scale separation and the other without. The overlapping domain decomposition is as follows. Let ω denote a subregion of Ω , without scale separation and consider two overlapping subdomains, ω_1 and ω_2 , with $\omega \cup \omega_2 = \Omega$, and $\omega \subset \omega_1$. The overlapping region, also referred to by interface region, is denoted by ω_0 , and it holds $\omega_1 \cap \omega_2 = \omega_0$; Figure 6.2 illustrates possible domain decompositions. Let the tensor a^ε be given by

$$a^\varepsilon(x) = a_\omega^\varepsilon(x) \mathbb{1}_\omega(x) + a_2^\varepsilon(x) \mathbb{1}_{\omega_2}(x), \quad x \in \Omega,$$

where $\mathbb{1}_\omega$ (resp. $\mathbb{1}_{\omega_2}$) denotes the characteristic function associated to the subdomain ω (resp. ω_2). Further, the tensor a_2^ε has scale separation, e.g. $a_2^\varepsilon(x) = a_2(x, x/\varepsilon)$, and is locally periodic in the fast variable, with period $\varepsilon > 0$, whereas the tensor a_ω^ε is a highly heterogeneous tensor without spacial assumptions. From the theory of H -convergence [90], the heterogeneous tensor a_2^ε H -converges towards a homogenized tensor a_2^0 . In ω , the scales are not well separated in the tensor a_ω^ε , which prevents the use of numerical homogenization methods. The heterogeneities can also be present in the right hand side f , and following homogenization theory, the smooth part of f converges to a function f^0 , when the size of the heterogeneities goes to zero, see the treatment in [80, 44]. However, depending on the nature of the heterogeneities — if the source term is singular for example — resolving the micro problems around quadrature points on a sampling domain of size δ could give unacceptable results, and one might have to increase the size δ . By doing so, the fine scale are needed almost everywhere in Ω and the cost becomes prohibitive.

Let $\Gamma_1 = \partial\omega_1 \setminus \Gamma$ and $\Gamma_2 = \partial\omega_2 \setminus \Gamma$ be Lipschitz continuous boundaries; see Figure 6.2. The optimization based problem is as follows; in ω_1 we solve a heterogeneous boundary value problem, and in ω_2 we solve a homogenized boundary value problem. The two solutions are

defined on the overlap ω_0 , and to couple the two boundary value problems, we minimize a cost function involving the difference between the two solutions. In this thesis, we will consider two costs;

Case 1. Minimization in $L^2(\omega_0)$, with

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)}^2. \quad (\text{case 1})$$

Case 2. Minimization in $L^2(\Gamma_1 \cup \Gamma_2)$, with

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2. \quad (\text{case 2})$$

The first one minimizes the difference over the overlap ω_0 and is the subject of this chapter. The second cost minimizes the difference over the boundary $\Gamma_1 \cup \Gamma_2$ of ω_0 ; this presents numerical advantages as it reduces the number of degrees of freedom of the coupling problem. This second cost is treated in Chapter 7.

The problems reads: find $u_1^\varepsilon \in H^1(\omega_1)$ and $u_2^0 \in H^1(\omega_2)$, such that the cost function

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)}^2,$$

is minimized under the following constraints, for $i = 1, 2$,

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla u_i(\theta_i)) &= f, & \text{in } \omega_i, \\ u_i(\theta_i) &= \theta_i, & \text{on } \Gamma_i, \\ u_i(\theta_i) &= g_D, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n_i \cdot (a_i(x)\nabla u_i(\theta_i)) &= g_N, & \text{on } \partial\omega_i \cap \Gamma_N, \end{aligned} \quad (6.4)$$

where $u_1 = u_1^\varepsilon$ and $u_2 = u_2^0$, and where the boundary conditions θ_i , which we refer to as virtual (interface) controls, are to be determined. Here and in what follows, we will sometimes use the short hand notations u_i to denote $u_i(\theta_i)$, for $i = 1, 2$, and

$$\begin{aligned} a_1 &= a_1^\varepsilon = a_\omega^\varepsilon \mathbb{1}_\omega + a_2^\varepsilon \mathbb{1}_{\omega_0}, & u_1 &= u_1^\varepsilon, \\ a_2 &= a_2^0, & u_2 &= u_2^0. \end{aligned}$$

We define the space of admissible Dirichlet controls \mathcal{U}_i^D for equations (6.4) by

$$\mathcal{U}_i^D = \{\mu_i \in H^{1/2}(\Gamma_i) \mid \exists u \in H^1(\omega_i), u|_{\Gamma_i} = \mu_i, \text{ in the sense of the trace}\}.$$

Here, we made the assumptions that the boundary controls are in $H^{1/2}$, in order to have $u_i \in H^1(\omega_i)$. By taking less regularity for the controls, we lose the guarantee that $u_i \in H^1(\omega_i)$. However, by the transposition method [86], one can prove the existence and uniqueness of a very weak solution u_i satisfying a very weak formulation of problem (6.4). One could also

Chapter 6. Optimization based coupling method

consider Neumann boundary controls instead of Dirichlet controls, and follow the theory with some adjustments.

For simplicity, we set $\mathcal{U}_{ad} := \mathcal{U}_1^D \times \mathcal{U}_2^D$, and define for $i = 1, 2$

$$\begin{aligned} H_D^1(\omega_i) &= \{u \in H^1(\omega_i) \mid u = 0 \text{ on } \partial\omega_i \cap \Gamma_D\}, \\ H_{D,\Gamma_i}^1(\omega_i) &= \{u \in H^1(\omega_i) \mid u = 0 \text{ on } \partial\omega_i \cap \Gamma_D \text{ and } \Gamma_i\}. \end{aligned}$$

Let $\gamma_D : H^1(\Omega) \rightarrow H^{1/2}(\Gamma_D)$ denote a linear continuous map, called the trace map. As g_D is in $H^{1/2}(\Gamma_D)$, there exists a function $R_{g_D} \in H^1(\Omega)$, called a lifting of the boundary data g_D , such that $\gamma_D(R_{g_D}) = g_D$. Further, there exists a constant $C(\Omega)$ depending on Ω such that

$$\|R_{g_D}\|_{H^1(\Omega)} \leq C(\Omega) \|g_D\|_{H^{1/2}(\Gamma_D)}.$$

The strategy is to solve a minimization problem in the space of admissible controls \mathcal{U}_{ad} , as explained in Section 6.2; find $(\theta_1, \theta_2) \in \mathcal{U}_{ad}$ such that

$$J(\theta_1, \theta_2) = \min \frac{1}{2} \|u_1^\varepsilon(\mu_1) - u_2^0(\mu_2)\|_{L^2(\omega_0)}^2, \quad \forall (\mu_1, \mu_2) \in \mathcal{U}_{ad}, \quad (6.5)$$

where u_1^ε and u_2^0 are the solutions of (6.4). From the theory of optimal control given in Section 6.2, the problem (6.5) admits a unique couple (θ_1, θ_2) in \mathcal{U}_{ad} if the cost functional J is coercive on a space of controls \mathcal{U} , with $\mathcal{U}_{ad} \subseteq \mathcal{U}$. In Section 6.5 we prove that J is a norm over \mathcal{U}_{ad} . Thus we will set $\mathcal{U}_{ad} = \mathcal{U}$ leading to an unconstrained optimization based problem.

Following the virtual control method exposed in [65], we split the solutions in two parts,

$$u_1^\varepsilon(\theta_1) = u_{1,0}^\varepsilon + v_1^\varepsilon(\theta_1), \quad u_2^0(\theta_2) = u_{2,0}^0 + v_2^0(\theta_2),$$

The functions $u_{i,0}$ are independent of the controls, whereas the functions v_i are dependent of θ_i , for $i = 1, 2$. One calls $(v_1^\varepsilon, v_2^0) \in H_D^1(\omega_1) \times H_D^1(\omega_2)$ the state variables, and they satisfy, for $i = 1, 2$,

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla v_i) &= 0, & \text{in } \omega_i, \\ v_i &= \theta_i, & \text{on } \Gamma_i, \\ v_i &= 0, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n_i \cdot (a_i(x)\nabla v_i) &= 0, & \text{on } \partial\omega_i \cap \Gamma_N, \end{aligned} \quad (6.6)$$

where $v_1 = v_1^\varepsilon$, and $v_2 = v_2^0$. Here again, we have made an abuse of notations and use v_i to denote $v_i(\theta_i)$, for $i = 1, 2$. The function $u_{i,0} \in H_{D,\Gamma_i}^1(\omega_i)$ satisfies,

$$B_i(u_{i,0}, w) = F(w_i), \quad \forall w \in H_{D,\Gamma_i}^1(\omega_i), \quad (6.7)$$

where the bilinear form $B_i : H^1(\omega_i) \times H^1(\omega_i) \rightarrow \mathbb{R}$ is given by

$$B_i(u, w) = \int_{\omega_i} a_i(x) \nabla u \cdot \nabla w \, dx, \quad (6.8)$$

and the right-hand side $F_i : H^1(\omega_i) \rightarrow \mathbb{R}$, by

$$F_i(w) = \int_{\omega_i} f w \, dx - \int_{\omega_i} a_i(x) \nabla R_{g_D} \cdot \nabla w \, dx + \int_{\partial\omega_i \cap \Gamma_N} g_N w \, ds.$$

The state solution $v_i \in H_D^1(\omega_i)$ verifies, for $i = 1, 2$,

$$B_i(v_i, w) = 0, \quad \forall w \in H_{D, \Gamma_i}^1(\omega_i).$$

Thanks to the Lax–Milgram lemma, there exist a unique solution $u_{2,0}^0$ and a family of solutions $\{u_{1,0}^\varepsilon\}$ indexed by ε . Moreover, if the virtual controls θ_1 and θ_2 are given, the solutions v_1^ε and v_2^0 can be uniquely determined. The solutions $u_{1,0}^\varepsilon$ and $u_{2,0}^0$ can be computed before the coupling as they are independent of the virtual controls (θ_1, θ_2) .

6.3 Existence and uniqueness

The well-posedness of the minimization problem (6.5) is proved following the virtual control theory established by Lions [85] and recalled in Section 6.2. The cost is given by

$$J(\mu_1, \mu_2) = \frac{1}{2} \|u_1^\varepsilon(\mu_1) - u_2^0(\mu_2)\|_{L^2(\omega_0)}^2, \quad (\mu_1, \mu_2) \in \mathcal{U}.$$

Using the splitting of the solutions $u_i(\mu_i)$ into $v_i(\mu_i)$ and $u_{i,0}$, the cost functions can be written as

$$\begin{aligned} J(\mu_1, \mu_2) &= \frac{1}{2} \|v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)\|_{L^2(\omega_0)}^2 + \frac{1}{2} \|u_{1,0}^\varepsilon - u_{2,0}^0\|_{L^2(\omega_0)}^2 \\ &\quad + \int_{\omega_0} (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) (u_{1,0}^\varepsilon - u_{2,0}^0) \, dx. \end{aligned}$$

Let us define a bilinear form $\pi : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$ and a form $F : \mathcal{U} \rightarrow \mathbb{R}$; for (μ_1, μ_2) and $(\theta_1, \theta_2) \in \mathcal{U}$, set

$$\pi((\theta_1, \theta_2), (\mu_1, \mu_2)) = \int_{\omega_0} (v_1^\varepsilon(\theta_1) - v_2^0(\theta_2)) (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) \, dx, \quad (6.9)$$

$$F(\mu_1, \mu_2) = - \int_{\omega_0} (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) (u_{1,0}^\varepsilon - u_{2,0}^0) \, dx. \quad (6.10)$$

Then, one can write the cost functions in terms of π and F ,

$$J(\mu_1, \mu_2) = \frac{1}{2} \pi((\mu_1, \mu_2), (\mu_1, \mu_2)) - F(\mu_1, \mu_2) + \frac{1}{2} \|u_{1,0}^\varepsilon - u_{2,0}^0\|_{L^2(\omega_0)}^2.$$

Chapter 6. Optimization based coupling method

Following [85], the existence and uniqueness of the optimal controls hold if the form π is a scalar product on the space of admissible controls. In that sense, the optimal control problem is unconstrained and Theorem 6.2.1 can be applied.

To prove the coercivity of the form π , we need some preliminary inequalities. The first set of inequalities are related to the Caccioppoli inequalities are stated in Section 5.1 and they bound a gradient norm by an L^2 norm on a slightly larger domain. The second inequality is a strong version of the Cauchy–Schwarz inequality.

Caccioppoli related inequalities

We recall that for a tensor $a \in (L^\infty(\Omega))^{d \times d}$ verifying (5.2), the set of a -harmonic functions is defined by $\mathcal{H}(\omega_1)$, which consists of functions $u \in L^2(\omega_1) \cap H_{\text{loc}}^1(\omega_1)$ such that

$$B(u, v) = \int_{\omega_1} a(x) \nabla u \cdot \nabla v \, dx = 0, \quad \forall v \in \mathcal{C}_0^\infty(\omega_1).$$

Lemma 6.3.1. *Let $\omega_0 = \omega_1 \setminus \overline{\omega}$ and $u \in \mathcal{H}(\omega_1)$. Then*

$$\|\nabla u\|_{L^2(\omega)} \leq \frac{\Lambda^{1/2}}{\lambda^{1/2}\tau} \|u\|_{L^2(\omega_0)},$$

where λ and Λ are given by (6.2) and τ is the width of the overlapping domain ω_0 .

Proof. Let $\eta \in \mathcal{C}_0^1(\omega_1)$ be a cutoff function with $\eta = 1$ in $\overline{\omega}$, $\eta = 0$ in $\partial\omega_1$, and $|\nabla\eta| \leq 1/\tau$. Further, $\eta = 0$ on Γ_1 and $\text{supp}(\nabla\eta) \subset \omega_0$. Then, it holds that $\eta^2 u \in H_0^1(\omega_1)$ and

$$\int_{\omega_1} a \nabla u \cdot \nabla (\eta^2 u) \, dx = 0.$$

Then,

$$0 = \int_{\omega_1} a \nabla u \cdot \nabla (\eta^2 u) \, dx = 2 \int_{\omega_1} a \nabla u \cdot \nabla \eta \eta u \, dx + \int_{\omega_1} a \nabla u \cdot \nabla u \eta^2 \, dx.$$

Using the ellipticity of a and the definition of η , it holds

$$\lambda \|\nabla u\|_{L^2(\omega)}^2 \leq \int_{\omega_1} a \nabla (\eta u) \cdot \nabla (\eta u) \, dx,$$

and

$$\begin{aligned}
 \int_{\omega_1} a \nabla(\eta u) \cdot \nabla(\eta u) dx &= \int_{\omega_1} a \nabla(\eta u) \cdot \nabla(\eta u) dx - \int_{\omega_1} a \nabla u \cdot \nabla(\eta^2 u) dx \\
 &= \int_{\omega_1} a \nabla(\eta u) \cdot \nabla(\eta u) dx - 2 \int_{\omega_1} a \nabla u \cdot \nabla \eta \eta u dx - \int_{\omega_1} a \nabla u \cdot \nabla u \eta^2 dx \\
 &= \int_{\omega_1} a \nabla \eta \cdot \nabla \eta u^2 dx \\
 &= \int_{\omega_0} a \nabla \eta \cdot \nabla \eta u^2 dx \\
 &\leq \frac{\Lambda}{\tau^2} \int_{\omega_0} u^2 dx = \frac{\Lambda}{\tau^2} \|u\|_{L^2(\omega_0)}^2.
 \end{aligned}$$

□

Lemma 6.3.2. *Let v_1^ε and v_2^0 be the solutions of (6.6), for $i = 1, 2$, respectively. The following bounds hold*

$$\begin{aligned}
 \|v_1^\varepsilon\|_{L^2(\omega)} &\leq \frac{C}{\tau} \|v_1^\varepsilon\|_{L^2(\omega_0)}, \\
 \|v_2^0\|_{L^2(\Omega \setminus \omega_1)} &\leq \frac{C}{\tau} \|v_2^0\|_{L^2(\omega_0)},
 \end{aligned}$$

where τ is the width of the overlap and C is a constant depending on λ, Λ , and the Poincaré constant associated to ω_1 and ω_2 , respectively.

Proof. We prove the Lemma for the function v_1^ε . Let η be a cutoff function such that $\eta = 1$ in $\bar{\omega}$, $\eta = 0$ in $\Omega \setminus \omega_1$ and $|\nabla \eta| \leq 1/\tau$. Further, we have $\text{supp}(\nabla \eta) \subset \omega_0$. Then, $\eta v_1^\varepsilon \in H_0^1(\omega_1)$, and due to (6.6) and using Poincaré inequality, it holds

$$\|v_1^\varepsilon\|_{L^2(\omega)} \leq \|\eta v_1^\varepsilon\|_{L^2(\omega_1)} \leq C_{\omega_1} \|\nabla(\eta v_1^\varepsilon)\|_{L^2(\omega_1)}.$$

The proof follows from the lines of the Caccioppoli inequality, Lemma 6.3.1, as

$$\|\nabla(\eta v_1^\varepsilon)\|_{L^2(\omega_1)} \leq \frac{\Lambda^{1/2}}{\lambda^{1/2} \tau} \|v_1^\varepsilon\|_{L^2(\omega_0)}.$$

We obtain

$$\|v_1^\varepsilon\|_{L^2(\omega)} \leq C_{\omega_1} \frac{\Lambda^{1/2}}{\lambda^{1/2} \tau} \|v_1^\varepsilon\|_{L^2(\omega_0)}.$$

The proof is similar for v_2^0 .

□

Strong Cauchy–Schwarz inequality

We recall the problems for the state variables: find $v_i \in H_D^1(\omega_i)$ such that

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla v_i) &= 0, & \text{in } \omega_i, \\ v_i &= \theta_i, & \text{on } \Gamma_i, \\ v_i &= 0, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n_i \cdot (a_i(x)\nabla v_i) &= 0, & \text{on } \partial\omega_i \cap \Gamma_N, \end{aligned} \tag{6.11}$$

where $a_1 = a_1^\varepsilon$ and $a_2 = a_2^0$.

Lemma 6.3.3 (Strong Cauchy–Schwarz). *Let $v_1^\varepsilon \in H_D^1(\omega_1)$ and $v_2^0 \in H_D^1(\omega_2)$ be the solutions of (6.11), for $i = 1, 2$, respectively. Then, there exist an $\varepsilon_0 > 0$ and a positive constant $C_s < 1$ such that for all $\varepsilon \leq \varepsilon_0$, it holds*

$$\int_{\omega_0} v_1^\varepsilon v_2^0 dx \leq C_s \|v_1^\varepsilon\|_{L^2(\omega_0)} \|v_2^0\|_{L^2(\omega_0)}.$$

Proof. We reason by contradiction. Suppose that there exist a sequence of $\{\varepsilon_n\}_{n \geq 1}$ that tends to zero such that

$$\int_{\omega_0} v_1^{\varepsilon_n} v_2^0 dx > C_n \|v_1^{\varepsilon_n}\|_{L^2(\omega_0)} \|v_2^0\|_{L^2(\omega_0)}, \quad \forall n \geq 1,$$

for any sequence $\{C_n\}_{n \geq 1}$ that tends to 1, with $C_n < 1$. Without loss of generality, we can normalize the vectors $v_1^{\varepsilon_n}$ and v_2 , and obtain

$$\|v_1^{\varepsilon_n}\|_{L^2(\omega_0)} = 1, \quad \|v_2^0\|_{L^2(\omega_0)} = 1 \quad \text{and} \quad (v_1^{\varepsilon_n}, v_2^0)_{L^2(\omega_0)} := \int_{\omega_0} v_1^{\varepsilon_n} v_2^0 dx \rightarrow 1.$$

As the sequence of tensors $\{a_1^{\varepsilon_n}\}_{n \geq 1} \in (L^\infty(\omega_1))^{d \times d}$ is bounded, and uniformly elliptic, by the H -convergence, there exists a subsequence of $\{\varepsilon_n\}_{n \geq 1}$ still denoted by $\{\varepsilon_n\}_{n \geq 1}$ and a tensor $a_1^0 \in (L^\infty(\omega_1))^{d \times d}$ bounded, and uniformly elliptic such that $\{a_1^{\varepsilon_n}\}_{n \geq 1}$ H -converges to a_1^0 . By definition of the H -convergence, the solution $v_1^{\varepsilon_n}$ of (6.11) — for the subsequence $\{\varepsilon_n\}$ — is such that

- i) $v_1^{\varepsilon_n} \rightharpoonup v_1^0$ in $H^1(\omega_1)$ and,
- ii) $a_1^{\varepsilon_n} \nabla v_1^{\varepsilon_n} \rightharpoonup a_1^0 \nabla v_1^0$ in $L^2(\omega_1)^d$,

where v_1^0 is the unique solution of

$$\begin{aligned} -\operatorname{div}(a_1^0(x)\nabla v_1^0) &= 0, & \text{in } \omega_1, \\ v_1^0 &= \theta_1, & \text{on } \Gamma_1, \\ v_1^0 &= 0, & \text{on } \partial\omega_1 \cap \Gamma_D, \\ n_1 \cdot (a_1^0(x)\nabla v_1^0) &= 0, & \text{on } \partial\omega_1 \cap \Gamma_N. \end{aligned}$$

As $H^1(\omega_1)$ is compactly embedded in $L^2(\omega_1)$, strong convergence in L^2 of $v_1^{\varepsilon_n}$ to v_1^0 , for a

subsequence of $\{\varepsilon_n\}_{n \geq 1}$, is achieved, i.e.,

$$v_1^{\varepsilon_n} \rightarrow v_1^0 \text{ in } L^2(\omega_1).$$

By the continuity of the norm, we have that

$$\lim_{n \rightarrow \infty} (v_1^{\varepsilon_n}, v_2)_{L^2(\omega_0)} = (v_1^0, v_2)_{L^2(\omega_0)}, \quad \|v_1^0\|_{L^2(\omega_0)} \leq 1 \quad \text{and} \quad (v_1^0, v_2)_{L^2(\omega_0)} = 1.$$

As

$$1 = (v_1^0, v_2)_{L^2(\omega_0)} \leq \|v_1^0\|_{L^2(\omega_0)} \|v_2\|_{L^2(\omega_0)} \leq 1,$$

we must have that $\|v_1^0\|_{L^2(\omega_0)} \|v_2\|_{L^2(\omega_0)} = 1$ and hence $\|v_1^0\|_{L^2(\omega_0)} = 1$. The previous inequalities become equalities; i.e.,

$$1 = (v_1^0, v_2)_{L^2(\omega_0)} = \|v_1^0\|_{L^2(\omega_0)} \|v_2\|_{L^2(\omega_0)}.$$

An equality in Cauchy–Schwarz is possible if and only if v_1^0 and v_2 are linearly dependent, that is there exist a constant $c > 0$ such that $v_1^0 = c v_2$ a.e. in ω_0 . As the norms of v_1^0 and v_2 are equal to 1, we can easily conclude that $c = \pm 1$ and that $v_1^0 = \pm v_2$ a.e. in ω_0 . Finally, as $(v_1^0, v_2)_{L^2(\omega_0)} = 1$ it holds that $v_1^0 = v_2$.

Both v_1^0 and v_2 are the solutions of a homogenized equation and are equal on the overlap, so we can combine them into a homogenized solution on the entire domain Ω . Further, the tensor a_2^0 and a_1^0 are equal in ω_0 . Indeed, let us continuously extend the tensors a_2^ε and a_1^ε to the domain Ω . The tensor a_1^ε H -converge to the tensor a_1^0 and the tensor a_2^ε H -converge to a_2^0 , in Ω . It holds that $a_2^\varepsilon = a_1^\varepsilon$ in ω_0 , and using the locality of H -convergence [90, 44], we can conclude that $a_2^0 = a_1^0$ in ω_2 . Thus they are equal in the overlap.

Let us split ω_0 into two disjoint sets ω_0^1 and ω_0^2 such that $\omega \subset \omega \cup \omega_0^1 \subset \omega \cup \omega_0$. As the solutions v_1^0 and v_2 are equal in ω_0 , we can construct a smooth function \bar{v} over Ω as

$$\bar{v}(x) = \begin{cases} v_1^0(x), & \text{if } x \in \omega \cup \omega_0^1, \\ v_2(x), & \text{if } x \in \omega_2 \setminus \omega_0^1. \end{cases}$$

The function \bar{v} is in $H_D^1(\Omega)$, has zero Neumann boundary condition on Γ_N , and satisfies

$$\int_{\Omega} \bar{a}^0 \nabla \bar{v} \cdot \nabla w \, dx = 0, \quad \forall w \in H_D^1(\Omega),$$

where the tensor \bar{a}^0 is given by

$$\bar{a}^0 = \begin{cases} a_1^0 & \text{in } \omega \cup \omega_0^1, \\ a_2^0 & \text{in } \omega_2 \setminus \omega_0^1. \end{cases}$$

Chapter 6. Optimization based coupling method

The solution \bar{v} must be zero everywhere in Ω , i.e., $\bar{v} \equiv 0$, which is a contradiction with $\|\bar{v}\|_{L^2(\omega_0)} = 1$. \square

Thanks to the strong Cauchy–Schwarz inequality, we are now able to prove that the form $\pi : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$ is an inner product over \mathcal{U} .

Lemma 6.3.4. *Let ε_0 be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and assume that $\varepsilon \leq \varepsilon_0$. Then, the form π defines an inner product on \mathcal{U} .*

Proof. The bilinearity, symmetry, and positivity are clear. We prove that the form is definite, i.e., $\pi((\mu_1, \mu_2), (\mu_1, \mu_2)) = 0$ if and only if $(\mu_1, \mu_2) = (0, 0)$.

On the one hand, if the virtual controls are zero traces or fluxes, the state functions v_1^ε and v_2^0 must be zero everywhere, as they are solutions of boundary value problems with zero right hand side and boundary conditions. Thus $\pi((\mu_1, \mu_2), (\mu_1, \mu_2)) = 0$.

On the other hand, using the strong Cauchy–Schwarz lemma, Lemma 6.3.3,

$$\begin{aligned} 0 &= \pi((\mu_1, \mu_2), (\mu_1, \mu_2)) = \|v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)\|_{L^2(\omega_0)}^2 \\ &= \|v_1^\varepsilon(\mu_1)\|_{L^2(\omega_0)}^2 + \|v_2^0(\mu_2)\|_{L^2(\omega_0)}^2 - 2 \int_{\omega_0} v_1^\varepsilon(\mu_1) v_2^0(\mu_2) dx \\ &\geq \|v_1^\varepsilon(\mu_1)\|_{L^2(\omega_0)}^2 + \|v_2^0(\mu_2)\|_{L^2(\omega_0)}^2 - 2C_S \|v_1^\varepsilon(\mu_1)\|_{L^2(\omega_0)} \|v_2^0(\mu_2)\|_{L^2(\omega_0)} \\ &\geq (1 - C_S) \left(\|v_1^\varepsilon(\mu_1)\|_{L^2(\omega_0)}^2 + \|v_2^0(\mu_2)\|_{L^2(\omega_0)}^2 \right). \end{aligned}$$

As $C_S < 1$, it holds that $\|v_1^\varepsilon(\mu_1)\|_{L^2(\omega_0)} = \|v_2^0(\mu_2)\|_{L^2(\omega_0)} = 0$ which implies that $v_1^\varepsilon = v_2^0 = 0$, a.e. in ω_0 . By Lemma 6.3.2, we have then that $\|v_1^\varepsilon(\mu_1)\|_{L^2(\omega)} = 0$ and $\|v_2^0(\mu_2)\|_{L^2(\Omega \setminus \omega_1)} = 0$, thus $v_i = 0$ a.e. in ω_i , for $i = 1, 2$. Then, we obtain, for $i = 1, 2$,

$$\|\mu_i\|_{H^{1/2}(\Gamma_i)} \leq C_1 \|v_i(\mu_i)\|_{H^1(\omega_i)} = 0,$$

where the constants depends on ω_i , and the trace operators $\gamma_i : H^{1/2}(\Gamma_i) \rightarrow H^1(\omega_i)$. Thus, $\mu_i = 0$ on Γ_i and the form π is an inner product on \mathcal{U} . \square

We can then define a norm on \mathcal{U} induced by the inner product π . For a pair $(\mu_1, \mu_2) \in \mathcal{U}$, we set

$$\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})} := \|v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)\|_{L^2(\omega_0)}. \quad (6.12)$$

The space \mathcal{U} might not be complete with respect to this norm, but we can construct a completion of \mathcal{U} , and solve the minimization problem in the completed space. Let us denote the completed control space by $\hat{\mathcal{U}}$. Using the Hahn–Banach theorem, the inner product π and the functional F can be continuously extended in a unique way on $\hat{\mathcal{U}}$ and we denote by $\hat{\pi}$ and

\hat{F} , these extensions. The form $\hat{\pi}$ is continuous, symmetric, and coercive in $\hat{\mathcal{U}}$. The existence and uniqueness of the optimal pair in $\hat{\mathcal{U}}$ is given in the next theorem.

Theorem 6.3.5. *The minimization problem (6.5) has a unique solution $(\theta_1, \theta_2) \in \hat{\mathcal{U}}$, that satisfies the Euler–Lagrange equation*

$$\hat{\pi}((\theta_1, \theta_2), (\mu_1, \mu_2)) = \hat{F}(\mu_1, \mu_2), \quad \forall (\mu_1, \mu_2) \in \hat{\mathcal{U}}, \quad (6.13)$$

where $\hat{\pi}$ and \hat{F} are the continuous extensions of π and F given by (6.9) and (6.10).

Proof. The existence and uniqueness of $(\theta_1, \theta_2) \in \hat{\mathcal{U}}$ follows from [85, Theorem I. 1.1], as the form $\hat{\pi}$ is symmetric, continuous, and coercive, and \hat{F} is continuous. \square

The optimal pair $(\theta_1, \theta_2) \in \hat{\mathcal{U}}$ minimizes the cost function, but in general there exist no functions $u_i \in H^1(\omega_i)$ that satisfy (6.4). However, there exists an embedding $\sigma : \mathcal{U} \rightarrow \hat{\mathcal{U}}$ such that $\sigma(\mathcal{U})$ is dense in $\hat{\mathcal{U}}$. Further, we can identify \mathcal{U} with $\sigma(\mathcal{U})$ and conclude that (θ_1, θ_2) is the limit of a sequence $(\theta_{1n}, \theta_{2n})_{n \in \mathbb{N}}$ with $u_i(\theta_{in}) \in H^1(\omega_i)$ satisfying (6.4). In the sequel, for simplicity, we assume that the optimal pair is in \mathcal{U} and hence $u_i(\theta_i) \in H^1(\omega_i)$, for $i = 1, 2$ (we then also have $v_i(\theta_i) \in H^1(\omega_i)$).

6.4 Optimality systems

The solution of the minimization problem (6.5) is found by solving an optimality system; this is the system used in the fully discrete coupling method. We emphasize that the existence and uniqueness of the optimal controls have already been established as the minimization problem (6.5) is well-posed (see Section 6.3). We give here the optimality system as an alternative approach to find the optimal controls (and the state variables).

In this section, we will derive the optimality system following two approaches. The first one is derived from the Euler–Lagrange equation (6.13) and is referred to as "à la Lions". The second approach is derived from a Lagrangian functional by taking the Gâteaux derivatives. The existence and uniqueness of the optimal controls can be proved if the optimality system is well-posed.

We recall that $\theta_i \in \mathcal{U}$ is a Dirichlet boundary data on Γ_i , and that the state variable $v_i \in H_D^1(\omega_i)$ satisfies

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla v_i) &= 0, & \text{in } \omega_i, \\ v_i &= \theta_i, & \text{on } \Gamma_i, \\ v_i &= 0, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n_i \cdot (a_i(x)\nabla v_i) &= 0, & \text{on } \partial\omega_i \cap \Gamma_N. \end{aligned}$$

6.4.1 À la Lions

The solution of the minimization problem can be found by solving an optimality problem derived from the Euler–Lagrange equation (6.13); i.e.,

$$\pi((\theta_1, \theta_2), (\mu_1, \mu_2)) = F(\mu_1, \mu_2), \quad \forall (\mu_1, \mu_2) \in \mathcal{U}.$$

For the minimization with the cost function (6.5), the optimality system reads: find $(v_1, v_2) \in H_D^1(\omega_1) \times H_D^1(\omega_2)$, $(\theta_1, \theta_2) \in \mathcal{U}$, and $(\lambda_1, \lambda_2) \in H_0^1(\omega_1) \times H_0^1(\omega_2)$ such that

$$B_i(v_i(\theta_i), w) = 0, \text{ in } \omega_i, \quad v_i = \theta_i, \text{ on } \Gamma_i, \quad \forall w \in H_0^1(\omega_i), \quad (6.14)$$

$$B_i(w, \lambda_i) = (-1)^{i+1} \int_{\omega_0} ((v_1^\varepsilon(\theta_1) + u_{1,0}^\varepsilon) - (v_2^0(\theta_2) + u_{2,0}^0)) w dx, \quad \forall w \in H_0^1(\omega_i), \quad (6.15)$$

$$\begin{aligned} & \int_{\omega_0} ((v_1^\varepsilon(\theta_1) + u_{1,0}^\varepsilon) - (v_2^0(\theta_2) + u_{2,0}^0)) (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) dx \\ &= - \int_{\Gamma_1} a_1 \nabla \lambda_1 n_1 \mu_1 ds - \int_{\Gamma_2} a_2 \nabla \lambda_2 n_2 \mu_2 ds = 0, \quad \forall (\mu_1, \mu_2) \in \mathcal{U}. \end{aligned} \quad (6.16)$$

Equation (6.14) is the state equation for v_i , equation (6.15) is the adjoint problem corresponding to (6.14), and equation (6.16) is derived from the adjoint problem and using integration by parts. We recognize the Euler–Lagrange equation. For the derivation of this system, we refer to [85, 49].

A sufficient and necessary for (θ_1, θ_2) to be optimal controls is that the set of equations (6.14), (6.15), and (6.16) form a well-posed system [85, Chapter 2, Theorem 1.4].

6.4.2 Lagrangian functional

Solving optimal control problems can be done using Lagrange multipliers. Assuming that the admissible set \mathcal{U}_{ad} is equal to \mathcal{U} , the solution of the optimal control problem can be obtained as an unconstrained critical point of a Lagrange functional.

Let λ_i , $i = 1, 2$, be Lagrange multipliers associated to the constraints in ω_i . Here one could consider $(v_i, \theta_i, \lambda_i) \in H_D^1 \times \mathcal{U} \times H_0^1(\omega_i)$ for the unknowns or $(v_i, \lambda_i) \in H_D^1(\omega_i) \times H_0^1(\omega_i)$. We will derive the optimality system with the unknown (v_i, λ_i) , and refer to [101] for the derivation of the optimality system for $(v_i, \theta_i, \lambda_i)$. The solution $(v_1^\varepsilon, \lambda_1, v_2^0, \lambda_2)$ is the critical point of the Lagrangian functional \mathcal{L} given by

$$\begin{aligned} \mathcal{L}(v_1^\varepsilon, \lambda_1, v_2^0, \lambda_2) &= \frac{1}{2} \|v_1^\varepsilon - v_2^0\|_{L^2(\omega_0)}^2 + \frac{1}{2} \|u_{1,0}^\varepsilon - u_{2,0}^0\|_{L^2(\omega_0)}^2 + \int_{\omega_0} (v_1^\varepsilon - v_2^0)(u_{1,0}^\varepsilon - u_{2,0}^0) dx \\ &\quad + \langle f + \operatorname{div}(a_1^\varepsilon(x) \nabla (v_1^\varepsilon + u_{1,0}^\varepsilon)), \lambda_1 \rangle_{H^{-1}, H^1} \\ &\quad + \langle f + \operatorname{div}(a_2^0(x) \nabla (v_2^0 + u_{2,0}^0)), \lambda_2 \rangle_{H^{-1}, H^1}, \end{aligned}$$

with $v_i \in H_D^1(\omega_i)$ and $\lambda_i \in H_0^1(\omega_i)$ with $n_i \cdot (a_i \nabla \lambda_i) = 0$ on $\partial\omega_i \cap \Gamma_N$, for $i = 1, 2$. Computing the

Gâteaux derivatives for each of the unknowns leads to the optimality system.

$$\int_{\omega_0} (v_1^\varepsilon - v_2^0) w_1 dx - \int_{\omega_1} a_1^\varepsilon(x) \nabla w_1 \cdot \nabla \lambda_1 dx = - \int_{\omega_0} (u_{1,0}^\varepsilon - u_{2,0}^0) w_1 dx, \quad \forall w_1 \in H_D^1(\omega_1), \quad (6.17)$$

$$- \int_{\omega_0} (v_1^\varepsilon - v_2^0) w_2 dx - \int_{\omega_2} a_2^0(x) \nabla w_2 \cdot \nabla \lambda_2 dx = \int_{\omega_0} (u_{1,0}^\varepsilon - u_{2,0}^0) w_2 dx, \quad \forall w_2 \in H_D^1(\omega_2), \quad (6.18)$$

$$\int_{\omega_1} a_1^\varepsilon(x) \nabla v_1^\varepsilon \cdot \nabla \xi_1 dx = F_1(\xi_1), \quad \forall \xi_1 \in H_0^1(\omega_1), \quad (6.19)$$

$$\int_{\omega_2} a_2^0(x) \nabla v_2^0 \cdot \nabla \xi_2 dx = F_2(\xi_2), \quad \forall \xi_2 \in H_0^1(\omega_2), \quad (6.20)$$

The boundary conditions (θ_1, θ_2) are hidden in the state functions (v_1^ε, v_2^0) , and if the system is well-posed, it admits a unique solution (θ_1, θ_2) which are the optimal controls.

One notices that the equations (6.17) and (6.18) are the adjoint equations (6.15), and that equations (6.19) and (6.20) are the state equations (6.14).

Remark 6.4.1. When considering $(v_i, \theta_i, \lambda_i)$ as an unknown, the optimality system has two more equations, obtained by the Gâteaux derivatives for θ_1 and θ_2 . These two equations are in fact given in the Lions' approach in equation (6.16).

The optimality system can be written in a matrix form as a saddle point problem: find $U = (v_1^\varepsilon, v_2^0, \lambda_1, \lambda_2)^\top$ such that

$$\begin{pmatrix} M & -B^\top \\ B & 0 \end{pmatrix} U = G, \quad (6.21)$$

where

$$M(\{v_1^\varepsilon, v_2^0\}, \{w_1^\varepsilon, w_2^0\}) = \begin{pmatrix} \int_{\omega_0} v_1^\varepsilon w_1^\varepsilon dx & - \int_{\omega_0} v_2^0 w_1^\varepsilon dx \\ - \int_{\omega_0} v_1^\varepsilon w_2^0 dx & \int_{\omega_0} v_2^0 w_2^0 dx \end{pmatrix},$$

$$B(\{v_1^\varepsilon, v_2^0\}, \{\lambda_1, \lambda_2\}) = \begin{pmatrix} B_1(v_1^\varepsilon, \lambda_1) & 0 \\ 0 & B_2(v_2^0, \lambda_2) \end{pmatrix}.$$

The operator M is defined on the space $(H_D^1(\omega_1) \times H_D^1(\omega_2))^2$, and the operator B on the space $H_D^1(\omega_1) \times H_D^1(\omega_2) \times H_0^1(\omega_1) \times H_0^1(\omega_2)$.

Well-posedness of the optimality system

To prove the well-posedness of the saddle point problem (6.21), one needs to show that the form M is coercive and that the bilinear form B is bounded and satisfies an inf-sup condition.

We follow [34] and write the forms M and B as

$$\begin{aligned} M(\{v_1^\varepsilon, v_2^0\}, \{w_1^\varepsilon, w_2^0\}) &= \int_{\omega_0} (v_1^\varepsilon - v_2^0)(w_1^\varepsilon - w_2^0) dx \\ B(\{v_1^\varepsilon, v_2^0\}, \{\lambda_1, \lambda_2\}) &= B_1(v_1^\varepsilon, \lambda_1) + B_2(v_2^0, \lambda_2). \end{aligned}$$

One has to prove that

1. M is coercive on $H_D^1(\omega_1) \times H_D^1(\omega_2)$;
2. B is bounded on $H_D^1(\omega_1) \times H_D^1(\omega_2) \times H_0^1(\omega_1) \times H_0^1(\omega_2)$ and satisfies an inf-sup condition; for all $(\xi_1, \xi_2) \in H_0^1(\omega_1) \times H_0^1(\omega_2)$

$$\sup_{w_1^\varepsilon, w_2^0} \frac{B(\{w_1^\varepsilon, w_2^0\}, \{\xi_1, \xi_2\})}{\|w_1^\varepsilon\|_{H^1(\omega_1)} + \|w_2^0\|_{H^1(\omega_2)}} \geq C(\|\xi_1\|_{H^1(\omega_1)} + \|\xi_2\|_{H^1(\omega_2)}).$$

As the state v_i and the controls θ_i are linked through an elliptic boundary value problem, one could consider the unknown $U = (\theta_1, \theta_2, \lambda_1, \lambda_2)^\top \in \mathcal{U}_1^D \times \mathcal{U}_2^D \times H_0^1(\omega_1) \times H_0^1(\omega_2)$ instead of $(v_1^\varepsilon, v_2^0, \lambda_1, \lambda_2)$. The optimality system remains unchanged; the forms M and B are given by

$$\begin{aligned} M(\{\theta_1, \theta_2\}, \{\mu_1, \mu_2\}) &= \int_{\omega_0} (v_1^\varepsilon(\theta_1) - v_2^0(\theta_2))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) dx \\ B(\{\theta_1, \theta_2\}, \{\lambda_1, \lambda_2\}) &= B_1(v_1^\varepsilon(\theta_1), \lambda_1) + B_2(v_2^0(\theta_2), \lambda_2). \end{aligned}$$

and prove

1. M is coercive on \mathcal{U} ;
2. B is bounded on $\mathcal{U}_1^D \times \mathcal{U}_2^D \times H_0^1(\omega_1) \times H_0^1(\omega_2)$ and satisfies an inf-sup condition; for all $(\xi_1, \xi_2) \in H_0^1(\omega_1) \times H_0^1(\omega_2)$

$$\sup_{\mu_1, \mu_2} \frac{B(\{v_1^\varepsilon(\mu_1), v_2^0(\mu_2)\}, \{\xi_1, \xi_2\})}{\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}} \geq C(\|\xi_1\|_{H^1(\omega_1)} + \|\xi_2\|_{H^1(\omega_2)}).$$

In both approaches, we end up with the need to bound an H^1 norm (over ω_i or ω_0) by an L^2 norm over ω_i . The well-posedness of the optimality system derived from the Lagrange functional remains an open question, as the constant can not be determined. However, this is not an issue as we already proved that the minimization problem is well-posed (see Section 6.3). Further we note that in the discrete coupling method the well-posedness of the discrete optimality system can be successfully proved.

6.4.3 Transposition method

When the regularity on the controls (θ_1, θ_2) is lessened, the existence and uniqueness of a (weak) solution v_i of problem (6.6) is not guaranteed any more. In order to construct and optimality system, one needs to find a well-posed weak formulation of equation (6.6).

Previously, we made the strong assumption that the optimal controls (θ_1, θ_2) are in the space \mathcal{U} , which ensures that the states v_i are in $H^1(\omega_i)$, for $i = 1, 2$. The bilinear form π is an inner product on \mathcal{U} and using a completion argument on the space \mathcal{U} , we obtained that the completed space \mathcal{U} equipped with the inner product π is an Hilbert space; the control space is denoted by $\hat{\mathcal{U}}$ and it holds $\mathcal{U} \subset \hat{\mathcal{U}}$. Lion's theory assures the existence and uniqueness of the optimal controls in $\hat{\mathcal{U}}$ (see Section 6.2). However, if $\theta_i \in \hat{\mathcal{U}}$, we lose the guarantee that $v_i \in H^1(\Omega)$, for $i = 1, 2$.

For simplicity, we will not use the splitting $u_i = u_{i,0} + v_i$ and derive the optimality system using the transposition method [86] for the unknown u_i .

We multiply problem (6.4) with a test function w in $H^1(\omega_i)$ vanishing on $\partial\omega_i \cap \Gamma_D$ and Γ_i , and using Green formula, it holds

$$\int_{\omega_i} a_i(x) \nabla u_i \cdot \nabla w dx = \int_{\omega_i} f w dx + \int_{\partial\omega_i} a_i(x) \nabla u_i n_i w ds,$$

where the boundary integral regroupes the Dirichlet and Neumann data; i.e,

$$\int_{\partial\omega_i} a_i(x) \nabla u_i n_i w ds = \int_{\Gamma_i} a_i(x) \nabla u_i n_i w ds + \int_{\partial\omega_i \cap \Gamma_D} a_i(x) \nabla u_i n_i w ds + \int_{\partial\omega_i \cap \Gamma_N} g_N w ds.$$

For each smooth test function w vanishing on $\partial\omega_i \cap \Gamma_D$ and Γ_i , we obtain

$$\int_{\omega_i} a_i(x) \nabla u_i \cdot \nabla w dx = \int_{\omega_i} f w dx + \int_{\partial\omega_i \cap \Gamma_N} g_N w ds. \quad (6.22)$$

Considering now a lifting $R_{g_D}^i \in H^1(\omega_i)$ of the Dirichlet data g_D , it hold $u_i = \bar{u}_i + R_{g_D}^i$, where $\bar{u}_i = 0$ on $\partial\omega_i \cap \Gamma_D$. For simplicity of notations, let u_i denote \bar{u}_i . We can reformulate equation (6.22) as

$$\int_{\omega_i} a_i(x) \nabla u_i \cdot \nabla w dx = \int_{\omega_i} f w dx - \int_{\omega_i} a_i(x) \nabla R_{g_D}^i \cdot \nabla w dx + \int_{\partial\omega_i \cap \Gamma_N} g_N w ds. \quad (6.23)$$

For standard weak solution of (6.23), one needs the existence of a continuous lifting of the boundary controls θ_i . The lifting should live in $H^1(\omega_i)$, however, such lifting might not exist due the regularity of $\theta_i \in \hat{\mathcal{U}}$. To find a weak solution of (6.23), we follow the method of transposition [86] Let us integrate by part equation (6.23), this gives

$$\begin{aligned} - \int_{\omega_i} u_i \nabla \cdot (a_i(x) \nabla w) dx + \int_{\partial\omega_i} a_i(x) \nabla w n_i u_i ds &= \int_{\omega_i} f w dx - \int_{\omega_i} a_i(x) \nabla R_{g_D}^i \cdot \nabla w dx \\ &+ \int_{\partial\omega_i \cap \Gamma_N} g_N w ds, \end{aligned}$$

with

$$\begin{aligned} \int_{\partial\omega_i} a_i(x) \nabla w n_i u_i ds &= \int_{\partial\omega_i \cap \Gamma_N} a_i(x) \nabla w n_i u_i ds + \int_{\partial\omega_i \cap \Gamma_D} a_i(x) \nabla w n_i g_D ds \\ &\quad + \int_{\Gamma_i} a_i(x) \nabla w n_i \theta_i ds. \end{aligned}$$

This gives a *very* weak formulation: find $u_i \in L^2_{\partial\omega_i \cap \Gamma_D}(\omega_i)$ such that

$$\begin{aligned} - \int_{\omega_i} u_i \nabla \cdot (a_i(x) \nabla w) dx &= \int_{\omega_i} f w dx - \int_{\omega_i} a_i(x) \nabla R_{g_D}^i \cdot \nabla w dx + \int_{\partial\omega_i \cap \Gamma_N} g_N w ds \\ &\quad - \int_{\partial\omega_i \cap \Gamma_N} a_i(x) \nabla w n_i u_i ds - \int_{\Gamma_i} a_i(x) \nabla w n_i \theta_i ds, \end{aligned}$$

for any test function $w \in H^2(\omega_i)$ with $w = 0$ on $\partial\omega_i \cap \Gamma_D$ and Γ_i . We can further assume that v has homogeneous Neumann boundary condition on $\partial\omega_i \cap \Gamma_N$, i.e., $n_1 \cdot (a_i(x) \nabla w) = 0$. We remark that the test function w plays the role of the adjoint.

Let λ_i be Lagrange multipliers associated to the constraints in ω_i , $i = 1, 2$, and look for the critical point of the Lagrangian functional

$$\begin{aligned} \mathcal{L}(u_1^\varepsilon, \theta_1, \lambda_1, u_2^0, \theta_2, \lambda_2) &= \frac{1}{2} \|u_1^\varepsilon - u_2^0\|_{L^2(\omega_0)}^2 + \langle f + \operatorname{div}(a_1^\varepsilon(x) \nabla u_1^\varepsilon), \lambda_1 \rangle_{H^{-1}, H^1} \\ &\quad + \langle f + \operatorname{div}(a_2^0(x) \nabla u_2^0), \lambda_2 \rangle_{H^{-1}, H^1}, \end{aligned}$$

with $u_i \in H^1(\omega_i)$, $\theta_i \in \hat{\mathcal{U}}_i$, and $\lambda_i \in H^2(\omega_i)$ with $\lambda_i = 0$ on $\partial\omega_i \cap \Gamma_D$ and Γ_i , and $n_i \cdot (a_i \nabla \lambda_i) = 0$ on $\partial\omega_i \cap \Gamma_N$, for $i = 1, 2$.

The Lagrangian can be reformulated using the transposition method exposed above; i.e.,

$$\begin{aligned} \mathcal{L}(u_1^\varepsilon, \theta_1, \lambda_1, u_2^0, \theta_2, \lambda_2) &= \frac{1}{2} \|u_1^\varepsilon - u_2^0\|_{L^2(\omega_0)}^2 + \int_{\omega_1} f \lambda_1 dx + \int_{\omega_2} f \lambda_2 dx + \int_{\partial\omega_1 \cap \Gamma_N} g_N \lambda_1 ds \\ &\quad + \int_{\partial\omega_2 \cap \Gamma_N} g_N \lambda_2 ds - \int_{\Gamma_1} a^\varepsilon(x) \nabla \lambda_1 n_1 \theta_1 ds - \int_{\Gamma_2} a^0(x) \nabla \lambda_2 n_2 \theta_2 ds \\ &\quad + \int_{\omega_1} u_1^\varepsilon \nabla \cdot (a_1^\varepsilon(x) \nabla \lambda_1) dx + \int_{\omega_2} u_2^0 \nabla \cdot (a_2^0(x) \nabla \lambda_2) dx \\ &\quad - \int_{\omega_1} a_1^\varepsilon(x) \nabla R_{g_D}^1 \cdot \nabla \lambda_1 dx - \int_{\omega_2} a_2^0(x) \nabla R_{g_D}^2 \cdot \nabla \lambda_2 dx. \end{aligned}$$

Let us denote by Λ^i the space of the test functions in the transposition method, i.e.

$$\Lambda^i = \{v \in H^2(\omega_i) \mid v = 0, \text{ on } \Gamma_i \text{ and } \partial\omega_i \cap \Gamma_D, \text{ and } n_i(a_i \nabla v) = 0\},$$

Computing the Gâteaux derivatives for each of the unknowns leads to the optimality system;

$$\begin{aligned}
 & \int_{\omega_0} (u_1^\varepsilon - u_2^0) w_1 dx + \int_{\omega_1} w_1 \nabla \cdot (a_1^\varepsilon(x) \nabla \lambda_1) dx = 0, & \forall w_1 \in H_D^1(\omega_1), \\
 & - \int_{\omega_0} (u_1^\varepsilon - u_2^0) w_2 dx + \int_{\omega_2} w_2 \nabla \cdot (a_2^0(x) \nabla \lambda_2) dx = 0, & \forall w_2 \in H_D^1(\omega_2), \\
 & \int_{\omega_1} u_1^\varepsilon \nabla \cdot (a_1^\varepsilon(x) \nabla \xi_1) dx - \int_{\Gamma_1} a_1^\varepsilon(x) \nabla \xi_1 n_1 \theta_1 ds = F_1(\xi_1), & \forall \xi_1 \in \Lambda^1, \\
 & \int_{\omega_2} u_2^0 \nabla \cdot (a_2^0(x) \nabla \xi_2) dx - \int_{\Gamma_2} a_2^0(x) \nabla \xi_2 n_2 \theta_2 ds = F_2(\xi_2), & \forall \xi_2 \in \Lambda^2, \\
 & - \int_{\Gamma_1} a_1^\varepsilon(x) \nabla \lambda_1 n_1 \mu_1 ds = 0, & \forall \mu_1 \in \hat{\mathcal{U}}_1, \\
 & - \int_{\Gamma_2} a_2^0(x) \nabla \lambda_2 n_2 \mu_2 ds = 0, & \forall \mu_2 \in \hat{\mathcal{U}}_2,
 \end{aligned}$$

where, for $i = 1, 2$,

$$F_i(\xi_i) = \int_{\omega_i} f \xi_i dx + \int_{\partial \omega_i \cap \Gamma_N} g_N \xi_i ds - \int_{\omega_i} a_i(x) \nabla R_{g_D}^i \cdot \nabla \xi_i dx.$$

6.5 A priori error analysis

In this section, we give an a priori error analysis of the optimization based method. The analysis is separated into a fine and a coarse scale error estimate. The solution of the minimization problem with constraints (6.4) gives us a fine scale solution in ω_1 and a coarse scale solution in ω_2 . Looking at the error between the solution of the coupling and the exact fine scale solution u^ε , the solution of problem (6.1), on either ω_1 or ω_2 , leads to the estimation of terms on the boundary Γ_1 or Γ_2 , respectively. In order to avoid such additional error terms, we introduce an intermediate domain ω^+ with $\omega \subset \omega^+ \subset \omega_1$; see Figure 6.3. Then given $u_1^\varepsilon(\theta_1)$ and $u_2^0(\theta_2)$, the solutions of the optimization based coupling method, we define

$$\bar{u}^\varepsilon = \begin{cases} u_1^\varepsilon(\theta_1), & \text{in } \omega^+, \\ u_2^{rec}(\theta_2), & \text{in } \Omega \setminus \omega^+, \end{cases} \quad (6.24)$$

where u_2^{rec} stands for a correction to the homogenized solution $u_2^0(\theta_2)$ given below. The main convergence results are

$$\begin{aligned}
 \|u^\varepsilon - \bar{u}^\varepsilon\|_{H^1(\omega^+)} &\leq C\varepsilon, \\
 \|u^\varepsilon - \bar{u}^\varepsilon\|_{H^1(\Omega \setminus \omega^+)} &\leq C\varepsilon^{1/2},
 \end{aligned}$$

where the constants depend on the width of ω^+ and the ellipticity constants of the tensor a^ε , see equation (6.2). For the analysis, we assume that the tensor is locally periodic in Y , and consider the classical locally periodic correctors χ^j the solutions of (2.17), but other

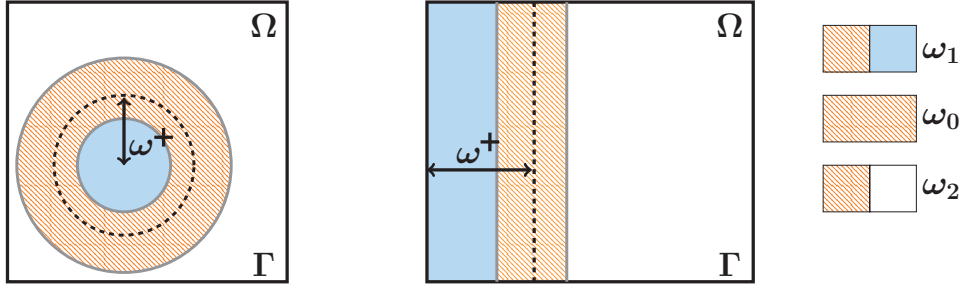


Figure 6.3 – Illustration of two possible domains decompositions considered in the a priori error analysis, with interior domains (left) and with shared boundaries (right).

post-processing procedure could be used. The correction $u_2^{rec}(\theta_2)$ is given by

$$u_2^{rec}(x) = u_2^0(x) + \varepsilon \sum_{j=1}^d \chi^j(x, x/\varepsilon) \frac{\partial u_2^0(x)}{\partial x_j}, \quad x \in \Omega \setminus \omega^+, \quad (6.25)$$

where $u_2^0 = u_2^0(\theta_2)$. We sometimes use $u_2^0(\theta_2)$ and $u_2^{rec}(\theta_2)$ to emphasize the dependency on θ_2 . We will however avoid the heavy notation $u_2^0(\theta_2)(x)$ (or $u_2^{rec}(\theta_2)(x)$) and drop the dependency on θ_2 when writing such maps as functions of x .

6.5.1 A priori error estimates to the fine scale solver in ω^+ .

The coupled solution restricted to the subregion ω^+ is given by the fine scale solution $u_1^\varepsilon(\theta_1)$, hence the error becomes

$$\|u^\varepsilon - \bar{u}^\varepsilon\|_{H^1(\omega^+)} = \|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{H^1(\omega^+)}.$$

Let τ denote the width of the overlap ω_0 and recall that the heterogeneous tensor a_2^ε satisfies the ellipticity condition (6.2). Further, we denote by τ^+ the distance between $\partial\omega^+$ and $\partial\omega$; it holds $\tau^+ < \tau$. Moreover, we suppose that there exists $\varepsilon_0 > 0$, such that the strong Cauchy–Schwarz lemma, Lemma 6.3.3, holds, for all $\varepsilon \leq \varepsilon_0$.

Let $\gamma_i : H^1(\omega_i) \rightarrow H^{1/2}(\Gamma_i)$, $i = 1, 2$, be trace operators and consider the solution u^ε restricted to the domain ω_2 ,

$$\begin{aligned} -\operatorname{div}(a_2^\varepsilon(x)\nabla u^\varepsilon) &= f, & \text{in } \omega_2, \\ u^\varepsilon &= \gamma_2(u^\varepsilon), & \text{on } \Gamma_2, \\ u^\varepsilon &= g_D, & \text{on } \partial\omega_2 \cap \Gamma_D, \\ n_2 \cdot (a_2^\varepsilon(x)\nabla u^\varepsilon) &= g_N, & \text{on } \partial\omega_2 \cap \Gamma_N. \end{aligned}$$

Further, for a fixed $\varepsilon \leq \varepsilon_0$, we introduce $u^0 \in H^1(\omega_2)$, the homogenized solution of

$$\begin{aligned} -\operatorname{div}(a_2^0(x)\nabla u^0) &= f, & \text{in } \omega_2, \\ u^0 &= \gamma_2(u^\varepsilon), & \text{on } \Gamma_2, \\ u^0 &= g_D, & \text{on } \partial\omega_2 \cap \Gamma_D, \\ n_2 \cdot (a_2^0(x)\nabla u^0) &= g_N, & \text{on } \partial\omega_2 \cap \Gamma_N. \end{aligned} \quad (6.26)$$

We assume that strong convergence in the L^2 norm is available [80, Sect. 1.4]; i.e.,

$$\|u^\varepsilon - u^0\|_{L^2(\omega_2)} \leq C\varepsilon. \quad (6.27)$$

We follow the framework introduced in [98] and define an operator $P : \mathcal{U} \rightarrow H^1(\omega_1) \times H^1(\Omega \setminus \omega_1)$ by

$$(\mu_1, \mu_2) \mapsto P(\mu_1, \mu_2) = \begin{cases} u_{1,0}^\varepsilon + v_1^\varepsilon(\mu_1), & \text{in } \omega_1, \\ u_{2,0}^0 + v_2^0(\mu_2), & \text{in } \Omega \setminus \omega_1, \end{cases}$$

where v_i are solutions of (6.6), for $i = 1, 2$. We note that for the traces $(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon))$ of the exact solution u^ε , we obtain

$$P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) = \begin{cases} u^\varepsilon, & \text{in } \omega_1, \\ u^0, & \text{in } \Omega \setminus \omega_1. \end{cases}$$

The operator P can be split into $P(\mu_1, \mu_2) = U_0 + Q(\mu_1, \mu_2)$, for $(\mu_1, \mu_2) \in \mathcal{U}$, where we define

$$U_0 = \begin{cases} u_{1,0}^\varepsilon, & \text{in } \omega_1, \\ u_{2,0}^0, & \text{in } \Omega \setminus \omega_1, \end{cases} \quad \text{and} \quad Q(\mu_1, \mu_2) = \begin{cases} v_1^\varepsilon(\mu_1), & \text{in } \omega_1, \\ v_2^0(\mu_2), & \text{in } \Omega \setminus \omega_1. \end{cases} \quad (6.28)$$

Theorem 6.5.1. *Let u^ε and u^0 be the solution of (6.1) and (6.26), respectively, and let \tilde{u}^ε be given by (6.24). Suppose that u^0 and χ^j are regular enough so that (6.27) holds. Let ε_0 be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and assume that $\varepsilon \leq \varepsilon_0$. Then, we have*

$$\|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{H^1(\omega^+)} \leq C\varepsilon,$$

where the constant C depends on τ , τ^+ , λ , Λ , and on the domains ω_1 and ω_2 .

Proof. The difference $u^\varepsilon - u_1^\varepsilon(\theta_1)$ is a^ε -harmonic in ω_1 and Caccioppoli inequality, Lemma 5.1.1, can be applied; that is

$$\|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{H^1(\omega^+)} \leq \frac{C}{\tau - \tau^+} \|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{L^2(\omega_1)},$$

where the constant C depends on the ellipticity constants of the tensor a^ε . Let us focus on the

Chapter 6. Optimization based coupling method

L^2 norm; recalling that $u_1^\varepsilon(\theta_1) = P(\theta_1, \theta_2)$, it holds that

$$\begin{aligned} \|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{L^2(\omega_1)} &= \|u^\varepsilon - P(\theta_1, \theta_2)\|_{L^2(\omega_1)} \\ &\leq \|u^\varepsilon - P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon))\|_{L^2(\omega_1)} + \|P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - P(\theta_1, \theta_2)\|_{L^2(\omega_1)}. \end{aligned}$$

By the definitions of P and u^ε , the first L^2 error is zero and it remains to bound the second error. Using the splitting of P into U_0 and Q , defined in equation (6.28), it holds

$$\begin{aligned} \|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{L^2(\omega_1)} &\leq \|P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - P(\theta_1, \theta_2)\|_{L^2(\omega_1)} \\ &= \|U_0 - Q(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - U_0 + Q(\theta_1, \theta_2)\|_{L^2(\omega_1)} \\ &\leq \|Q\| \|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})}, \end{aligned}$$

where the norm $\|\cdot\|_{L^*(\mathcal{U})}$ is induced by the inner product π and is defined in (6.12). Using Lemmas 6.5.2 and 6.5.3 given below proves the result. \square

Lemma 6.5.2. *Let u^ε and u^0 solve (6.1) and (6.26) respectively, and let $(\theta_1, \theta_2) \in \mathcal{U}$ be the optimal virtual controls. Then*

$$\|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})} \leq \|u^\varepsilon - u^0\|_{L^2(\omega_0)}.$$

Proof. By definition, we have

$$\begin{aligned} &\|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})} = \\ &\sup_{(\mu_1, \mu_2) \in \mathcal{U}} \frac{|\pi((\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)), (\mu_1, \mu_2)) - \pi((\theta_1, \theta_2), (\mu_1, \mu_2)))|}{\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}}. \end{aligned}$$

We look at the numerator. As the pair (θ_1, θ_2) minimizes the cost function J , the Euler–Lagrange formulation (6.13) holds and

$$\begin{aligned} &\pi((\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)), (\mu_1, \mu_2)) - \pi((\theta_1, \theta_2), (\mu_1, \mu_2)) = \\ &= \int_{\omega_0} (v_1^\varepsilon(\gamma_1(u^\varepsilon)) - v_2^0(\gamma_2(u^\varepsilon)))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) dx \\ &\quad + \int_{\omega_0} (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2))(u_{1,0}^\varepsilon - u_{2,0}^0) dx \\ &= \int_{\omega_0} ((v_1^\varepsilon(\gamma_1(u^\varepsilon)) + u_{1,0}^\varepsilon) - (v_2^0(\gamma_2(u^\varepsilon)) + u_{2,0}^0))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) dx \\ &= \int_{\omega_0} (u^\varepsilon - u^0)(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) dx \leq \|u^\varepsilon - u^0\|_{L^2(\omega_0)} \|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}. \end{aligned}$$

The result follows. \square

To complete the a priori error analysis in the continuous case, we need to bound the norm of the operator Q .

Lemma 6.5.3. *Let ε_0 be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and assume that $\varepsilon \leq \varepsilon_0$. The operator Q , defined by (6.28), is bounded from $L^2(\Omega)$ to $L^*(\mathcal{U})$; i.e.,*

$$\|Q\| \leq C,$$

where the constant C depends on ω_1 , ω_2 , τ , and the strong Cauchy–Schwarz constant, see Lemma 6.3.3.

Proof. By definition, the norm of the operator Q is given by

$$\|Q\| := \sup_{(\mu_1, \mu_2) \in \mathcal{U}} \frac{\|Q(\mu_1, \mu_2)\|_{L^2(\Omega)}}{\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}}.$$

For $(\mu_1, \mu_2) \in \mathcal{U}$, we show the existence of a positive constant C such that

$$\|Q(\mu_1, \mu_2)\|_{L^2(\Omega)}^2 \leq C \|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}^2.$$

For simplicity, we set $v_i = v_i(\mu_i)$, $i = 1, 2$. Using Lemma 6.3.2, we have

$$\begin{aligned} \|Q(\mu_1, \mu_2)\|_{L^2(\Omega)}^2 &= \|v_1^\varepsilon\|_{L^2(\omega_1)}^2 + \|v_2^0\|_{L^2(\Omega \setminus \omega_1)}^2 \\ &\leq \frac{C(\omega_1; \omega_2)}{\tau^2} \left(\|v_1^\varepsilon\|_{L^2(\omega_0)}^2 + \|v_2^0\|_{L^2(\omega_0)}^2 \right). \end{aligned}$$

Next, using the strong Cauchy–Schwarz lemma, Lemma 6.3.3, we obtain

$$\begin{aligned} \|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}^2 &= \|v_1^\varepsilon - v_2^0\|_{L^2(\omega_0)}^2 = \|v_1^\varepsilon\|_{L^2(\omega_0)}^2 + \|v_2^0\|_{L^2(\omega_0)}^2 - 2 \int_{\omega_0} v_1^\varepsilon v_2^0 dx \\ &\geq \|v_1^\varepsilon\|_{L^2(\omega_0)}^2 + \|v_2^0\|_{L^2(\omega_0)}^2 - 2C_s \|v_1^\varepsilon\|_{L^2(\omega_0)} \|v_2^0\|_{L^2(\omega_0)} \\ &\geq (1 - C_s) \left(\|v_1^\varepsilon\|_{L^2(\omega_0)}^2 + \|v_2^0\|_{L^2(\omega_0)}^2 \right). \end{aligned}$$

Summarizing, this gives

$$\|Q(\mu_1, \mu_2)\|_{L^2(\Omega)}^2 \leq \frac{C(\omega_1; \omega_2)}{\tau^2(1 - C_s)} \|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}^2.$$

□

6.5.2 A priori error estimates to the reconstructed coarse scale solver in $\Omega \setminus \omega^+$.

In this section, we give an a priori error estimate in the coarse scale region $\Omega \setminus \omega^+$. The coupled solution restricted to the subregion $\Omega \setminus \omega^+$ is given by $u_2^{rec}(\theta_2)$.

Lemma 6.5.4. *Let u^ε and u_2^0 be the solutions of problems (6.1) and (6.4), respectively. Assume that (6.27) holds, we obtain*

$$\|u^\varepsilon - u_2^0(\theta_2)\|_{L^2(\omega_2)} \leq C\varepsilon.$$

Proof. Let us define an operator $P: \mathcal{U} \rightarrow H^1(\omega) \times H^1(\omega_2)$ by

$$P(\mu_1, \mu_2) = \begin{cases} u_{1,0}^\varepsilon + v_1^\varepsilon(\mu_1), & \text{in } \omega, \\ u_{2,0}^0 + v_2^0(\mu_2), & \text{in } \omega_2, \end{cases}$$

and consider the decomposition $P = U_0 + Q$, following (6.28). It holds $u_2^0(\theta_2) = P(\theta_1, \theta_2)|_{\omega_2}$, and

$$\begin{aligned} \|u^\varepsilon - u_2^0(\theta_2)\|_{L^2(\omega_2)} &\leq \|u^\varepsilon - P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon))\|_{L^2(\omega_2)} \\ &\quad + \|P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - P(\theta_1, \theta_2)\|_{L^2(\omega_2)}. \end{aligned}$$

The term $P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon))$, restricted to ω_2 , is equal to $u_2^0(\gamma_2(u^\varepsilon))$, which is defined as the homogenized solution u^0 obtained in (6.26). Using (6.27), we have

$$\begin{aligned} \|u^\varepsilon - u_2^0(\theta_2)\|_{L^2(\omega_2)} &\leq \|u^\varepsilon - u^0\|_{L^2(\omega_2)} + \|P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - P(\theta_1, \theta_2)\|_{L^2(\omega_2)} \\ &\leq C\varepsilon + \|Q\| \|\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})}. \end{aligned}$$

Following the proof of Lemma 6.5.3, we can show that $\|Q\|$ is bounded, and using Lemma 6.5.2, we obtain

$$\|u^\varepsilon - u_2^0(\theta_2)\|_{L^2(\omega_2)} \leq C_1\varepsilon + C_2\|u^\varepsilon - u^0\|_{L^2(\omega_0)} \leq C\varepsilon.$$

□

Theorem 6.5.5. *Let u^ε be the solution of (6.1) and $u_2^{rec}(\theta_2)$ be given by (6.25). Let $a_2(x, y) \in \mathcal{C}(\overline{\omega_2}; L_{per}^\infty(Y))$ and $\chi^j \in W_{per}(Y)$, $j = 1, \dots, d$. If in addition, $u^\varepsilon \in H^2(\Omega)$, $u_2^0(\theta_2) \in H^2(\omega_2)$, and $\chi^j \in W^{1,\infty}(Y)$, $j = 1, \dots, d$, it holds*

$$\|u^\varepsilon - u_2^{rec}(\theta_2)\|_{H^1(\Omega \setminus \omega^+)} \leq C\varepsilon^{1/2},$$

where the constant C is independent of ε , but depends on τ , τ^+ , and the ellipticity constants of a_2^ε .

Proof. Recall that u^0 is the homogenized solution of (6.26), and using the periodic correctors

χ^j , we have that the reconstructed solution $u^{0,rec}$ is given by

$$u^{0,rec}(x) = u^0(x) + \varepsilon \sum_{j=1}^d \chi^j(x, x/\varepsilon) \frac{\partial u^0(x)}{\partial x_j}.$$

Using the triangular inequality with $u^{0,rec}$, we have

$$\|u^\varepsilon - u_2^{rec}(\theta_2)\|_{H^1(\Omega \setminus \omega^+)} \leq \|u^\varepsilon - u^{0,rec}\|_{H^1(\Omega \setminus \omega^+)} + \|u^{0,rec} - u_2^{rec}(\theta_2)\|_{H^1(\Omega \setminus \omega^+)}.$$

The first norm is bounded by $C\varepsilon^{1/2}$; this follows from [80]. The second norm can be bounded by

$$\begin{aligned} \|u^{0,rec} - u_2^{rec}(\theta_2)\|_{H^1(\Omega \setminus \omega^+)} &\leq \|u^0 - u_2^0(\theta_2)\|_{H^1(\Omega \setminus \omega^+)} \\ &\quad + \varepsilon \left\| \sum_{j=1}^d \chi^j(x, x/\varepsilon) \left(\frac{\partial u^0(x)}{\partial x_j} - \frac{\partial u_2^0(x)}{\partial x_j} \right) \right\|_{H^1(\Omega \setminus \omega^+)}. \end{aligned}$$

Each of the term can be bounded by $C\varepsilon$, using the Caccioppoli inequality, Lemma 5.1.1, on the difference $u^0 - u_2^0(\theta_2)$ together with Lemma 6.5.4. \square

6.6 Summary

In this chapter, we gave a new global to local method based on optimal controls and minimization techniques. The method is based on an overlapping domain decomposition of the physical domain Ω into a region where a fine scale solver is used and a region where a coarse scale solver is preferred. The two solvers are coupled through a minimization problem, where we minimize the difference between the two solutions over the overlapping region; i.e., find (θ_1, θ_2) such that

$$J(\theta_1, \theta_2) = \min_{\mu_1, \mu_2} \frac{1}{2} \|u_1^\varepsilon(\mu_1) - u_2^0(\mu_2)\|_{L^2(\omega_0)}^2.$$

The well-posedness of the method is proved using Caccioppoli inequalities and a strong Cauchy–Schwarz inequality. The key to the existence and uniqueness of the optimal controls is to show that the cost functional is a norm over the space of controls.

Optimality systems can be derived from the minimization problem leading to a system of equations which can be used in the numerical coupling method. At last, an a priori error analysis of the optimization based method is proposed and error estimates are given; i.e.,

$$\begin{aligned} \|u^\varepsilon - u_1^\varepsilon(\theta_1)\|_{H^1(\omega^+)} &\leq C\varepsilon, \\ \|u^\varepsilon - u_2^{rec}(\theta_2)\|_{H^1(\Omega \setminus \omega^+)} &\leq C\varepsilon^{1/2}. \end{aligned}$$

Fully discrete optimization based coupling method

In here, we derive a numerical method to solve the optimization based coupling method. To fully resolve the fine scales in ω_1 , we need a partition with mesh size that resolves the fine scales, whereas the partition of $\Omega \setminus \omega_1$ can be coarse and independent of the smallest scale thanks to numerical homogenization techniques. In Figure 6.4, we recall the domain decomposition. In order to allow for flexible meshing, we do not impose continuity of the numerical homogenization method on Γ_1 . We choose to use a discontinuous Galerkin method on ω_2 and conform FEM on ω_1 . Imposing a discontinuity at the interface Γ_1 is a consequence of the choice to consider the same FE in the coarse and fine meshes over the overlap ω_0 . This choice is not optimal as it leads to the presence of small FE in the mesh over ω_2 and increases the number of degrees of freedom. To solve that issue, one could do an interpolation between the coarse and fine mesh in the overlap; this is treated in Chapter 7.

In what follows, we restrict the family of problems (6.3) to homogeneous Dirichlet problems, i.e., we set $g_D = 0$ and $\Gamma_N = \{\emptyset\}$; we seek u^ε such that

$$\begin{aligned} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, & \text{in } \Omega, \\ u^\varepsilon &= 0, & \text{on } \Gamma_D. \end{aligned} \tag{6.29}$$

We denote by $H_D^1(\omega_i)$ the set of functions in $H^1(\omega_i)$ that vanish on $\partial\omega_i \cap \Gamma_D$, for $i = 1, 2$. Further the analysis is conducted for piecewise macro and micro FE; however, the error estimates in the fine scale region ω^+ can be easily adapted to higher order FE, following the Sections 2.1 and 2.5.

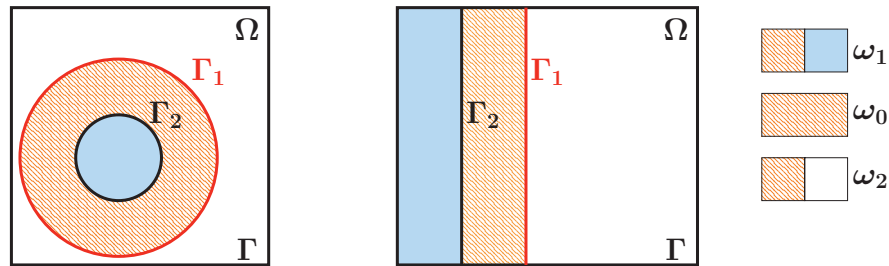


Figure 6.4 – Two possible domains decompositions with Γ_1 (in red) and Γ_2 (in black) for interior domains (left) and for domains with shared boundaries (right).

Chapter 6. Optimization based coupling method

Further, we assume that the strong Cauchy–Schwarz lemma, Lemma 6.3.3, holds, and assume that $\varepsilon \leq \varepsilon_0$. For simplicity of notations, we omit the superscripts ε and 0 in the numerical approximations of u_1^ε (v_1^ε) and u_2^0 (v_2^0).

We recall the coupling in its continuous form: find $u_1^\varepsilon \in H^1(\omega_1)$ and $u_2^0 \in H^1(\omega_2)$, such that the cost function

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)}^2,$$

is minimized under the following constraints, for $i = 1, 2$,

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla u_i(\theta_i)) &= f, & \text{in } \omega_i, \\ u_i(\theta_i) &= \theta_i, & \text{on } \Gamma_i, \\ u_i(\theta_i) &= g_D, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n_i \cdot (a_i(x)\nabla u_i(\theta_i)) &= g_N, & \text{on } \partial\omega_i \cap \Gamma_N, \end{aligned} \tag{6.30}$$

where $u_1 = u_1^\varepsilon$ and $u_2 = u_2^0$.

Outline. In Section 6.7, we give the numerical method for the fine and coarse scale solvers, and give the numerical algorithm in 6.7.3. In Section 6.8, we give discrete inequalities useful for well-posedness and a priori error analysis. The existence and uniqueness of the optimal controls are proved in Section 6.9 and error estimates are given in Section 6.10. In Section 6.11, we give numerical examples to assess the convergence rates and compare our method with other goal-oriented numerical methods.

6.7 Numerical method for the optimization based coupling.

The optimization based method couples a fine scale solver with a coarse scale solver. We choose to couple the FEM in ω_1 with the DG-FE-HMM in ω_2 , and recall in this section the two numerical methods.

6.7.1 Numerical method for the fine scale solver

Let $\{\mathcal{T}_{\tilde{h}}\}$ be a family of partitions over ω_1 , in simplicial or quadrilateral elements, with mesh size $\tilde{h} \ll \varepsilon$ where $\tilde{h} = \max_{K \in \mathcal{T}_{\tilde{h}}} h_K$, and h_K is the diameter of the element K . In addition, we suppose that the family of partitions $\{\mathcal{T}_{\tilde{h}}\}$ is admissible (T1) and shape-regular (T2). For simplicity, we consider, for each partition $\mathcal{T}_{\tilde{h}}$ in the family $\{\mathcal{T}_{\tilde{h}}\}$, a piecewise FE in ω_1 , given by

$$V_D^1(\omega_1, \mathcal{T}_{\tilde{h}}) = \{w \in H_D^1(\omega_1) \mid w|_K \in \mathcal{R}^1(K), \quad \forall K \in \mathcal{T}_{\tilde{h}}\},$$

where \mathcal{R}^1 is the space of piecewise linear polynomials on K . Further, we denote by $V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ the functions in $V_D^1(\omega_1, \mathcal{T}_{\tilde{h}})$ that vanish on $\partial\omega_1$.

Let $u_{1,\tilde{h}}$ be the numerical approximation of u_1^ε , the solution of (6.4) for $i = 1$. We decompose

6.7. Numerical method for the optimization based coupling.

$u_{1,\tilde{h}}$ into $u_{1,\tilde{h}} = u_{1,0,\tilde{h}} + v_{1,\tilde{h}}$, where $v_{1,\tilde{h}} \in V_D^1(\omega_1, \mathcal{T}_{\tilde{h}})$ is obtained by the optimization method and $u_{1,0,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ is the solution of

$$B_1(u_{1,0,\tilde{h}}, w_{1,\tilde{h}}) = F_1(w_{1,\tilde{h}}), \quad \forall w_{1,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}), \quad (6.31)$$

where B_1 is defined by (6.8), and F_1 is given by

$$F_1(w_{1,\tilde{h}}) = \int_{\omega_1} f w_{1,\tilde{h}} dx.$$

Thanks to the Poincaré inequality, the bilinear form B_1 is coercive and bounded over the space $V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$, and the existence and uniqueness of $u_{1,0,\tilde{h}}$ follows from the Lax–Milgram lemma. We note that a quadrature formula should be considered for the bilinear form B_1 and for the right hand side F_1 , and we refer to 2.1.1.

6.7.2 Discontinuous Galerkin (DG) method for the coarse scale problem

Let $\{\mathcal{T}_H\}$ be a family of partitions over ω_2 , with discontinuity in Γ_1 and mesh size $H = \max_{K \in \mathcal{T}_H} h_K$; further we assume that the family of partitions $\{\mathcal{T}_H\}$ are shape-regular (T2). For each partition \mathcal{T}_H in the family $\{\mathcal{T}_H\}$, we denote by E the set of $(d-1)$ dimensional elements of \mathcal{T}_H that form the boundary Γ_1 — it will be edges (for $d=2$) or faces (for $d=3$). Further, assume that the set E is composed of the smallest common interface between two elements K_+ and K_- of \mathcal{T}_H , with intersection in Γ_1 ; that is e is in E if $e = \min K_+ \cap K_-$ and $e \subset \Gamma_1$. As the solutions of problem (6.4), for $i=2$, are assumed to be continuous in $\omega_2 \setminus \Gamma_1$, we construct a piecewise FE space as

$$V_D^1(\omega_2, \mathcal{T}_H) = \{v \in H_D^1(\omega_2 \setminus \Gamma_1) \cap L^2(\omega_2) \mid v|_K \in \mathcal{R}^1(K), \quad \forall K \in \mathcal{T}_H\},$$

we denote by $V_0^1(\omega_2, \mathcal{T}_H)$ the set of functions of $V_D^1(\omega_2, \mathcal{T}_H)$ that vanish over $\partial\omega_2$. For $v \in V_D^1(\omega_2, \mathcal{T}_H)$, we consider its average $\{\cdot\}$ and its jump $[\![\cdot]\!]$ given by

$$\{v\} = \frac{1}{2}(v_+ + v_-), \quad \text{and} \quad [\![v]\!] = v_+ n_+ + v_- n_-,$$

where $v_{\pm} := v|_{K_{\pm}}$ denotes the trace of v from within K_{\pm} and n_{\pm} stands for the unit outward normal in K_{\pm} .

Quadrature formula. For piecewise FE spaces, a quadrature formula is given by the pair $(x_K, |K|)$, where x_K is the barycenter of K . The sampling domain of size δ around each point x_K is denoted by $K_{\delta} = x_K + \delta[-1/2, 1/2]^2$.

The numerically homogenized tensor $a_2^{0,h}(x_K)$, around the quadrature point x_K , is obtained using numerical solutions of micro problems defined in the sampling domains K_{δ} ; we note that a numerical approximation of f^0 can be obtained similarly. Let us consider a partition \mathcal{T}_h of K_{δ} in simplicial or quadrilateral elements K of diameter h_K ; the mesh size is $h = \max_{K \in \mathcal{T}_h} h_K$

Chapter 6. Optimization based coupling method

and, as the fine scales should be resolved in K_δ , we impose $h \leq \varepsilon$. The piecewise micro FE space is given by

$$V^1(K_\delta, \mathcal{T}_h) = \{v^h \in W(K_\delta) \mid v|_K \in \mathcal{R}^1(K), \quad \forall K \in \mathcal{T}_h\},$$

where $W(K_\delta)$ depends on the boundary conditions imposed on the micro problems (2.21) or (2.22). We introduce discrete micro problems: find $\psi_{K_\delta}^{i,h} \in V^1(K_\delta, \mathcal{T}_h)$, $i = 1, \dots, d$, the solution of

$$\int_{K_\delta} a_2^\varepsilon(x) \nabla \psi_{K_\delta}^{i,h} \cdot \nabla z^h dx = - \int_{K_\delta} a_2^\varepsilon(x) e_i \cdot \nabla z^h dx, \quad \forall z^h \in V^1(K_\delta, \mathcal{T}_h). \quad (6.32)$$

The numerically homogenized tensor at a quadrature point x_K in a macro element K , is computed by

$$a_2^{0,h}(x_K) = \frac{1}{|K_\delta|} \int_{K_\delta} a_2^\varepsilon(x) \left(I + \nabla \psi_{K_\delta}^h \right) dx, \quad (6.33)$$

where $\nabla \psi_{K_\delta}^h = (\nabla \psi_{K_\delta}^{1,h}, \dots, \nabla \psi_{K_\delta}^{d,h})$. Following [5], we define a DG macro bilinear form $B_{2,H} : V_D^1(\omega_2, \mathcal{T}_H) \times V_D^1(\omega_2, \mathcal{T}_H) \rightarrow \mathbb{R}$ by

$$\begin{aligned} B_{2,H}(v_{2,H}, w_{2,H}) = & \sum_{K \in \mathcal{T}_H} |K| a_2^{0,h}(x_K) \nabla v_{2,H}(x_K) \cdot \nabla w_{2,H}(x_K) \\ & + \sum_{e \in E} \int_e \mu_e \llbracket v_{2,H} \rrbracket \llbracket w_{2,H} \rrbracket ds \\ & - \sum_{e \in E} \int_e \left(\{a_2^{0,h}(x_K) \nabla v_{2,H}(x_K)\} \llbracket w_{2,H} \rrbracket \right. \\ & \left. + \{a_2^{0,h}(x_K) \nabla w_{2,H}(x_K)\} \llbracket v_{2,H} \rrbracket \right) ds, \end{aligned} \quad (6.34)$$

where the functions μ_e stand for weighting functions that penalize the jumps of $v_{2,H}$ and $w_{2,H}$ over the element e in E . They are given by

$$\mu_e = \kappa h_e^{-1}, \quad (6.35)$$

with $\kappa > 0$, and h_e is the size of the interface e .

The numerical homogenized solution $u_{2,H}$ is split into $u_{2,H} = u_{2,0,H} + v_{2,H}$, where $v_{2,H} \in V_D^1(\omega_2, \mathcal{T}_H)$ is given by the coupling and $u_{2,0,H} \in V_0^1(\omega_2, \mathcal{T}_H)$ is the solution of

$$B_{2,H}(u_{2,0,H}, w_{2,H}) = F_2(w_{2,H}), \quad \forall w_{2,H} \in V_0^1(\omega_2, \mathcal{T}_H). \quad (6.36)$$

The right hand side F_2 is given by

$$F_2(w_{2,H}) = \sum_{K \in \mathcal{T}_H} |K| f(x_K) w_{2,H}(x_K).$$

Remark 6.7.1. Considering non-homogeneous Dirichlet boundary condition $g_D \neq 0$ on Γ_D and Neumann condition on $\Gamma_N \neq \{0\}$ leads to some additional terms in the right hand sides F_1 and F_2 of problems (6.31) and (6.36), respectively. In particular, one should construct a lifting

6.7. Numerical method for the optimization based coupling.

of the Dirichlet data as explained in Section 6.5.

Remark 6.7.2. Higher order FE spaces can be considered and we note that the macro FEM over ω_1 and the micro FEM over K_{δ_j} can be easily generalized to higher order FEM. For the DG-FE-HMM some work needs to be done on the average of the fluxes, and we refer to [5].

6.7.3 Numerical Algorithm

In here, we state the discrete optimization based coupling, give the algorithm, and present the main convergence results. The well-posedness and the proofs of the errors estimates are given in the next sections.

The solution $(u_{1,\tilde{h}}, u_{2,H}) \in V_D^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_D^1(\omega_2, \mathcal{T}_H)$ satisfies

$$\min_{\mu_{1,\tilde{h}}, \mu_{2,H}} \frac{1}{2} \|u_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - u_{2,H}(\mu_{2,H})\|_{L^2(\omega_0)}^2 \text{ subject to } \begin{cases} B_1(u_{1,\tilde{h}}, w_{1,\tilde{h}}) = F_1(w_{1,\tilde{h}}), \\ B_{2,H}(u_{2,H}, w_{2,H}) = F_2(w_{2,H}), \end{cases}$$

for all $w_{1,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ and $w_{2,H} \in V_0^1(\omega_2, \mathcal{T}_H)$.

Introducing discrete Lagrange multipliers $\lambda_{1,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ and $\lambda_{2,H} \in V_0^1(\omega_2, \mathcal{T}_H)$ for each of the constraint, leads to a discrete optimality system:

find $(v_{1,\tilde{h}}, \lambda_{1,\tilde{h}}, v_{2,H}, \lambda_{2,H}) \in V_D^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_D^1(\omega_2, \mathcal{T}_H) \times V_0^1(\omega_2, \mathcal{T}_H)$ satisfying

$$\int_{\omega_0} (v_{1,\tilde{h}} - v_{2,H}) w_{1,\tilde{h}} dx - B_1(w_{1,\tilde{h}}, \lambda_{1,\tilde{h}}) = - \int_{\omega_0} (u_{1,0,\tilde{h}} - u_{2,0,H}) w_{1,\tilde{h}} dx, \quad (6.37)$$

$$B_1(v_{1,\tilde{h}}, \xi_{1,\tilde{h}}) = F_1(\xi_{1,\tilde{h}}) - B_1(u_{1,0,\tilde{h}}, \xi_{1,\tilde{h}}), \quad (6.38)$$

$$\int_{\omega_0} (v_{2,H} - v_{1,\tilde{h}}) w_{2,H} dx - B_{2,H}(w_{2,H}, \lambda_{2,H}) = \int_{\omega_0} (u_{1,0,\tilde{h}} - u_{2,0,H}) w_{2,H} dx, \quad (6.39)$$

$$B_{2,H}(v_{2,H}, \xi_{2,H}) = F_2(\xi_{2,H}) - B_{2,H}(u_{2,0,H}, \xi_{2,H}), \quad (6.40)$$

for all $w_{1,\tilde{h}} \in V_D^1(\omega_1, \mathcal{T}_{\tilde{h}})$, $\xi_{1,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$, $w_{2,H} \in V_D^1(\omega_2, \mathcal{T}_H)$, and $\xi_{2,H} \in V_0^1(\omega_2, \mathcal{T}_H)$. Notice that in the system (6.37) to (6.40), the discrete controls $(\theta_{1,\tilde{h}}, \theta_{2,H})$ are not in the unknown but they are hidden in the state variables $v_{1,\tilde{h}}$ and $v_{2,H}$.

The optimality system (6.37) to (6.40) can be written in matrix form, for the unknown vector $U = (v_{1,\tilde{h}}, v_{2,H}, \lambda_{1,\tilde{h}}, \lambda_{2,H})^\top$, as

$$\begin{pmatrix} M & -B^\top \\ B & 0 \end{pmatrix} U = G, \quad (6.41)$$

where

$$M(\{v_{1,\tilde{h}}, v_{2,H}\}, \{w_{1,\tilde{h}}, w_{2,H}\}) = \begin{pmatrix} \int_{\omega_0} v_{1,\tilde{h}} w_{1,\tilde{h}} dx & - \int_{\omega_0} v_{2,H} w_{1,\tilde{h}} dx \\ - \int_{\omega_0} v_{1,\tilde{h}} w_{2,H} dx & \int_{\omega_0} v_{2,H} w_{2,H} dx \end{pmatrix},$$

$$B(\{v_{1,\tilde{h}}, v_{2,H}\}, \{\lambda_{1,\tilde{h}}, \lambda_{2,H}\}) = \begin{pmatrix} B_1(v_{1,\tilde{h}}, \lambda_{1,\tilde{h}}) & 0 \\ 0 & B_{2,H}(v_{2,H}, \lambda_{2,H}) \end{pmatrix}.$$

The algorithm for the numerical coupling method is given below.

1. Find $u_{1,0,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ such that

$$B_1(u_{1,0,\tilde{h}}, w_{1,\tilde{h}}) = F_1(w_{1,\tilde{h}}), \quad \forall w_{1,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}). \quad (6.42)$$

2. Find $u_{2,0,H} \in V_0^1(\omega_2, \mathcal{T}_H)$ such that

$$B_{2,H}(u_{2,0,H}, w_{2,H}) = F_2(w_{2,H}), \quad \forall w_{2,H} \in V_0^1(\omega_2, \mathcal{T}_H). \quad (6.43)$$

3. Find $v_{1,\tilde{h}} \in V_D^1(\omega_1, \mathcal{T}_{\tilde{h}})$ and $v_{2,H} \in V_D^1(\omega_2, \mathcal{T}_H)$ by solving the saddle point problem (6.41).

Remark 6.7.3. The bilinear form B_1 and the right-hand side F_1 should be replaced by a discrete bilinear form $B_{1,\tilde{h}}$ and discrete right-hand side $F_{1,\tilde{h}}$ where a quadrature formula is used. But for simplicity of notations, we keep B_1 and F_1 .

We state the two main convergence results for the fully discrete coupling. The optimization based method relies on the DG-FE-HMM, thus one should expect to find the DG-FE-HMM error in the a priori estimates. Further, as we use the Caccioppoli inequalities, we give the L^2 error for the DG-FE-HMM. The a priori error is split into a macro, micro and modeling error; i.e.,

$$e_{\text{HMM},L^2} \leq e_{\text{MAC},L^2} + e_{\text{MIC}} + e_{\text{MOD}}.$$

The macro and micro errors correspond to FE errors due to the choice of macro and micro FE methods respectively. The modeling error is due to the upscaling procedure, and will be influenced by the choice of boundary conditions for (6.32), the size of the sampling domain δ , and whether we consider collocation in the macro and micro bilinear forms of the slow variable x to the quadrature points in the tensor a_2^ε . Details about the DG-FE-HMM error are given in the Section 2.5. We recall that (θ_1, θ_2) is the optimal couple of controls that minimize the cost J and that $u_1^\varepsilon(\theta_1)$ and $u_2^0(\theta_2)$ are the solutions of (6.30). Let $(\theta_{1,\tilde{h}}, \theta_{2,H})$ be the discrete couple of boundary conditions given by the minimization problem (6.5). We recall

6.7. Numerical method for the optimization based coupling.

the notations

$$\begin{aligned} u_{1,\tilde{h}}(\theta_{1,\tilde{h}}) & \text{ denotes the fine scale numerical solution in } \omega_1, \\ u_{2,H}(\theta_{2,H}) & \text{ denotes the coarse scale numerical solution in } \omega_2. \end{aligned}$$

The coupling solution, denoted by $\bar{u}_{\tilde{h}H}$, is defined as

$$\bar{u}_{\tilde{h}H} = \begin{cases} u_{1,\tilde{h}}(\theta_{1,\tilde{h}}), & \text{in } \omega^+, \\ u_{2,H}^{rec}(\theta_{2,H}), & \text{in } \Omega \setminus \omega^+, \end{cases} \quad (6.44)$$

where $u_{2,H}^{rec}(\theta_{2,H})$ corresponds to the reconstructed coarse scale solution $u_{2,H}(\theta_{2,H})$ and is defined by

$$u_{2,H}^{rec}(x) = u_{2,H}(x) + \sum_{j=1}^d \psi_{K_\varepsilon}^{j,h}(x) \frac{\partial u_{2,H}}{\partial x_j}(x), \quad x \in K,$$

where $\psi_{K_\varepsilon}^{j,h}$ are the micro solutions of (6.32). As the reconstructed numerical solution might be discontinuous across elements in ω_2 , we consider a broken H^1 semi-norm,

$$\|v\|_{\tilde{H}^1(\Omega)}^2 := \sum_{K \in \mathcal{T}_h(\omega^+)} \|\nabla v\|_{L^2(K)}^2 + \sum_{K \in \mathcal{T}_H(\Omega \setminus \omega^+)} \|\nabla v\|_{L^2(K)}^2.$$

We next state our main convergence result for the optimization based numerical method. Let u^0 be the homogenized solution of

$$\begin{aligned} -\operatorname{div}(a_2^0(x)\nabla u^0) &= f, & \text{in } \omega_2, \\ u^0 &= \gamma_2(u^\varepsilon), & \text{on } \Gamma_2, \\ u^0 &= 0, & \text{on } \partial\omega_2 \cap \Gamma_D, \\ n_2 \cdot (a_2^0(x)\nabla u^0) &= 0, & \text{on } \partial\omega_2 \cap \Gamma_N. \end{aligned} \quad (6.45)$$

We first have an error estimate in the fine scale region.

Theorem 6.7.4 (A priori error analysis in ω^+). *Let ε_0 be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and consider $\varepsilon \leq \varepsilon_0$. Let u^ε and u^0 be the exact solutions of problems (6.29) and (6.45), respectively, and $\bar{u}_{\tilde{h}H}$ be the numerical solution of the coupling (6.44). Assume $u^\varepsilon \in H^{s+1}(\Omega)$, with $s \leq 1$, $u^0 \in H^2(\omega_2)$, and assume that (6.27) holds, then*

$$\|u^\varepsilon - \bar{u}_{\tilde{h}H}\|_{\tilde{H}^1(\omega^+)} \leq C_1 \tilde{h}^s |u^\varepsilon|_{H^{s+1}(\omega_1)} + \frac{C_2}{\tau - \tau^+} (\tilde{h}^{s+1} |u^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + e_{\text{HMM},L^2}),$$

where the constants are independent of ε , H , \tilde{h} , and h . The DG-FE-HMM error e_{HMM,L^2} is given in Lemmas 6.10.2, 6.10.3, and 6.10.5.

Next, we state an error estimates in the coarse scale region $\Omega \setminus \omega^+$.

Theorem 6.7.5 (Error estimates in $\Omega \setminus \omega^+$). *Let u^ε be the exact solution of problem (6.29) and $\bar{u}_{\tilde{h}H}$ be the numerical solution of the coupling (6.44). Let $a_2^\varepsilon(x) = a_2(x, x/\varepsilon)$, where $a_2(x, y)$ is Y -periodic in y and satisfies $a_2(x, y) \in \mathcal{C}(\bar{\omega}_2; L_{\text{per}}^\infty(Y))$. Let $\psi_{K_\varepsilon}^j(x) \in W_{\text{per}}^1(K_\varepsilon)$, $j = 1, \dots, d$. If in addition, $u^\varepsilon \in H^2(\Omega)$, $u_2^0(\theta_2) \in H^2(\omega_2)$, $u_1^\varepsilon \in H^{s+1}(\omega_1)$, with $s \leq 1$, and $\psi_{K_\varepsilon}^j(x) \in W^{1,\infty}(K_\varepsilon)$, $j = 1, \dots, d$. It holds,*

$$\begin{aligned} \|u_2^{\text{rec}}(\theta_2) - \bar{u}_{\tilde{h}H}\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 \varepsilon^{1/2} + C_2 \left(\frac{h}{\varepsilon} \right) + C_3 H |u_2^0|_{H^2(\omega_2)} \\ &\quad + \frac{C_4}{\tau^+} (\tilde{h}^{s+1} |u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2 |u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

where the constants are independent of H, \tilde{h}, h , and ε .

6.8 Discrete inequalities

Let $\omega \subset \omega_1 \subset \Omega$, with $\tau = \text{dist}(\partial\omega_1, \partial\omega)$ and consider a partition \mathcal{T}_h of Ω in simplicial or quadrilateral elements K , with mesh size $h = \max_{K \in \mathcal{T}_h} h_K$, where h_K is the diameter of the element K . Further, we assume that h is smaller than τ and that \mathcal{T}_h is admissible (T1) and shape-regular (T2). The inequalities are given for general FE spaces of degree $p \geq 1$.

We give a discrete Caccioppoli inequality for functions $v^h \in V^p(\omega_1, \mathcal{T}_h)$ such that

$$B_1(v^h, w^h) := \int_{\omega_1} a(x) \nabla v^h \cdot \nabla w^h dx = 0, \quad \forall w^h \in V_0^p(\omega_1, \mathcal{T}_h). \quad (6.46)$$

Discrete versions of the Caccioppoli inequality were first given by Nitsche and Schatz [92] and Nitsche and Wahlbin [103]. We start this section by recalling the discrete Caccioppoli inequality and give its proof following the theory given in [92, 103] and more recently in [47]. Let us denote by I_h the Lagrange interpolant, and state a super approximation useful in the proof of the discrete Caccioppoli inequality.

Lemma 6.8.1 ([47]). *Let $\eta \in \mathcal{C}^1(\omega_1)$ with $|\nabla \eta| \leq C\tau^{-1}$. Then for each $v^h \in V^p(\omega_1, \mathcal{T}_h)$ and $K \in \mathcal{T}_h$, with $h_K \leq \tau$, it holds,*

$$\|\eta^2 v^h - I_h(\eta^2 v^h)\|_{H^1(K)} \leq C \left(\frac{h_K}{\tau} \|\nabla(\eta v^h)\|_{L^2(K)} + \frac{h_K}{\tau^2} \|v^h\|_{L^2(K)} \right).$$

Proof. See [47, Theorem 2.1]. □

We recall that local inverse inequalities are valid for functions $v^h \in V^p(\omega_1, \mathcal{T}_h)$; that is

$$\|\nabla v^h\|_{L^2(K)} \leq C h_K^{-1} \|v^h\|_{L^2(K)}, \quad (6.47)$$

where the constant C is independent of h_K .

We give here the proof of the discrete Caccioppoli inequality adapted to our problem and notations, but we precise that it follows the steps of the proof of [47, Lemma 3.3].

Lemma 6.8.2 (Discrete Caccioppoli inequality for interior domains). *Let $v^h \in V^p(\omega_1, \mathcal{T}_h)$ satisfy equation (6.46) for all $w^h \in V_0^p(\omega_1, \mathcal{T}_h)$; it holds*

$$\|\nabla v^h\|_{L^2(\omega)} \leq C \frac{1}{\tau} \|v^h\|_{L^2(\omega_1)},$$

where the constant C is independent of h , but depends on λ, Λ , and τ .

Proof. Let $\eta \in \mathcal{C}_0^1(\omega_1)$ be a cutoff function with $|\nabla \eta| \leq C\tau^{-1}$. We have that η satisfies $\eta \equiv 0$ in $\Omega \setminus \omega_1$, $\eta \equiv 1$ in ω , and $|\nabla \eta| \leq 1/\tau$ for points in ω_0 . By the uniform ellipticity of the tensor a , it holds

$$\lambda \|\nabla v^h\|_{L^2(\omega)}^2 \leq \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx.$$

Using $\eta^2 v^h$ as a test function in (6.46), and expanding the integral, we obtain

$$\int_{\omega_1} a \nabla v^h \cdot \nabla (\eta^2 v^h) dx = \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx + 2 \int_{\omega_1} a \eta \nabla v^h \cdot \nabla \eta v^h dx,$$

and thus

$$\begin{aligned} \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx &= \int_{\omega_1} a \nabla v^h \cdot \nabla (\eta^2 v^h) dx - 2 \int_{\omega_1} (\eta a^{1/2} \nabla v^h) \cdot (v^h a^{1/2} \nabla \eta) dx \\ &\leq \int_{\omega_1} a \nabla v^h \cdot \nabla (\eta^2 v^h) dx + 2 \int_{\omega_1} (\eta a^{1/2} \nabla v^h) \cdot (v^h a^{1/2} \nabla \eta) dx \\ &\leq B_1(v^h, \eta^2 v^h) + \zeta \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx + \frac{1}{\zeta} \int_{\omega_1} a v^h \nabla \eta \cdot \nabla \eta v^h dx \\ &\leq B_1(v^h, \eta^2 v^h) + \zeta \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx + \frac{\Lambda}{\zeta \tau^2} \|v^h\|_{L^2(\omega_1)}^2. \end{aligned}$$

The last step is to bound the quantity $B_1(v^h, \eta^2 v^h)$. Let us consider $I_h(\eta^2 v^h) \in V^p(\omega_1, \mathcal{T}_h)$, it holds

$$B_1(v^h, I(\eta^2 v^h)) = 0,$$

and then

$$\begin{aligned} B_1(v^h, \eta^2 v^h) &= B_1(v^h, \eta^2 v^h - I(\eta^2 v^h)) = \int_{\omega_1} a \nabla v^h \nabla (\eta^2 v^h - I(\eta^2 v^h)) dx \\ &\leq \Lambda \|\nabla v^h\|_{L^2(\omega_1)} \|\nabla (\eta^2 v^h - I(\eta^2 v^h))\|_{L^2(\omega_1)} \\ &\leq \Lambda \sum_{K \in \mathcal{T}_h} \|\nabla v^h\|_{L^2(K)} \|\nabla (\eta^2 v^h - I(\eta^2 v^h))\|_{L^2(K)}. \end{aligned}$$

Chapter 6. Optimization based coupling method

Using the local inverse inequality (6.47) and Lemma 6.8.1, we obtain

$$\begin{aligned}
B_1(v^h, \eta^2 v^h) &\leq C\Lambda \sum_{K \in \mathcal{T}_h} \frac{1}{h_K} \|v^h\|_{L^2(K)} \left(\frac{h_K}{\tau} \|\nabla(\eta v^h)\|_{L^2(K)} + \frac{h_K}{\tau^2} \|v^h\|_{L^2(K)} \right) \\
&= \Lambda \sum_{K \in \mathcal{T}_h} \|v^h\|_{L^2(K)} \frac{C}{\tau} \|\nabla(\eta v^h)\|_{L^2(K)} + \frac{C}{\tau^2} \|v^h\|_{L^2(K)}^2 \\
&\leq \Lambda \sum_{K \in \mathcal{T}_h} \frac{C}{\tau^2} \left(\frac{1}{\zeta} + 1 \right) \|v^h\|_{L^2(K)}^2 + \zeta \|\nabla(\eta v^h)\|_{L^2(K)}^2 \\
&\leq \frac{C\Lambda}{\tau^2} \left(\frac{1}{\zeta} + 1 \right) \|v^h\|_{L^2(\omega_1)}^2 + \Lambda\zeta \|\eta \nabla v^h\|_{L^2(\omega_1)}^2 + \Lambda\zeta \|v^h \nabla \eta\|_{L^2(\omega_1)}^2 \\
&\leq \Lambda \left(\frac{C}{\tau^2} \left(\frac{1}{\zeta} + 1 + \zeta \right) \|v^h\|_{L^2(\omega_1)}^2 + \zeta \|\eta \nabla v^h\|_{L^2(\omega_1)}^2 \right).
\end{aligned}$$

Recalling that

$$\|\eta \nabla v^h\|_{L^2(\omega_1)}^2 = \int_{\omega_1} \nabla v^h \cdot \nabla v^h \eta^2 dx \leq \frac{1}{\lambda} \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx,$$

and collecting the previous bounds, it holds

$$\begin{aligned}
\int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx &\leq C \frac{\Lambda}{\tau^2} \left(\frac{2}{\zeta} + 1 + \zeta \right) \|v^h\|_{L^2(\omega_1)}^2 \\
&\quad + \zeta \left(\frac{\Lambda}{\lambda} + 1 \right) \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx.
\end{aligned}$$

This gives, for $\zeta \neq 1/(\Lambda/\lambda + 1)$,

$$(1 - \zeta(\Lambda/\lambda + 1)) \int_{\omega_1} a \nabla v^h \cdot \nabla v^h \eta^2 dx \leq C \frac{\Lambda}{\tau^2} \left(\frac{2}{\zeta} + 1 + \zeta \right) \|v^h\|_{L^2(\omega_1)}^2,$$

and finally

$$\|\nabla v^h\|_{L^2(\omega)}^2 \leq \frac{C}{(1 - \zeta(\Lambda/\lambda + 1))} \left(\frac{2}{\zeta} + 1 + \zeta \right) \frac{\Lambda}{\lambda \tau^2} \|v^h\|_{L^2(\omega_1)}^2.$$

□

Assume now that $\partial\omega \cap \Gamma \neq \emptyset$. A discrete Caccioppoli inequality can be proved.

Lemma 6.8.3 (Discrete Caccioppoli inequality for domains with shared boundaries). *Let $v^h \in V^p(\omega_1, \mathcal{T}_h)$ satisfy equation (6.46) for all $w^h \in V_0^p(\omega_1, \mathcal{T}_h)$. Further assume that $v^h = 0$ on $\partial\omega_1 \cap \Gamma$. Then it holds*

$$\|\nabla v^h\|_{L^2(\omega)} \leq C \frac{1}{\tau} \|v^h\|_{L^2(\omega_1)},$$

where the constant C is independent of h , but depends on λ, Λ , and τ .

Proof. We consider now a cutoff function η such that $\eta \equiv 1$ in $\bar{\omega}$, $\eta \equiv 0$ in $\Omega \setminus \omega_1$, and with $\eta \equiv 0$

on $\partial\omega_1 \cap \Omega$. We can then follow the proof of Lemma 6.8.2, as

$$B_1(v^h, \eta^2 v^h) = \int_{\omega_1} a \nabla v^h \cdot \nabla (\eta^2 v^h) dx = 0.$$

□

We show that the strong Cauchy–Schwarz lemma, Lemma 6.3.3, is valid for discrete functions.

Lemma 6.8.4. *Let $\varepsilon < \varepsilon_0$ and $C_s < 1$ be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and let $v_{1,\tilde{h}} \in V_D^p(\omega_1, \mathcal{T}_{\tilde{h}})$ and $v_{2,H} \in V_D^p(\omega_2, \mathcal{T}_H)$ be numerical solutions of (6.41). There exist $\tilde{h}_0 > 0$ and $H_0 > 0$ such that*

$$\int_{\omega_0} v_{1,\tilde{h}} v_{2,H} dx \leq C_s \|v_{1,\tilde{h}}\|_{L^2(\omega_0)} \|v_{2,H}\|_{L^2(\omega_0)}, \quad \forall \tilde{h} < \tilde{h}_0, H < H_0.$$

Proof. Let $\{\tilde{h}_n, H_n\}_{n \geq 1}$ be a sequence of mesh sizes converging to zero. We have strong convergence in L^2 , for a subsequence of $\{\tilde{h}_n, H_n\}_{n \geq 1}$ still denoted by $\{\tilde{h}_n, H_n\}_{n \geq 1}$, of the numerical solutions v_{1,\tilde{h}_n} and v_{2,H_n} to the exact solutions v_1^ε and v_2^0 respectively. Thus

$$\lim_{n \rightarrow \infty} \int_{\omega_0} v_{1,\tilde{h}_n} v_{2,H_n} dx = \int_{\omega_0} v_1^\varepsilon v_2^0 dx$$

and

$$\lim_{n \rightarrow \infty} \|v_{1,\tilde{h}_n}\|_{L^2(\omega_0)} = \|v_1^\varepsilon\|_{L^2(\omega_0)},$$

$$\lim_{n \rightarrow \infty} \|v_{2,H_n}\|_{L^2(\omega_0)} = \|v_2^0\|_{L^2(\omega_0)}.$$

We recall that the strong Cauchy–Schwarz lemma, Lemma 6.8.4, is valid for v_1^ε et v_2^0 ; there exists an ε_0 and a constant $0 < C_s < 1$, such that for all $\varepsilon \leq \varepsilon_0$, it holds

$$\int_{\omega_0} v_1^\varepsilon v_2^0 dx \leq C_s \|v_1^\varepsilon\|_{L^2(\omega_0)} \|v_2^0\|_{L^2(\omega_0)}.$$

Then, using the strong Cauchy–Schwarz inequality, Lemma 6.3.3, for v_1^ε and v_2^0 , it holds

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\omega_0} v_{1,\tilde{h}_n} v_{2,H_n} dx &= \int_{\omega_0} v_1^\varepsilon v_2^0 dx \\ &\leq C_s \|v_1^\varepsilon\|_{L^2(\omega_0)} \|v_2^0\|_{L^2(\omega_0)} \\ &= \lim_{n \rightarrow \infty} C_s \|v_{1,\tilde{h}_n}\|_{L^2(\omega_0)} \|v_{2,H_n}\|_{L^2(\omega_0)}. \end{aligned}$$

Then, there exist an $\varepsilon_0 > 0$ and a constant $0 < C_s < 1$, such that for all $\varepsilon \leq \varepsilon_0$, there exist $\tilde{h}_0 > 0$ and $H_0 > 0$, such that

$$\int_{\omega_0} v_{1,\tilde{h}} v_{2,H} dx \leq C_s \|v_{1,\tilde{h}}\|_{L^2(\omega_0)} \|v_{2,H}\|_{L^2(\omega_0)}, \quad \forall \tilde{h} \leq \tilde{h}_0, H \leq H_0.$$

□

6.9 Well-posedness of the discrete coupling method

In this section, we prove the well-posedness of the discrete coupling problem. The well-posedness of the optimization based coupling method can be established using Brezzi's theory [34] and the well-posedness of problems (6.42) and (6.43). The Lax–Milgram lemma implies the existence and uniqueness of $u_{1,0,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$.

Due to the discontinuity in \mathcal{T}_H , the space $V_0^1(\omega_2, \mathcal{T}_H)$ is not a subspace of $H_0^1(\omega_2)$, however, it will lie in the piecewise Sobolev space

$$H^2(\mathcal{T}_H) := \prod_{K \in \mathcal{T}_H} H^2(K) = \{v \in L^1(\omega_2) \mid v|_K \in H^2(K), \forall K \in \mathcal{T}_H\}.$$

Suppose that the exact solution $u_{2,0}$ of problem (6.7) is in the space $H_0^1(\omega_2) \cap H^2(\omega_2)$, we define the proper space for the analysis as $V(\omega_2) := V_0^1(\omega_2, \mathcal{T}_H) + H_0^1(\omega_2) \cap H^2(\omega_2) \subset H^2(\mathcal{T}_H)$, see discussions in [26, 48]. The space $V(\omega_2)$ is equipped with the norm

$$\|v\|_{\omega_2} := \left(\|\nabla v\|_{L^2(\omega_2)}^2 + \sum_{K \in \mathcal{T}_H} h_K^2 |v|_{2,K}^2 + |v|_*^2 \right)^{1/2}, \quad (6.48)$$

where

$$\|\nabla v\|_{L^2(\omega_2)}^2 = \sum_{K \in \mathcal{T}_H} |v|_{1,K}^2, \quad |v|_{2,K}^2 = \sum_{|r|=2} \|\partial^r v\|_{L^2(K)}^2, \quad \text{and} \quad |v|_*^2 = \sum_{e \in E} \|\mu_e^{1/2} \llbracket v \rrbracket\|_{L^2(e)}^2.$$

One can prove that (6.48) is a norm over $V(\omega_2)$, using the discrete Poincaré–Friedrichs inequality [34],

$$\|v\|_{L^2(\omega_2)}^2 \leq C(\|\nabla v\|_{L^2(\omega_2)}^2 + |v|_*^2). \quad (6.49)$$

Thanks to local inverse inequalities [41], restricting the space $V(\omega_2)$ to $V_0^1(\omega_2, \mathcal{T}_H)$, reduces the norm (6.48) to

$$\|v\|_{\omega_2} = \left(\|\nabla v\|_{L^2(\omega_2)}^2 + |v|_*^2 \right)^{1/2}.$$

Proposition 6.9.1. *There exists a value κ_0 , that depends only on the properties of the tensor a^ε given in (6.2), the shape regularity of \mathcal{T}_H , and the dimension d , such that for all $\kappa \geq \kappa_0$, κ being defined in (6.35), the bilinear form $B_{2,H}$ (6.34) is stable in $V_0^1(\omega_2, \mathcal{T}_H)$; i.e.,*

$$B_{2,H}(v_H, v_H) \geq C_1 \|v_H\|_{\omega_2}^2, \quad \forall v_H \in V_0^1(\omega_2, \mathcal{T}_H).$$

Furthermore, the bilinear form is bounded; i.e.,

$$B_{2,H}(v_H, w_H) \leq C_2 \|v_H\|_{\omega_2} \|w_H\|_{\omega_2}, \quad \forall v_H, w_H \in V_0^1(\omega_2, \mathcal{T}_H).$$

The constants C_1 and C_2 are independent of H, \tilde{h}, h , and ε .

Proof. See [5, Lemmas 4.3, 4.4, and 5.18]. □

6.9. Well-posedness of the discrete coupling method

Theorem 6.9.2. *Let assumption (6.2) holds. Then there exists a unique solution $u_{1,0,\tilde{h}}$ of problem (6.42) which satisfies $u_{1,0,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ and*

$$\|u_{1,0,\tilde{h}}\|_{H^1(\omega_1)} \leq C_1 \|F_1\|_{H^{-1}(\omega_1)},$$

with a constant C_1 independent of H, \tilde{h} , and ε .

Moreover, let κ_0 be given by Proposition 6.9.1. Then, the problem (6.43) admits a unique solution $u_{2,0,H} \in V_0^1(\omega_2, \mathcal{T}_H)$ and it holds

$$\|u_{2,0,H}\| \leq C_2 \|F_2\|_{H^{-1}(\omega_2)},$$

where the constant C_2 is independent of H, h, \tilde{h} , and ε .

Proof. The existence and uniqueness of $u_{1,0,\tilde{h}}$ and $u_{2,0,H}$ follows from Lax–Milgram lemma and Proposition 6.9.1. \square

We now prove that the saddle point problem 6.41 is well posed. We introduce $V^1(\Gamma_i)$ as the set of functions $\mu_i \in \mathcal{W}^i$ that are piecewise polynomials on the elements over Γ_i , $i = 1, 2$. Let us write the system of equations (6.37) to (6.40) in terms of the discrete virtual controls $\theta_{1,\tilde{h}}$ and $\theta_{2,H}$: find $(\theta_{1,\tilde{h}}, \lambda_{1,\tilde{h}}, \theta_{2,H}, \lambda_{2,H}) \in V^1(\Gamma_1) \times V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V^1(\Gamma_2) \times V_0^1(\omega_2, \mathcal{T}_H)$ satisfying

$$\pi((\theta_{1,\tilde{h}}, \theta_{2,H}), (\mu_{1,\tilde{h}}, \mu_{2,H})) - B((\mu_{1,\tilde{h}}, \mu_{2,H}), (\lambda_{1,\tilde{h}}, \lambda_{2,H})) = G(\mu_{1,\tilde{h}}, \mu_{2,H}), \quad (6.50)$$

$$B((\theta_{1,\tilde{h}}, \theta_{2,H}), (\xi_{1,\tilde{h}}, \xi_{2,H})) = 0, \quad (6.51)$$

for all $(\mu_{1,\tilde{h}}, \mu_{2,H}) \in V^1(\Gamma_1) \times V^1(\Gamma_2)$ and $(\xi_{1,\tilde{h}}, \xi_{2,H}) \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_0^1(\omega_2, \mathcal{T}_H)$. The forms π , B , and G are defined by

$$\pi((\theta_{1,\tilde{h}}, \theta_{2,H}), (\mu_{1,\tilde{h}}, \mu_{2,H})) = \int_{\omega_0} (v_{1,\tilde{h}}(\theta_{1,\tilde{h}}) - v_{2,H}(\theta_{2,H}))(v_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - v_{2,H}(\mu_{2,H})) dx,$$

$$B((\theta_{1,\tilde{h}}, \theta_{2,H}), (\xi_{1,\tilde{h}}, \xi_{2,H})) = B_1(\theta_{1,\tilde{h}}, \xi_{1,\tilde{h}}) + B_{2,H}(\theta_{2,H}, \xi_{2,H}),$$

$$G(\theta_{1,\tilde{h}}, \theta_{2,H}) = - \int_{\omega_0} (u_{1,0,\tilde{h}} - u_{2,0,H})(v_{1,\tilde{h}}(\theta_{1,\tilde{h}}) - v_{2,H}(\theta_{2,H})) dx.$$

(Note that, in order to avoid overloading of notation, we reuse the notation π in the discrete context, which should not be confused with (6.9).)

To prove the well-posedness of system (6.50)–(6.51), we need to show that

- The form π is continuous and coercive on $V^1(\Gamma_1) \times V^1(\Gamma_2)$ equipped with the inner product π .
- The form B is continuous and satisfies an inf-sup condition.

The continuity of π can be easily obtained with the Cauchy–Schwarz and the discrete Poincaré inequality (6.49).

Chapter 6. Optimization based coupling method

The coercivity of π can be proved similarly to the continuum case (cf. see Lemma 6.3.4), as is done in the next lemma.

Lemma 6.9.3. *Let ε_0 be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and assume that $\varepsilon \leq \varepsilon_0$. Then, the form π defines an inner product on $V^1(\Gamma_1) \times V^1(\Gamma_2)$.*

Proof. We will use the discrete Cauchy–Schwarz lemma, Lemma 6.8.4, with the same ε_0 and C_s , to prove that π is definite. Indeed, arguing as in Lemma 6.3.4 we assume that $(\mu_{1,\tilde{h}}, \mu_{2,H})$ is such that

$$\begin{aligned} 0 &= \pi((\mu_{1,\tilde{h}}, \mu_{2,H}), (\mu_{1,\tilde{h}}, \mu_{2,H})) = \|v_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - v_{2,H}(\mu_{2,H})\|_{L^2(\omega_0)}^2 \\ &\geq (1 - C_s) \left(\|v_{1,\tilde{h}}(\mu_{1,\tilde{h}})\|_{L^2(\omega_0)}^2 + \|v_{2,H}(\mu_{2,H})\|_{L^2(\omega_0)}^2 \right). \end{aligned}$$

As $C_s < 1$, it holds that $\|v_{1,\tilde{h}}(\mu_{1,\tilde{h}})\|_{L^2(\omega_0)} = \|v_{2,H}(\mu_{2,H})\|_{L^2(\omega_0)} = 0$ which implies that $v_{1,\tilde{h}}(\mu_{1,\tilde{h}}) = v_{2,H}(\mu_{2,H}) = 0$ in ω_0 and, in particular, $\mu_{1,\tilde{h}} = 0$ and $\mu_{2,H} = 0$. \square

Next, we prove the inf-sup condition for the bilinear form B .

Lemma 6.9.4. *The form B satisfies*

$$\sup_{(\mu_{1,\tilde{h}}, \mu_{2,H})} \frac{B((\mu_{1,\tilde{h}}, \mu_{2,H}), (\xi_{1,\tilde{h}}, \xi_{2,H}))}{\|(\mu_{1,\tilde{h}}, \mu_{2,H})\|_{L^*(\mathcal{U})}} \geq C \left(\|\xi_{1,\tilde{h}}\|_{H^1(\omega_1)} + \|\xi_{2,H}\|_{\omega_2} \right),$$

for all $(\xi_{1,\tilde{h}}, \xi_{2,H}) \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_0^1(\omega_2, \mathcal{T}_H)$. The constant C is independent of ε .

Proof. Let $(\xi_{1,\tilde{h}}, \xi_{2,H}) \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_0^1(\omega_2, \mathcal{T}_H)$. By the definition of B , we have

$$B((\mu_{1,\tilde{h}}, \mu_{2,H}), (\xi_{1,\tilde{h}}, \xi_{2,H})) = B_1(\mu_{1,\tilde{h}}, \xi_{1,\tilde{h}}) + B_{2,H}(\mu_{2,H}, \xi_{2,H}).$$

Take $(\mu_{1,\tilde{h}}, \mu_{2,H}) \in V^1(\Gamma_1) \times V^1(\Gamma_2)$ and by definition there exist $v_{\tilde{h}} \in V^1(\omega_1, \mathcal{T}_{\tilde{h}})$ and $v_H \in V^0(\omega_2, \mathcal{T}_H)$ such that $v_{\tilde{h}}(\mu_{1,\tilde{h}}) = \xi_{1,\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}})$ and $v_H(\mu_{2,H}) = \xi_{2,H} \in V_0^1(\omega_2, \mathcal{T}_H)$. Then,

$$B_1(\mu_{1,\tilde{h}}, \xi_{1,\tilde{h}}) = \int_{\omega_1} a_1^\varepsilon \nabla v_{\tilde{h}}(\mu_{1,\tilde{h}}) \cdot \nabla \xi_{1,\tilde{h}} dx = \int_{\omega_1} a_1^\varepsilon \nabla \xi_{1,\tilde{h}} \cdot \nabla \xi_{1,\tilde{h}} dx \geq C \|\xi_{1,\tilde{h}}\|_{H^1(\omega_1)}^2.$$

Similarly, by the coercivity of $B_{2,H}$, it holds

$$B_{2,H}(\mu_{2,H}, \xi_{2,H}) \geq C \|\xi_{2,H}\|_{\omega_2}^2.$$

Thus,

$$B((\mu_{1,\tilde{h}}, \mu_{2,H}), (\xi_{1,\tilde{h}}, \xi_{2,H})) \geq C \left(\|\xi_{1,\tilde{h}}\|_{H^1(\omega_1)} + \|\xi_{2,H}\|_{\omega_2} \right)^2,$$

where the constant is independent of H , h , \tilde{h} , and ε . We can conclude as

$$\begin{aligned} \|(\mu_{1,\tilde{h}}, \mu_{2,H})\|_{L^*(\mathcal{W})} &\leq \|\mu_{1,\tilde{h}}\|_{H^{1/2}(\Gamma_1)} + \|\mu_{2,H}\|_{H^{1/2}(\Gamma_2)} \\ &\leq C \left(\|v_{\tilde{h}}(\mu_{1,\tilde{h}})\|_{H^1(\omega_1)} + \|v_H(\mu_{2,H})\|_{\omega_2} \right) \\ &= C \left(\|\xi_{1,\tilde{h}}\|_{H^1(\omega_1)} + \|\xi_{2,H}\|_{\omega_2} \right). \end{aligned}$$

□

6.10 Fully discrete error estimates

In this section, we derive error estimates for the fully discrete optimization-based method. A post-processing procedure is used on the coarse solution $u_{2,H}(\theta_{2,H})$, in order to reach convergence to the exact solution u^ε . The norm considered is a broken H^1 semi-norm as we allow the corrected solution to be discontinuous across elements of $\Omega \setminus \omega$. The fully discrete analysis is then conducted for the error

$$\|u^\varepsilon - \tilde{u}_{\tilde{h}H}\|_{\tilde{H}^1(\Omega)} = \sum_{K \in \mathcal{T}_{\tilde{h}}(\omega^+)} \|\nabla(u^\varepsilon - \tilde{u}_{\tilde{h}H})\|_{L^2(K)}^2 + \sum_{K \in \mathcal{T}_H(\Omega \setminus \omega^+)} \|\nabla(u^\varepsilon - \tilde{u}_{\tilde{h}H})\|_{L^2(K)}^2.$$

where the numerical solution of the coupling $\tilde{u}_{\tilde{h}H}$ is given by (6.44). In the fully discrete analysis of the DG-FE-HMM method, the error between the homogenized solution and its approximation is decomposed into a macro, micro, and modeling error [3]. These errors will contribute to the a priori estimates of our method.

Remark 6.10.1. In Section 6.5, the error estimates depend on the bound of the operator Q (6.28). This bound was obtained in Lemma 6.5.3 using Caccioppoli inequalities. In the fully discrete case, we introduce a discrete operator $Q^{\tilde{h},H}$, which is a discrete version of the operator Q and the estimates will depend on $\|Q^{\tilde{h},H}\|$. For conforming FE spaces the norm of $Q^{\tilde{h},H}$ is bounded independently of the mesh sizes \tilde{h} , h , and H ; this can be seen by following the lines of Lemma 6.5.3. For non-conforming meshes, we will assume that $\|Q^{\tilde{h},H}\|$ is bounded independently of \tilde{h} , h , and H . In what follows, we will use the notations P , U_0 , and Q , previously used in the continuous analysis, to denote the operators in the discrete analysis.

We recall that u^0 , the solution of (6.26), denotes the homogenized solution over ω_2 with boundary condition on Γ_2 given by the trace of the physical solution u^ε , for a fixed ε . The DG-FE-HMM method gives us an approximation $u^H \in V_0^1(\omega_2, \mathcal{T}_H)$ of the homogenized solution u^0 . We state here the main results needed to bound $\|u^0 - u^H\|_{L^2(\omega_2)}$, for further details we refer to [1, 3, 4], and the references therein. We decompose the DG-FE-HMM error into the macro, micro, and modeling errors

$$e_{\text{HMM}, L^2} = \|u^0 - u^H\|_{L^2(\omega_2)} \leq e_{\text{MAC}} + e_{\text{MIC}} + e_{\text{MOD}}.$$

Macro Error

We define $u_H^0 \in V_0^1(\omega_2, \mathcal{T}_H)$ as the FEM approximation of the homogenized problem (6.26), i.e.,

$$B_{2,H}^0(u_H^0, w_H) = F_2(w_H), \quad \forall w_H \in V_0^1(\omega_2, \mathcal{T}_H), \quad (6.52)$$

where the bilinear form is given by

$$\begin{aligned} B_{2,H}^0(v_H, w_H) = & \sum_{K \in \mathcal{T}_H} |K| a_2^0(x_K) \nabla v^H \nabla w^H + \sum_{e \in E} \int_e \mu_e \llbracket v_H \rrbracket \llbracket w_H \rrbracket ds \\ & - \sum_{e \in E} \int_e (\{a_2^0 \nabla v_H\} \llbracket w_H \rrbracket + \{a_2^0 \nabla w_H\} \llbracket v_H \rrbracket) ds, \quad \forall v_H, w_H \in V_0^1(\omega_2, \mathcal{T}_H). \end{aligned} \quad (6.53)$$

Using a triangular inequality, the error can be formulated as

$$\begin{aligned} \|u^0 - u^H\|_{L^2(\omega_2)} & \leq \|u^0 - u_H^0\|_{L^2(\omega_2)} + \|u_H^0 - u^H\|_{L^2(\omega_2)} \\ & \leq \|u^0 - u_H^0\|_{L^2(\omega_2)} + \|u_H^0 - u^H\|_{\omega_2} \\ & = e_{\text{MAC}, L^2} + e_{\text{MIC}} + e_{\text{MOD}}. \end{aligned}$$

To simplify the analysis we make the following assumptions on the structure of the tensor a_2^ε ,

- (H1) $a_2^\varepsilon(x) = a_2(x, x/\varepsilon) = a_2(x, y)$ is Y -periodic in y and $a_2(\cdot, y)|_K$ is constant within each $K \in \mathcal{T}_H$.

Lemma 6.10.2 (Macro error). *Let u^0 and u_H^0 be the solutions of problems (6.26) and (6.52) respectively. Assume that (6.2) and (H1) hold, and that $u^0 \in H^2(\omega_2)$. Then,*

$$e_{\text{MAC}, L^2} = \|u^0 - u_H^0\|_{L^2(\omega_2)} \leq CH^2,$$

where the constant C is independent of H, \tilde{h}, h , and ε , but depends on the stability constant of the bilinear form $B_{2,H}^0$.

Proof. See [26]. □

Micro and modeling Errors

For the micro and modeling errors, we follow [5, Section 5]. We assume the following regularity on $\psi_{K_\delta}^i$, the non-discretized micro solutions of problem (6.32), in $W(K_\delta)$; i.e.,

- (H2) $|\psi_{K_\delta}^i|_{H^2(K_\delta)} \leq C\varepsilon^{-1} \sqrt{|K_\delta|}, \quad \text{for } i = 1, \dots, d.$

To discuss the micro and modeling errors, we recall that a_2^0 is the homogenized tensor on the domain ω_2 and that $a_2^{0,h}$ is the numerical homogenized tensor given by (6.33). Consider, further the tensor \bar{a}_2^0 defined by (6.33) using the non-discretized micro functions $\psi_{K_\delta}^i$, the

solutions of (6.32) in $W(K_\delta)$ instead of the discretized functions $\psi_{K_\delta}^{i,h}$. The error between the homogenized tensor a_2^0 and its numerical approximation $a_2^{0,h}$ can be bounded by

$$\sup_{K \in \mathcal{T}_H} \|a_2^0(x_K) - a_2^{0,h}(x_K)\|_F \leq \sup_{K \in \mathcal{T}_H} \|a_2^0(x_K) - \bar{a}_2^0(x_K)\|_F + \sup_{K \in \mathcal{T}_H} \|\bar{a}_2^0(x_K) - a_2^{0,h}(x_K)\|_F,$$

where the first term in the right-hand side of the above inequality is denoted by e_{MOD} (modeling error) and the second by e_{MIC} (micro error).

Lemma 6.10.3 (Micro and modeling errors). *Let u_H^0 be the solution of (6.52) and u^H be the DG-FE-HMM approximation of u^0 . Assume that (6.2) holds, then*

$$\|u_H^0 - u^H\|_{\omega_2} \leq C \sup_{K \in \mathcal{T}_H} \|a_2^0(x_K) - a_2^{0,h}(x_K)\|_F \|u^H\|,$$

where the constant C is independent of H , \tilde{h} , h , and ε . Further, assuming (H2), the Frobenius norm is bounded by

$$\sup_{K \in \mathcal{T}_H} \|a_2^0(x_K) - a_2^{0,h}(x_K)\|_F \leq e_{\text{MOD}} + C \left(\frac{h}{\varepsilon}\right)^2,$$

where the modeling error e_{MOD} is given in Lemma 6.10.5.

Proof. Follows from [5, Section 5]. □

Remark 6.10.4. Higher order micro error $\left(\frac{h}{\varepsilon}\right)^{2q}$ can be obtained for higher order micro FEM, provided higher order regularity of the micro functions,

$$|\psi_{K_\delta}^i|_{H^{q+1}(K_\delta)} \leq C\varepsilon^{-q} \sqrt{|K_\delta|} \quad \text{for } i = 1, \dots, d.$$

The modeling error e_{MOD, H^1} will depend on the choice of boundary condition on the micro problems. We recall that by collocation we mean that we collocate the slow variable x in the tensor a_2^ε to the quadrature points x_K in the definitions of the tensor a_2^0 and \bar{a}_2^0 .

Lemma 6.10.5 (Modeling error). *The modeling error is given by*

$$e_{\text{MOD}} = \begin{cases} 0, & V^1(K_\delta, \mathcal{T}_h) \subset W_{\text{per}}^1(K_\delta), \delta/\varepsilon \in \mathbb{N}, \text{ and collocation,} \\ C_1 \delta, & V^1(K_\delta, \mathcal{T}_h) \subset W_{\text{per}}^1(K_\delta), \delta/\varepsilon \in \mathbb{N}, \\ C_2 \frac{\varepsilon}{\delta}, & V^1(K_\delta, \mathcal{T}_h) \subset H_0^1(K_\delta), \delta/\varepsilon \notin \mathbb{N}, \text{ and collocation,} \\ C_3 \left(\delta + \frac{\varepsilon}{\delta}\right), & V^1(K_\delta, \mathcal{T}_h) \subset H_0^1(K_\delta), \delta/\varepsilon \notin \mathbb{N}. \end{cases}$$

Proof. see [1, 3]. □

6.10.1 A priori error estimates in the fine scale region

In here, we will prove Theorem 6.7.4.

Proof of Theorem 6.7.4. Let $u_{\tilde{h}} \in V_D^1(\omega_1, \mathcal{T}_{\tilde{h}})$ be the FE approximation of the physical solution u^ε over the mesh $\mathcal{T}_{\tilde{h}}$, i.e. $u_{\tilde{h}} = u_{1,0,\tilde{h}} + v_{1,\tilde{h}}(I^{\tilde{h}}\gamma_1(u))$, where $I^{\tilde{h}}$ is the Lagrange interpolant on Γ_1 . Classical FE estimate holds; i.e.,

$$\|u^\varepsilon - u_{\tilde{h}}\|_{H^1(\omega^+)} \leq C\tilde{h}^s |u^\varepsilon|_{H^{s+1}(\omega_1)},$$

where the constant C is independent of H, h, \tilde{h} , and ε . Applying a triangular inequality, we obtain

$$\|\nabla(u^\varepsilon - \bar{u}_{\tilde{h}H})\|_{L^2(\omega^+)} \leq C\tilde{h}^s |u^\varepsilon|_{H^{s+1}(\omega_1)} + \|\nabla(u_{\tilde{h}} - \bar{u}_{\tilde{h}H})\|_{L^2(\omega^+)}.$$

The numerical solution $\bar{u}_{\tilde{h}H}$ over ω^+ is equal to the numerical fine scale solution $u_{1,\tilde{h}}(\theta_{1,\tilde{h}})$, it holds

$$B_1(u_{\tilde{h}} - u_{1,\tilde{h}}, v_{\tilde{h}}) = 0, \quad \forall v_{\tilde{h}} \in V_0^1(\omega_1, \mathcal{T}_{\tilde{h}}),$$

i.e., the difference $u_{\tilde{h}} - u_{1,\tilde{h}}(\theta_{1,\tilde{h}})$ is a^ε -harmonic in ω_1 and thus the discrete Caccioppoli inequality, Lemma 6.8.2, can be applied,

$$\|\nabla(u_{\tilde{h}} - u_{1,\tilde{h}}(\theta_{1,\tilde{h}}))\|_{L^2(\omega^+)} \leq \frac{C}{(\tau - \tau^+)} \|u_{\tilde{h}} - u_{1,\tilde{h}}(\theta_{1,\tilde{h}})\|_{L^2(\omega_1)},$$

where the constant $C > 0$ is independent of H, \tilde{h}, h , and ε , but depends on the ellipticity constants of the tensor a^ε . Consider an operator $P : V^1(\Gamma_1) \times V^1(\Gamma_2) \rightarrow V_D^1(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_D^1(\Omega \setminus \omega_1, \mathcal{T}_H)$ defined as

$$P(\mu_{1,\tilde{h}}, \mu_{2,H}) = \begin{cases} u_{1,0,\tilde{h}} + v_{1,\tilde{h}}(\mu_{1,\tilde{h}}), & \text{in } \omega_1, \\ u_{2,0,H} + v_{2,H}(\mu_{2,H}), & \text{in } \Omega \setminus \omega_1. \end{cases}$$

As in the continuous case, we decompose the operator P as $P = U_0 + Q$. Over ω_1 , it holds $u_{1,\tilde{h}}(\theta_{1,\tilde{h}}) = P(\theta_{1,\tilde{h}}, \theta_{2,H})$ and $u_{\tilde{h}} = P(I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon))$. Then,

$$\begin{aligned} \|u_{\tilde{h}} - u_{1,\tilde{h}}(\theta_{1,\tilde{h}})\|_{L^2(\omega_1)} &= \|P(I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon)) - P(\theta_{1,\tilde{h}}, \theta_{2,H})\|_{L^2(\omega_1)} \\ &\leq \|Q\| \|(I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon)) - (\theta_{1,\tilde{h}}, \theta_{2,H})\|_{L^*(\mathcal{U})}. \end{aligned}$$

As $(\theta_{1,\tilde{h}}, \theta_{2,H})$ are the discrete optimal virtual controls, they satisfy

$$\begin{aligned} &\int_{\omega_0} (v_{1,\tilde{h}}(\theta_{1,\tilde{h}}) - v_{2,H}(\theta_{2,H}))(v_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - v_{2,H}(\mu_{2,H})) dx \\ &= - \int_{\omega_0} (v_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - v_{2,H}(\mu_{2,H}))(u_{1,0,\tilde{h}} - u_{2,0,H}) dx, \end{aligned}$$

for all $(\mu_{1,\tilde{h}}, \mu_{2,H}) \in V^1(\Gamma_1) \times V^1(\Gamma_2)$. Then,

$$\begin{aligned} & \| (I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon)) - (\theta_{1,\tilde{h}}, \theta_{2,H}) \|_{L^*(\mathcal{U})} \\ &= \sup_{(\mu_{1,\tilde{h}}, \mu_{2,H})} \frac{|\pi((I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon)), (\mu_{1,\tilde{h}}, \mu_{2,H})) - \pi((\theta_1, \theta_2), (\mu_{1,\tilde{h}}, \mu_{2,H}))|}{\|(\mu_{1,\tilde{h}}, \mu_{2,H})\|_{L^*(\mathcal{U})}}, \end{aligned}$$

and following the proof of Lemma 6.5.2,

$$\begin{aligned} & \pi((I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon)), (\mu_{1,\tilde{h}}, \mu_{2,H})) - \pi((\theta_1, \theta_2), (\mu_{1,\tilde{h}}, \mu_{2,H})) \\ &= \int_{\omega_0} (u_{\tilde{h}} - u^H)(v_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - v_{2,H}(\mu_{2,H})) dx \\ &\leq \|u_{\tilde{h}} - u^H\|_{L^2(\omega_0)} \|(\mu_{1,\tilde{h}}, \mu_{2,H})\|_{L^*(\mathcal{U})}, \end{aligned}$$

where $u^H = u_{2,0,H} + v_{2,H}(I^H\gamma_2(u))$. We obtain that

$$\|(I^{\tilde{h}}\gamma_1(u^\varepsilon), I^H\gamma_2(u^\varepsilon)) - (\theta_{1,\tilde{h}}, \theta_{2,H})\|_{L^*(\mathcal{U})} \leq \|u_{\tilde{h}} - u^H\|_{L^2(\omega_0)},$$

and summarizing, we have

$$\|\nabla(u_{\tilde{h}} - u_{1,\tilde{h}}(\theta_{1,\tilde{h}}))\|_{L^2(\omega^+)} \leq C \|u_{\tilde{h}} - u^H\|_{L^2(\omega_0)}.$$

Then, we decompose the error into

$$\|u_{\tilde{h}} - u^H\|_{L^2(\omega_0)} \leq \|u_{\tilde{h}} - u^\varepsilon\|_{L^2(\omega_0)} + \|u^\varepsilon - u^0\|_{L^2(\omega_0)} + \|u^0 - u^H\|_{L^2(\omega_0)}, \quad (6.54)$$

provided that the solutions u^ε and u^0 are smooth enough, standard FE estimates and (6.27) can be applied to bound the first two quantities in (6.54), i.e.,

$$\|u_{\tilde{h}} - u^H\|_{L^2(\omega_0)} \leq C\tilde{h}^{s+1}|u^\varepsilon|_{H^{s+1}(\omega_1)} + C\varepsilon + \|u^0 - u^H\|_{L^2(\omega_0)}.$$

We bound the error in ω_0 by the error in ω_2

$$\|u^0 - u^H\|_{L^2(\omega_0)} \leq \|u^0 - u^H\|_{L^2(\omega_2)} \leq \|u^0 - u_H^0\|_{L^2(\omega_2)} + \|u_H^0 - u^H\|_{L^2(\omega_2)}.$$

The two norms corresponds to the DG-FE-HMM error in the L^2 norm and are given by Lemmas 6.10.2, 6.10.3, and 6.10.5.

6.10.2 A priori error estimates in the scale separated region

We prove an a priori error bound between u^ε and $\bar{u}_{\tilde{h}H}$ in $\Omega \setminus \omega^+$, where $\bar{u}_{\tilde{h}H}$ is defined in (6.44). For simplicity, we assume that $\delta = \varepsilon$ and choose periodic coupling conditions between the macro and micro problems. We recall that the reconstructed homogenized solution u_2^{rec} , and

its numerical approximation $u_{2,H}^{rec}$, are given by

$$u_2^{rec}(x) = u_2^0(x) + \varepsilon \sum_{j=1}^d \chi^j(x, x/\varepsilon) \frac{\partial u_2^0(x)}{\partial x_j}, \quad (6.55)$$

$$u_{2,H}^{rec}(x) = u_{2,H}(x) + \sum_{j=1}^d \psi_{K_\varepsilon}^{j,h}(x) \frac{\partial u_{2,H}(x)}{\partial x_j}, \quad (6.56)$$

where $u_2^0 = u_2^0(\theta_2)$ and $u_{2,H} = u_{2,H}(\theta_{2,H})$ are the exact solution and numerical solution of the coupling in ω_2 , respectively, and $\psi_{K_\varepsilon}^{j,h}$ are the micro solutions of (6.32). We sometimes use $u_2^{rec}(\theta_2)$ and $u_{2,H}^{rec}(\theta_{2,H})$ to emphasize the dependence on θ_2 and $\theta_{2,H}$, respectively.

We introduce the discrete micro problems on K_ε ; find u^h such that $u^h - u_{2,H} \in V^1(K_\varepsilon, \mathcal{T}_h)$ and

$$\int_{K_\varepsilon} a_2^\varepsilon(x) \nabla v^h \cdot \nabla z^h dx = 0, \quad \forall z^h \in V^1(K_\varepsilon, \mathcal{T}_h). \quad (6.57)$$

From assumption (H1), the tensor a_2^ε is constant in each macro element $K \in \mathcal{T}_H$. This simplifies the analysis as the modeling error is zero. We introduce a semi-discrete problem over ω_2 : find $\bar{u}_{2,H} \in V_D^1(\omega_2, \mathcal{T}_H)$ the solution of

$$\begin{aligned} \bar{B}_{2,H}(\bar{u}_{2,H}, w_H) &= F_2(w_H), \quad \forall w_H \in V^1(\omega_2, \mathcal{T}_H), \\ \bar{u}_{2,H} &= \theta_{2,H}, \text{ on } \Gamma_2, \end{aligned}$$

where the bilinear form $\bar{B}_{2,H} : V^1(\omega_2, \mathcal{T}_H) \times V^1(\omega_2, \mathcal{T}_H) \rightarrow \mathbb{R}$ is given by

$$\begin{aligned} \bar{B}_{2,H}(v_H, w_H) &= \sum_{K \in \mathcal{T}_H} \frac{|K|}{|K_\varepsilon|} \int_{K_\varepsilon} a_2^\varepsilon(x) \nabla v \cdot \nabla w dx + \sum_{e \in E} \int_e \mu_e \llbracket v_H \rrbracket \llbracket w_H \rrbracket ds \\ &\quad - \sum_{e \in E} \int_e \left(\overline{\{a_2^\varepsilon \nabla v\}} \llbracket w_H \rrbracket + \overline{\{a_2^\varepsilon \nabla w\}} \llbracket v_H \rrbracket \right) ds, \end{aligned}$$

where v and w are solutions of (6.57) in the exact Sobolev space $W(K_\varepsilon)$.

For a vector valued function η , we define the average of the multiscale fluxes as

$$\overline{\{\eta\}} = \frac{1}{2} \left(\frac{1}{|K_\varepsilon^+|} \int_{K_\varepsilon^+} \eta_+ dx + \frac{1}{|K_\varepsilon^-|} \int_{K_\varepsilon^-} \eta_- dx \right).$$

We can then define $\bar{u}_{2,H}^{rec}$ by

$$\bar{u}_{2,H}^{rec}(x) = \bar{u}_{2,H}(x) + \sum_{j=1}^d \psi_{K_\varepsilon}^j(x) \frac{\partial \bar{u}_{2,H}(x)}{\partial x_j}, \quad x \in K, \quad (6.58)$$

where $\bar{u}_{2,H} = \bar{u}_{2,H}(\theta_{2,H})$. We use $\bar{u}_{2,H}^{rec}(\theta_{2,H})$ to denote the dependence on $\theta_{2,H}$.

We now give the proof of Theorem 6.7.5.

Proof of Theorem 6.7.5. We decompose the error into

$$\|u^\varepsilon - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} \leq \|u^\varepsilon - u_2^{rec}(\theta_2)\|_{\tilde{H}^1(\Omega \setminus \omega^+)} + \|u_2^{rec}(\theta_2) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)}.$$

From Theorem 6.5.5, it holds that $\|u^\varepsilon - u_2^{rec}(\theta_2)\|_{\tilde{H}^1(\Omega \setminus \omega^+)} \leq C_1 \varepsilon^{1/2}$. We focus on $\|u_2^{rec}(\theta_2) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)}$ and follow [3, Section 3.3.3]. Using the triangular inequality, we obtain

$$\begin{aligned} \|u^\varepsilon - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 \varepsilon^{1/2} + \|u_2^{rec}(\theta_2) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} \\ &\leq C_1 \varepsilon^{1/2} + \|u_2^{rec}(\theta_2) - \bar{u}_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} \\ &\quad + \|\bar{u}_{2,H}^{rec}(\theta_{2,H}) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)}. \end{aligned}$$

Lemma 6.10.7 gives us

$$\begin{aligned} \|u_2^{rec}(\theta_2) - \bar{u}_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_3 H |u_2^0|_{H^2(\omega_2)} + C_4 \varepsilon \\ &\quad + \frac{C_5}{\tau^+} (\tilde{h}^{s+1} |u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2 |u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

Further, Lemma 6.10.9 provides us with

$$\|\bar{u}_{2,H}^{rec}(\theta_{2,H}) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} \leq C_2 \left(\frac{h}{\varepsilon} \right).$$

Collecting the previous results gives

$$\begin{aligned} \|u_2^{rec}(\theta_2) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 \varepsilon^{1/2} + C_2 \left(\frac{h}{\varepsilon} \right) + C_3 H |u_2^0|_{H^2(\omega_2)} \\ &\quad + \frac{C_5}{\tau^+} (\tilde{h}^{s+1} |u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2 |u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

Remark 6.10.6. Theorem 6.7.5 can be adapted for general tensor $a_2^\varepsilon(x)$ without a two-scale structure. In that case, the modeling error is present in the last term of the error.

Recall that we assumed periodic coupling with $\delta = \varepsilon$ and that (H1) and (H2) hold. Further, we assume Lipschitz continuity of the tensor in the first variable, i.e. $a_2(x, y) \in W^{1,\infty}(\bar{\omega}_2, L^\infty(Y))$.

Lemma 6.10.7. Let $u_2^{rec}(\theta_2)$ and $\bar{u}_{2,H}^{rec}(\theta_{2,H})$ be given by (6.55) and (6.58). Assume that $u_2^0 \in H^2(\omega_2)$, $u_1^\varepsilon \in H^{s+1}(\omega_1)$, with $s \leq 1$, and that the exact solutions of the micro problems (6.32) verify (H2). Then

$$\begin{aligned} \|u_2^{rec}(\theta_2) - \bar{u}_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 H |u_2^0|_{H^2(\omega_2)} + C_2 \varepsilon \\ &\quad + \frac{C_3}{\tau^+} (\tilde{h}^{s+1} |u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2 |u_2^0|_{H^2(\omega_2)}), \end{aligned}$$

where the constants are independent of H, \tilde{h}, h , and ε .

Chapter 6. Optimization based coupling method

Proof. Using the definitions of $u_2^{rec}(\theta_2)$ and $\bar{u}_{2,H}^{rec}(\theta_{2,H})$, it holds

$$\begin{aligned} \|u_2^{rec}(\theta_2) - \bar{u}_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)}^2 &= \sum_{K \in \mathcal{T}_H(\Omega \setminus \omega^+)} \|\nabla(u_2^{rec}(\theta_2) - \bar{u}_{2,H}^{rec}(\theta_{2,H}))\|_{L^2(K)}^2 \\ &\leq \sum_{K \in \mathcal{T}_H(\Omega \setminus \omega^+)} \|\nabla(u_2^0 - \bar{u}_{2,H})\|_{L^2(K)}^2 \\ &+ \sum_{K \in \mathcal{T}_H(\Omega \setminus \omega^+)} \left\| \sum_{j=1}^d \nabla(\varepsilon \chi^j(x, x/\varepsilon) \frac{\partial u_2^0}{\partial x_j} - \psi_{K_\varepsilon}^j(x) \frac{\partial \bar{u}_{2,H}}{\partial x_j}) \right\|_{L^2(K)}^2. \end{aligned}$$

Thanks to (H1), it holds $\varepsilon \chi^j(x, x/\varepsilon) = \psi_{K_\varepsilon}^j(x)$, and the second norm is bounded by the first norm plus a term $C\varepsilon$. We recall the bilinear form (6.53) for the problem (6.4) with a quadrature formula,

$$\begin{aligned} B_{2,H}^0(v_H, w_H) &= \sum_{K \in \mathcal{T}_H} |K| a_2^0(x_K) \nabla v_H \cdot \nabla w_H + \sum_{e \in E} \int_e \mu_e \llbracket v_H \rrbracket \llbracket w_H \rrbracket ds \\ &- \sum_{e \in E} \int_e (\{a_2^0(x_K, x/\varepsilon) v\} \llbracket w_H \rrbracket + \{a_2^0(x_K, x/\varepsilon) w\} \llbracket v_H \rrbracket) ds, \end{aligned}$$

and define $\hat{u}_{2,H}(\theta_{2,H}) \in V_D^1(\omega_2, \mathcal{T}_H)$ solution of

$$B_{2,H}^0(\hat{u}_{2,H}, w_H) = F_2(w_H), \quad \forall w_H \in V_0^1(\omega_2, \mathcal{T}_H).$$

By [3, Proposition 14], it holds that $\bar{u}_{2,H} = \hat{u}_{2,H}$. By hypothesis $u_2^0(\theta_2)$ and $\bar{u}_{2,H}(\theta_{2,H})$ have zero boundary conditions on $\partial\omega_2 \cap \Gamma$, and we can use [38, Lemmas 4.1, 4.2],

$$\begin{aligned} \|u_2^0(\theta_2) - \bar{u}_{2,H}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 \inf_{w \in V_D^1(\omega_2, \mathcal{T}_H), w=I^H\theta_2 \text{ on } \Gamma_2} \|u_2^0(\theta_2) - w\|_{\tilde{H}^1(\omega_2)} \\ &+ \frac{C_2}{\tau^+} \|u_2^0(\theta_2) - \bar{u}_{2,H}(\theta_{2,H})\|_{L^2(\omega_2)}. \end{aligned}$$

The first norm can be bounded by

$$\inf_w \|u_2^0(\theta_2) - w\|_{\tilde{H}^1(\omega_2)} \leq \|u_2^0(\theta_2) - u_{2,H}(I^H\theta_2)\|_{\tilde{H}^1(\omega_2)} \leq C_1 H |u_2^0|_{H^2(\omega_2)},$$

where $u_{2,H}(I^H\theta_2)$ is the FEM solution with an interpolation of θ_2 on Γ_2 . Following the proof of Theorem 6.7.4, the second part is bounded by

$$\begin{aligned} \|u_2^0(\theta_2) - \bar{u}_{2,H}(\theta_{2,H})\|_{L^2(\omega_2)} &\leq \|u_2^0(\theta_2) - u_{2,H}(I^H\theta_2)\|_{L^2(\omega_2)} \\ &+ \|u_{2,H}(I^H\theta_2) - \bar{u}_{2,H}(\theta_{2,H})\|_{L^2(\omega_2)} \\ &\leq C_1 H^2 |u_2^0|_{H^2(\omega_2)} + \|Q(I^{\tilde{h}}\theta_1, I^H\theta_2) - Q(\theta_{1,\tilde{h}}, \theta_{2,H})\|_{L^2(\omega_2)} \\ &\leq C_1 H^2 |u_2^0|_{H^2(\omega_2)} + C_2 \|u_{1,\tilde{h}}(I^{\tilde{h}}\theta_1) - \bar{u}_{2,H}(I^H\theta_2)\|_{L^2(\omega_0)}, \end{aligned}$$

where we have used that $(\theta_{1,\tilde{h}}, \theta_{2,H})$ is the optimal couple of the discrete minimization problem

and that Q is bounded. Finally using the triangular inequality, we have

$$\begin{aligned} \|u_{1,\tilde{h}}(I^{\tilde{h}}\theta_1) - \bar{u}_{2,H}(I^H\theta_2)\|_{L^2(\omega_0)} &\leq \|u_{1,\tilde{h}}(I^{\tilde{h}}\theta_1) - u_1^\varepsilon(\theta_1)\|_{L^2(\omega_0)} \\ &\quad + \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)} \\ &\quad + \|u_2^0(\theta_2) - \bar{u}_{2,H}(I^H\theta_2)\|_{L^2(\omega_0)} \\ &\leq C(\tilde{h}^{s+1}|u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2|u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

Summarizing,

$$\begin{aligned} \|u_2^{rec}(\theta_2) - \bar{u}_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 H |u_2^0|_{H^2(\omega_2)} + C_2 \varepsilon \\ &\quad + \frac{C_3}{\tau^+} (\tilde{h}^{s+1}|u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2|u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

The result of the Lemma follows. \square

Remark 6.10.8. The proof of Lemma 6.10.7, can be generalized for functions with non homogeneous boundary conditions. This can be done by splitting the solutions into a function depending of the controls and a function independent of the controls. The proof follows the same lines.

Lemma 6.10.9. Let $\bar{u}_{2,H}^{rec}(\theta_{2,H})$ and $u_{2,H}^{rec}(\theta_{2,H})$ be defined by (6.58) and (6.56), respectively. Then

$$\|\bar{u}_{2,H}^{rec}(\theta_{2,H}) - u_{2,H}^{rec}(\theta_{2,H})\|_{\tilde{H}^1(\Omega \setminus \omega^+)} \leq C \left(\frac{h}{\varepsilon} \right).$$

Proof. Follows from [3, Section 3.3.3]. \square

6.11 Numerical experiments

In this section we present various numerical experiments to illustrate the convergence rates and the performance of our coupling method. To facilitate the numerical comparison, we assume that the meshes $\mathcal{T}_{\tilde{h}}$ and \mathcal{T}_H have the same finite elements in the overlap ω_0 . The implementations can be adapted to the case where the meshes are not equal in ω_0 , using interpolations between the two meshes. This is treated in Chapter 7.

Outline. In 6.11.1, we give the computational cost of the coupling method and the goal-oriented method in terms of macro and micro numbers of degrees of freedom (DOF). In 6.11.2, we assume that \mathcal{T}_H and $\mathcal{T}_{\tilde{h}}$ form a conform partition of the computational domain, and show the influence of ε and τ in the convergence rates. In 6.11.4 and 6.11.5 we take an elliptic problem with a crack and an elliptic problem with a singular source term, respectively. In 6.11.6, we consider a domain with a defect of size ε . We compare our coupling method with other global to local methods.

6.11.1 Computational costs of the methods

Here, we briefly comment the computational cost of the optimization based method and of the global to local method [95]. Both methods use the FE-HMM (or DG-FE-HMM) and the FEM.

Let N denote the total number of DOF of the initial triangulation over Ω , and N_{mic} denote the micro number of DOF to obtain the homogenized conductivity at the quadrature points of the macro mesh. Further, N_{ω_1} be the number of DOF of the fine triangulation in ω_1 , and $N_{\Omega \setminus \omega_1}, N_{\omega_2}$ be the number of DOF of the coarse triangulation over $\Omega \setminus \omega_1$ and $\omega_2 = \Omega \setminus \omega$, respectively.

For the classical global to local method, (DG-)FE-HMM provides us with a numerical homogenized solution u^H , which is used as boundary condition on Γ_1 and the total cost is $\mathcal{O}(N \cdot N_{mic}) + \mathcal{O}(N_{\omega_1})$.

For the optimization based method, we start by computing the numerical solutions $u_{1,0,\tilde{h}}$ and $u_{2,0,H}$, using FEM and (DG-)FE-HMM respectively. The cost is $\mathcal{O}(N_{\omega_1})$ for FEM and $\mathcal{O}(N_{\omega_2} \cdot N_{mic})$ for (DG-)FE-HMM. Then, we solve a saddle point problem with cost $\mathcal{O}(N_{\omega_1} + N_{\Omega \setminus \omega_1})$. We note that the cost of the optimization based method can further be reduced, see Chapter 7 and [15].

6.11.2 Influence of ε and τ in the convergence rates

In here, we conduct two experiments to see the influence of ε and τ in the convergence rates between the fine scale solution and the numerical solution obtained by the coupling. We consider an elliptic problem in $\Omega = [0, 1]^2$ with homogeneous Dirichlet boundary values and a right-hand side $f \equiv 1$. Further ω has width $1/8$ and is centred around $[1/2, 1/2]$.

In ω , we take a highly heterogeneous non-periodic tensor with oscillations at several non separated scales denoted by $a_1^\varepsilon(x_1, x_2)$ and, in $\Omega \setminus \omega$, we take a tensor with scale separation denoted by $a_2^\varepsilon(x_1, x_2)$ and with a locally periodic structure. At first, we take a tensor a_2^ε which is locally Y -periodic in the fast variable with period ε . The a priori error analysis for the FE-HMM with periodic tensors is well known and was used in the a priori error analysis of our coupling method given in Section 6.10. The analysis of the FE-HMM with general tensors with scale separation, has not yet been derived and the modeling error remains unknown. We thus test our coupling method to a problem where the tensor a_2^ε has scale separation but is not locally periodic in Y . We expect the presence of a modeling error due to the wrong boundary conditions used in the micro problems.

Influence of ε . We assume that the union of the meshes \mathcal{T}_H and $\mathcal{T}_{\tilde{h}}$ forms a conform partition of Ω , with large FE in $\Omega \setminus \omega_1$ and fine FE in ω , as illustrated in Figure 6.5a. We take different values of ε and see the influence of the locally periodic wavelength in the coupling strategy. We conduct the experiments for two tensors, see Figures 6.6a and 6.6b.

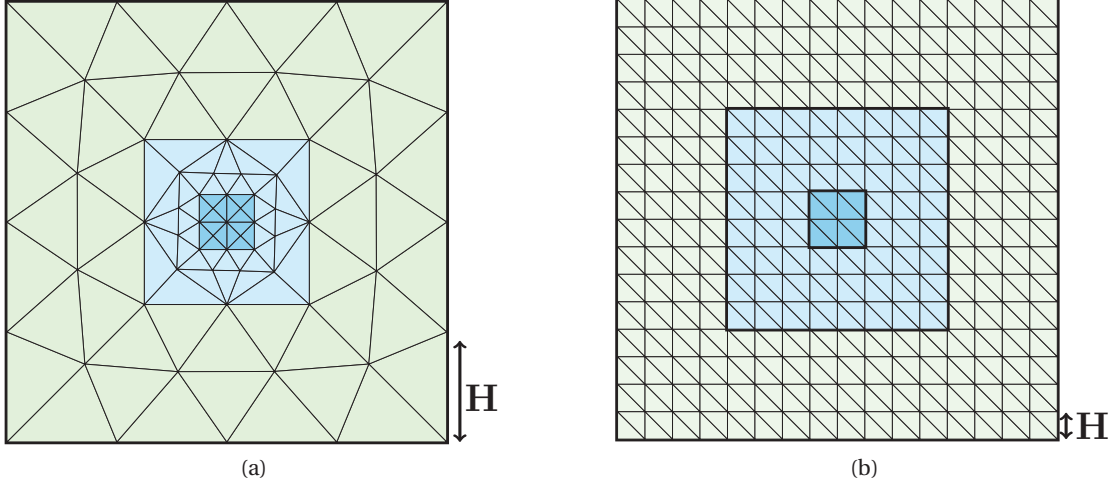


Figure 6.5 – Conform partitions over Ω for (a) $\tau = 1/8$, and (b) $\tau = 3/16$ with ω (in dark blue), ω_0 (in light blue), and $\Omega \setminus \omega_1$ (in light green).

Example 1. Let $a^\varepsilon(x_1, x_2)$ be given by

$$a_\omega^\varepsilon(x_1, x_2) = 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos \left(\left\lfloor 8 \left(ix_2 - \frac{x_1}{i+1} \right) \right\rfloor + [150ix_1] + [150x_2] \right)$$

$$a_2^\varepsilon(x_1, x_2) = \left(2.1 + \cos(2\pi x_1/\varepsilon) \cos(2\pi x_2/\varepsilon) + \sin(4x_1^2 x_2^2) \right) I.$$

Example 2. Let $a^\varepsilon(x_1, x_2)$ be given by

$$a_\omega^\varepsilon(x_1, x_2) = 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos \left(\left\lfloor 8 \left(ix_2 - \frac{x_1}{i+1} \right) \right\rfloor + [150ix_1] + [150x_2] \right)$$

$$a_2^\varepsilon(x_1, x_2) = \frac{1}{6} \left(\frac{1.1 + \sin(2\pi(x_1/\varepsilon)(x_2/\varepsilon))}{1.1 + \sin(2\pi x_2/\varepsilon)} + \sin(4x_1^2 x_2^2) + 2 \right) I.$$

The second tensor is locally periodic with period ε but not periodic in Y . Boundary layers are expected in the micro problems due to the wrong boundary conditions. With this tensor, we are able to test our coupling method to a problem with tensors with scale separation. In Figure 6.7, we illustrate the structure of the Y -periodic tensor $\sin(2\pi y_1)$ and the tensor $\sin(2\pi y_1 y_2)$, in the reference cell Y .

Consider an initial number of DOF per wavelength of $N_\varepsilon = 3$, with $\varepsilon = 1/8, 1/12$, and $\varepsilon = 3/16$, and refine the partition uniformly. The reference fine scale solution is obtained by the FEM over a very fine mesh. We use collocation in the tensors to the quadrature points. The mesh is uniformly refined and we expect the convergence rates to reach a threshold value depending on ε . As the tensor in the second example is not locally periodic, we expect a modeling error different than zero, whereas in the first example the modeling error is zero. In Figure 6.8a for the tensor of Example 1, and in Figure 6.8b, for the tensor of Example 2, we see the H^1 norm

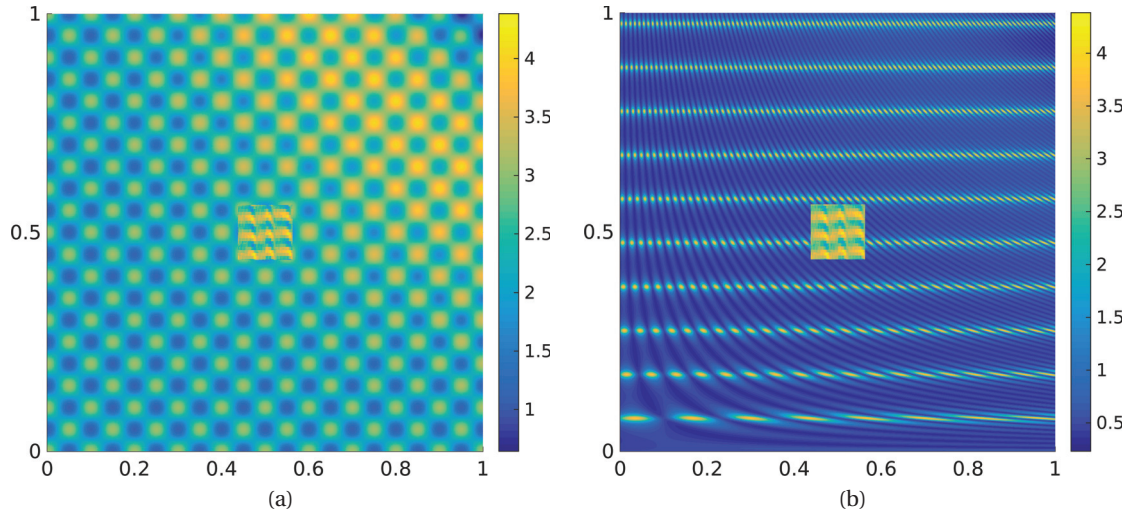


Figure 6.6 – Tensors with coefficients with and without scale separation for (a) the example 1, and (b) the example 2.

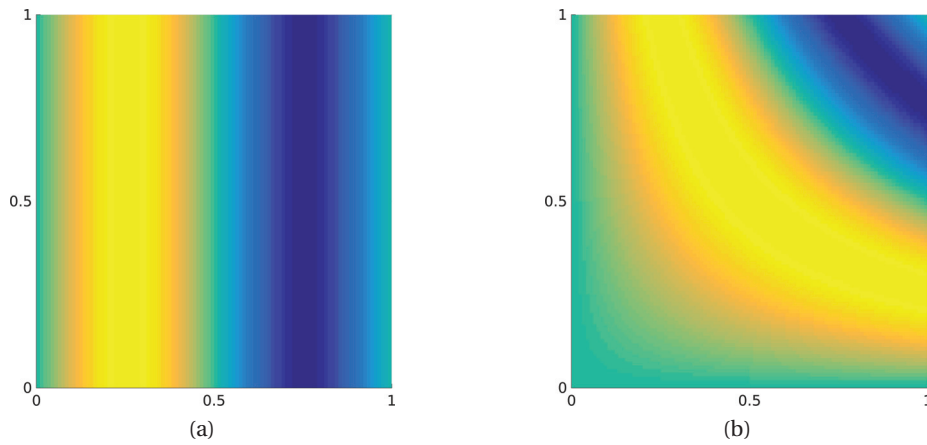


Figure 6.7 – (a) tensor $\sin(2\pi y_1)$ and (b) tensor $\sin(2\pi y_1 y_2)$ in the unit domain Y .

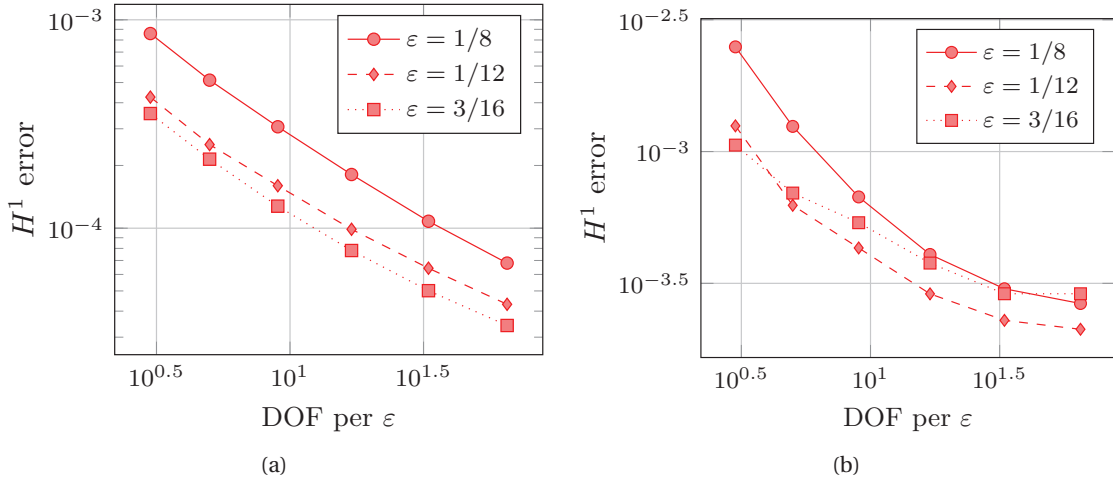


Figure 6.8 – H^1 error for different values of ε for the coupling with the tensor of Example 1(a) and Example 2(b).

between the reference solution and the numerical solution of the coupling, using P^1 macro and micro FE. The rates are similar for the three values of ε and they become smaller when ε is decreased. In Figure 6.8b, the saturation in the errors in ω is more pronounced due to the presence of a modeling error.

Influence of τ . We start by considering the tensor of Example 1 with an initial mesh size of $H = 1/16$, and set $\varepsilon = 1/4$. Then we compute a numerical coupling solution for different values of τ and plot the H^1 error between the fine scale solution u^ε and the numerical solution in ω . As no modeling error is present, we take a larger value of ε to enhance the effect of τ in the error between the reference solution u^ε and the numerical coupling solution in ω . We fix the number of DOF of the initial mesh and uniformly refine all elements. The initial mesh is represented in Figure 6.5b. This refinement is not optimal as the fine micro problems are resolved around each macro quadrature points, leading to a computationally expensive method. For $\tau = 1/16, 1/8, 3/16$, and $\tau = 1/4$, the errors are plotted in Figure 6.9a. We see that when τ is made smaller, then the error deteriorates.

We then consider the tensor of Example 2 with an initial mesh size of $H = 1/16$, and set $\varepsilon = 1/10$. For $\tau = 1/16, 1/8, 3/16$, and $\tau = 1/4$, the errors are plotted in Figure 6.9b. The convergence rate deteriorates when τ goes to zero, as expected from the Caccioppoli constant. The effect of τ on the convergence rates is enhanced due to the presence of the modeling error.

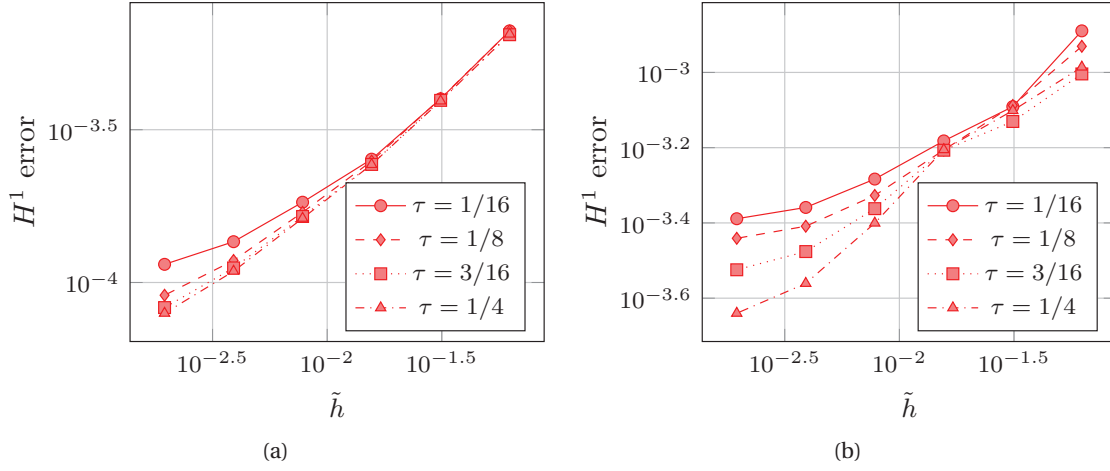


Figure 6.9 – H^1 error between the fine scale solution and the numerical coupling solution for different values of τ , using the tensors of (a) the Example 1 and (b) the Example 2.

6.11.3 Influence of micro HMM error

Consider an elliptic problem with a^ε given by

$$a_\omega^\varepsilon(x_1, x_2) = 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos \left(\left\lfloor 8 \left(ix_2 - \frac{x_1}{i+1} \right) \right\rfloor + \lfloor 150ix_1 \rfloor + \lfloor 150x_2 \rfloor \right)$$

$$a_2^\varepsilon(x_1, x_2) = \left(2.1 + \cos(2\pi x_1/\varepsilon) \cos(2\pi x_2/\varepsilon) + \sin(4x_1^2 x_2^2) \right) I.$$

Let $\varepsilon = 1/20$, $x_c = [1/2, 1/2]$ be the center of Ω , and set $\omega = x_c + 1/16[-1, 1]^d$, $\tau = 1/16$, $\tilde{h} = 1/16$ and $H = 1/3$. We use collocation and periodic coupling with $\delta = \varepsilon$, leading to zero modeling error. We refine the mesh uniformly and compute the error rates for $h/\varepsilon = 1/4, 1/8, 1/16$, and $1/32$. In Figures 6.10a and 6.10b, we plot the H^1 and L^2 errors, respectively. We see that the error rates reach a threshold value depending on h and ε . As ε is fixed, we can see that the error, for different values of h , is made smaller until ε is bigger and dominates the errors.

6.11.4 A domain with a crack

Consider an elliptic boundary value problem in $\Omega = [0, 1]^2$,

$$-\operatorname{div}(a^\varepsilon(x)\nabla u) = 0, \text{ in } \Omega,$$

with Dirichlet boundary condition $u = \varphi$ on Γ , where $\varphi \in [0, 2\pi]$ is the angle measured counterclockwise from the axis $\{(x, 0.5) : x \geq 0\}$. We add free Neumann boundary condition on the crack $\{x \in \Omega : x_1 \geq 0, x_2 = 0.5\}$. The homogenization model might not be accurate around the crack. A mesh refinement of the coarse model around the crack may lead to coarse meshes with mesh size smaller than ε , hence it requires more work around the crack than the FEM

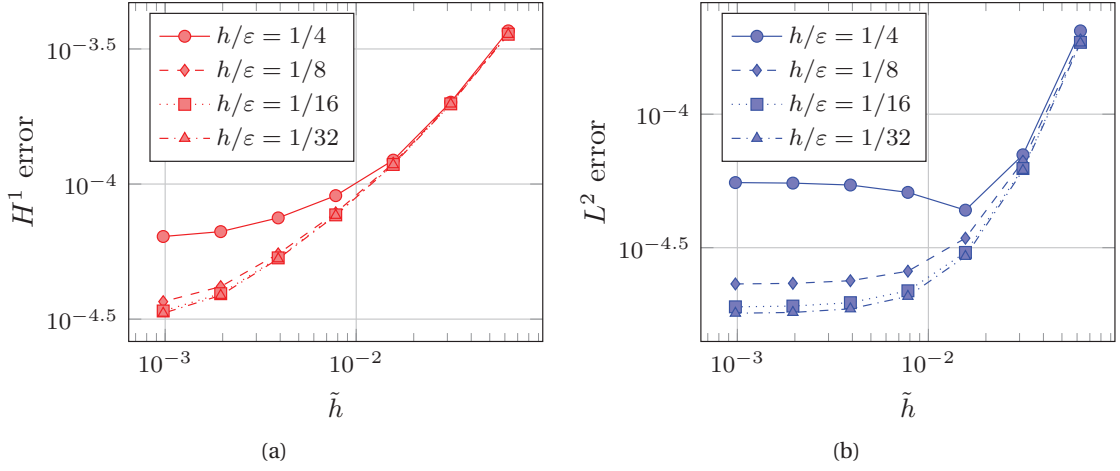


Figure 6.10 – (a) H^1 error and (b) L^2 error between the fine scale solution and the numerical coupling solution for different values of h/ε .

with scale resolution. For the treatment of crack problem with the FE-HMM, we cite [18]. We take a tensor a^ε — represented in Figure 6.11a for $\varepsilon = 1/10$ — with separation of scale and locally periodic in Y ,

$$a^\varepsilon(x_1, x_2) = \left(\frac{1}{(1.1 + \cos(2\pi \frac{x_1}{\varepsilon}))^2} + \frac{1}{(1.1 + \cos(2\pi \frac{x_2}{\varepsilon}))^2} \right)^{1/2}.$$

Let $x_c = [1/2, 1/2]$ be the center of Ω , and let $\omega_1 = x_c + \frac{1}{15}[-1, 1]^2$. The classical global to local numerical solution is the approximation of the following problem;

$$\begin{aligned} -\operatorname{div}(a^\varepsilon \nabla u) &= f, & \text{in } \omega_1, \\ u &= u^0, & \text{on } \Gamma_1, \end{aligned} \tag{6.59}$$

where u^0 is the homogenized solution. Recall that $\omega \subset \omega_1$, and define $\omega = x_c + \frac{1}{30}[-1, 1]^2$. We compute the numerical homogenized solution u^H over Ω on the coarse initial mesh, and use the value of u^H as Dirichlet boundary condition on Γ_1 and solve problem (6.59) with a fine scale FEM.

We refine uniformly in ω_1 and as the mesh size in ω should be small enough to capture the microscopic scales of the problem, it would be prohibitive to compute the numerical homogenized solution at each iteration. The coupling and the classical global to local method are both performed on the same mesh, where the coarse mesh in $\Omega \setminus \omega_1$ is left unchanged. We then compare the numerical solution with a reference solution obtained with a FEM on a very fine mesh. The reference solution is shown in Figure 6.11d and the numerical optimization based coupling solution in Figure 6.11c. We plot the H^1 semi-norm for the two methods in Figure 6.11b. We see that the global to local method (in black) reaches a threshold value, as

expected due to the use of the numerical homogenized function u^H as Dirichlet data on Γ_1 .

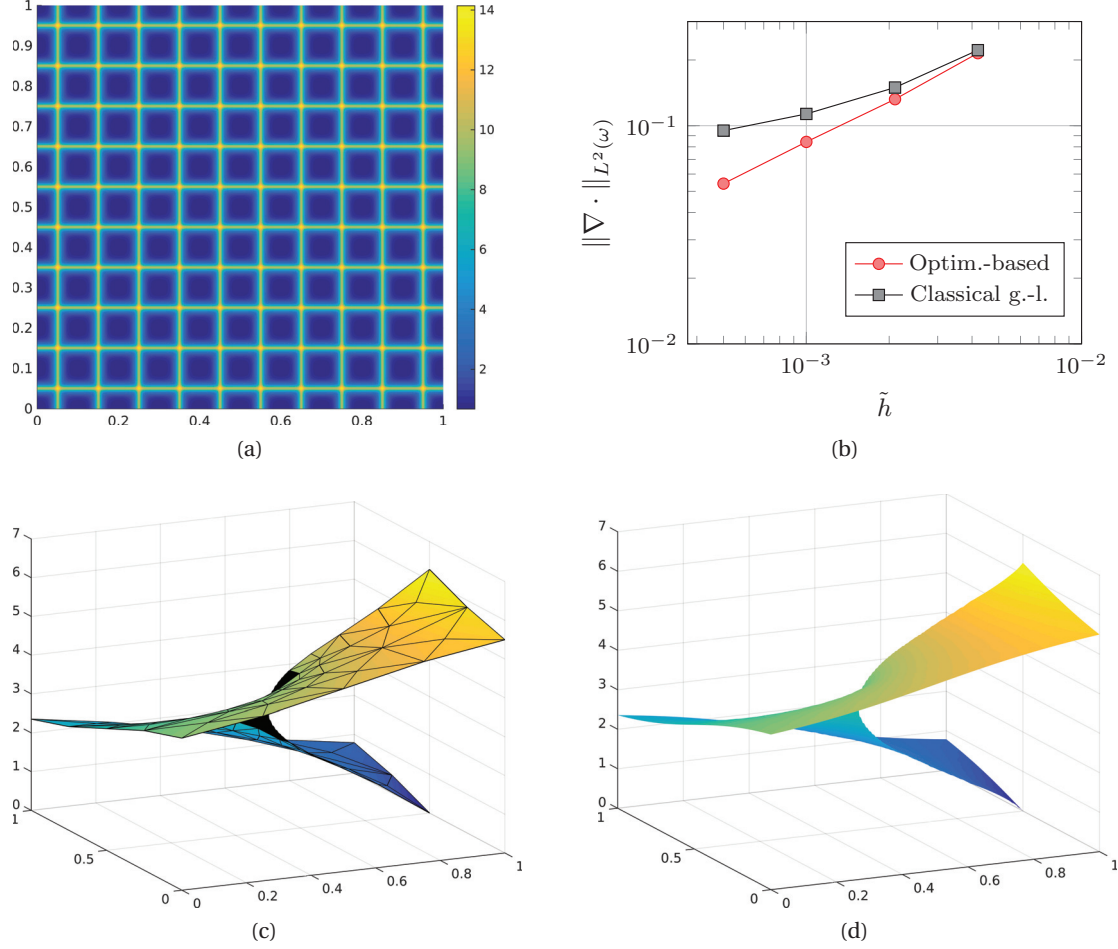


Figure 6.11 – Crack experiment: (a) tensor for $\varepsilon = 1/10$, (b) H^1 semi-norm in ω , for the optimization based coupling (bullet, red) and the classical coupling (square, black), (c) numerical optimization-based solution, and (d) reference solution.

6.11.5 Singular source term

In this experiment, we consider an elliptic problem with a singular source term given by random peaks. The tensor is assumed to have scale separation and is given by

$$a^\varepsilon(x) = \frac{1}{6} \left(\frac{1.1 + \sin\left(2\pi \frac{x_1}{\varepsilon} \frac{x_2}{\varepsilon}\right)}{1.1 + \sin\left(2\pi \frac{x_2}{\varepsilon}\right)} + \sin(4x_1^2 x_2^2) + 2 \right).$$

Depending on the location of the random peaks, the numerical homogenized right-hand side f^0 can be wrong, leading to an inaccurate approximation of u^0 . As in the crack experiments, we compute a numerical approximation of u^0 on a coarse initial mesh and then use it as boundary condition on Γ_1 . In Figure 6.12a we show the tensor for $\varepsilon = 1/25$. Let $x_c = [1/2, 1/2]$

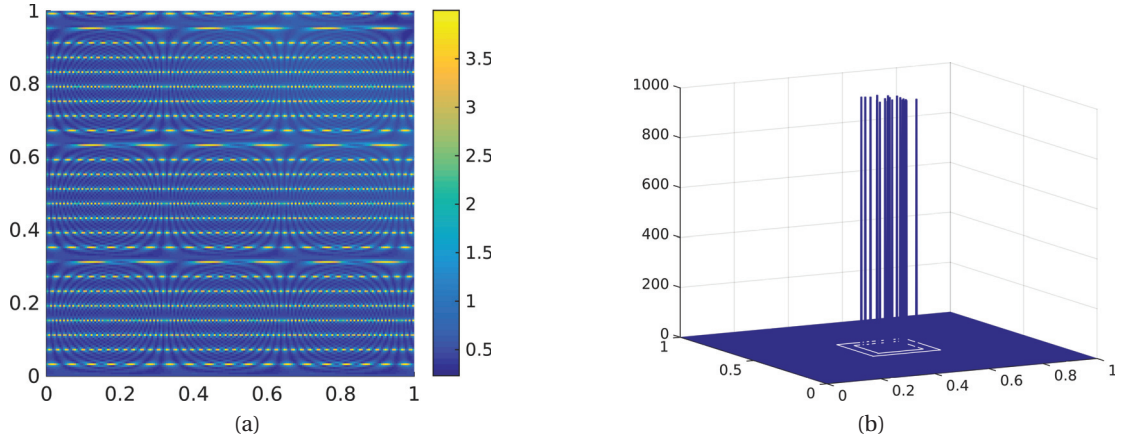


Figure 6.12 – Singular source term experiment: (a) tensor for $\varepsilon = 1/25$, (b) right hand side with 20 random peaks.

be the center of Ω , we set $\omega = x_c + \frac{1}{12}[-1, 1]^2$ and $\omega_1 = x_c + \frac{1}{4}[-1, 1]^2$. In Figure 6.12b, we illustrate the random source term f with 20 peaks. Figures 6.13a and 6.13b illustrate the reference solution and the optimization based solutions with the fine scale solution in ω and the coarse scale solution in $\Omega \setminus \omega$. The H^1 error to the reference solution, for $\varepsilon = 1/10$ and 100 random peaks, is shown in Figure 6.14, for the classical global to local method (in black) and the coupling (in red). While we observe a linear convergence rate for the optimization based method as predicted by Theorem 6.7.4, we see that the classical coupling leads to saturation in the error decay. This is due to inaccurate boundary conditions for the fine scale problems.

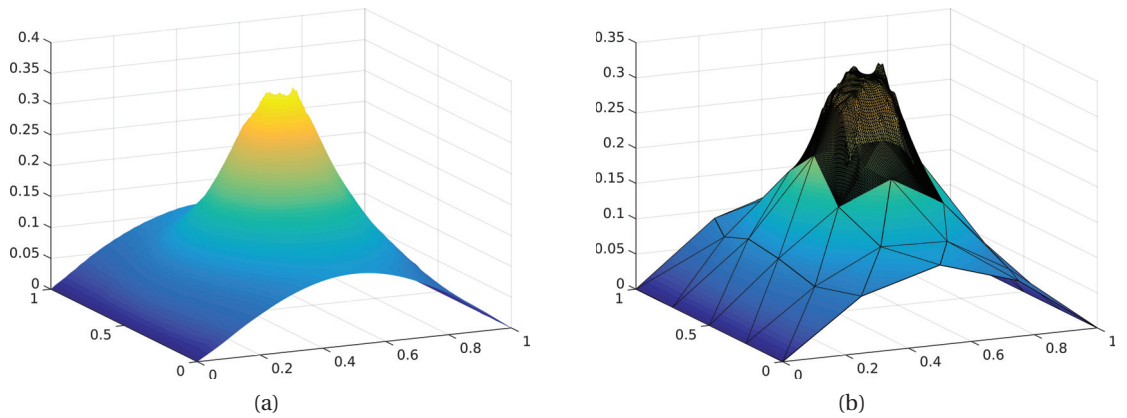


Figure 6.13 – Singular source term experiment: (a) reference solution, (b) optimization based numerical solution.

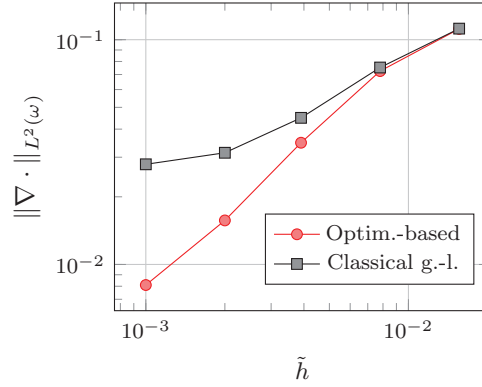


Figure 6.14 – Singular source term experiment: H^1 semi-norm in ω for the optimization based coupling (bullet, red) and the classical coupling (square, black).

6.11.6 A domain with a defect

We consider a homogenization problem with a local perturbation in the tensor, treated in [31]. The PDE is

$$\begin{aligned} -\operatorname{div}(a^\varepsilon \nabla u^\varepsilon) &= f, \text{ in } \Omega, \\ u^\varepsilon &= 0, \text{ on } \Gamma, \end{aligned}$$

where the tensor is of the form $a^\varepsilon = a_{per}^\varepsilon(x) + b^\varepsilon(x)$, with $a_{per}^\varepsilon(x) = a_{per}(x, x/\varepsilon)$ is (locally) periodic and $b^\varepsilon \in L^2(\Omega)^2$ is a local perturbation of size ε . A numerical homogenized solution u^H can be obtained with FE-HMM and produces a good approximation of u^ε in the L^2 norm. To obtain good approximation in the H^1 norm one needs to add correctors. However, the usual periodic cell problems are not valid as a^ε is not periodic. One could compute the periodic correctors corresponding to the tensor a_{per}^ε , and use them to correct the homogenized solution. This will be a good approximation at the large scale but will fail at the fine scale close to the defect. Following the approach in [31], a new corrector can be computed by adding a term to the periodic correctors as follows. Let $\chi^j \in W_{per}^1(Y)$ be the classical periodic correctors that satisfy the cell problems

$$\int_Y a_{per}^\varepsilon(x) \nabla \chi^j \cdot \nabla z dy = - \int_Y a_{per}^\varepsilon(x) e_j \cdot \nabla z dy, \quad \forall z \in W_{per}^1(Y).$$

Then, the additional term will be the solution of a Dirichlet boundary value problem in $K_n = [-n\varepsilon, n\varepsilon]^2$, where n is large enough so that the effect of the defect are negligible at the boundary of K_n . The problem reads: find $\chi_b^j \in H_0^1(K_n)$

$$\int_{K_n} a^\varepsilon(x) \nabla \chi_b^j \cdot \nabla z dx = - \int_{K_n} b^\varepsilon(x) (e_j + \nabla \chi^j) \cdot \nabla z dx, \quad \forall z \in H_0^1(K_n).$$

One can extend χ^j periodically to K_n and obtain a corrector $\tilde{\chi}^j(x) = \chi^j(x) + \chi_b^j(x)$ for all $x \in K_n$.

In this numerical example, we compute the FE-HMM solution and add to it either the periodic correctors χ or the modified correctors $\tilde{\chi}$. We then compare these two solutions with the optimization based solution presented in this paper. We will take similar oscillatory data as given in [31, Section 4.]. Let $\Omega = [-1, 1]^2$ and define

$$\begin{aligned} a_{per}^\varepsilon(x_1, x_2) &= 3 + \cos\left(2\pi \frac{x_1}{\varepsilon}\right) + \sin\left(2\pi \frac{x_2}{\varepsilon}\right), \\ b^\varepsilon(x_1, x_2) &= 10 \exp\left(-\left(\frac{x_1^2}{\varepsilon^2} + \frac{x_2^2}{\varepsilon^2}\right)\right), \\ f(x_1, x_2) &= \sin(\pi x_1) \cos(\pi x_2). \end{aligned}$$

We use a uniform triangular mesh and compute a reference solution on a very fine mesh. We compute the periodic correctors on $\mathcal{T}_h(Y)$ and extend it to $[-n\varepsilon, n\varepsilon]^2$ where n is sufficiently large. The terms χ_b are then computed on $[-n\varepsilon, n\varepsilon]^2$ with Dirichlet boundary conditions and adding them to χ , we obtain the non periodic correctors $\tilde{\chi}$. In each macro element K we defined a mesh $\mathcal{T}_{h_{rec}}(K)$, obtained by uniform refinement of K until the mesh size h_{rec} is smaller or equal to h . The two reconstructed solutions read

$$\begin{aligned} u_H^{\varepsilon, rec}(x) &= u^H(x) + \sum_{i=1}^d \varepsilon \chi^{i, h}(x, x/\varepsilon) \frac{\partial u^H(x)}{\partial x_i}, \\ \tilde{u}_H^{\varepsilon, rec}(x) &= u^H(x) + \sum_{i=1}^d \varepsilon \tilde{\chi}^{i, h}(x, x/\varepsilon) \frac{\partial u^H(x)}{\partial x_i}, \end{aligned}$$

where both correctors are defined on $[-n\varepsilon, n\varepsilon]^2$ with mesh size h and interpolated to $\mathcal{T}_{h_K}(K)$. In the coupling method, the fine scale region ω_1 will be centered around the defect, as its size is ε , we set $\omega = [-1/4, 1/4]^2$ and $\omega_1 = [-1/2, 1/2]^2$. The mesh size in ω_1 is equal to h_{rec} and the mesh size in the coarse region $\Omega \setminus \omega_1$ is H . We recall that the fine scale reference solution is given by

$$\bar{u}_{\tilde{h}H} = \begin{cases} u_{1, \tilde{h}}, & \text{in } \omega_+, \\ u_{2, H}^{rec}, & \text{in } \Omega \setminus \omega_+, \end{cases}$$

where we have chosen $\omega_+ = [-3/8, 3/8]^2$. We compute the error between the reference solution and the numerical solutions $u_H^{\varepsilon, rec}$, $\tilde{u}_H^{\varepsilon, rec}$, and $\bar{u}_{\tilde{h}H}$ in ω_1 and in $[-\varepsilon, \varepsilon]^2$. We first take $\varepsilon = 1/5$, $H = 1/16$ and a micro degree of freedom of $N_{micro} = \frac{1}{32^2}$.

We look at the relative error between the reference solution and the reconstructed solution $u_H^{\varepsilon, rec}$ (resp. $\tilde{u}_H^{\varepsilon, rec}$) for the periodic correctors (resp. non periodic),

$$\frac{\|\nabla(u^\varepsilon - u_H^{\varepsilon, rec})\|_{L^2(\omega_1)}}{\|\nabla u^\varepsilon\|_{L^2(\omega_1)}}.$$

As expected (see e.g. [31]), the errors with the two reconstructed solutions are similar in the far field, and one should look at the error around the defects to see the advantage of the correctors $\tilde{\chi}$. In Table 6.1, we see the relative errors for the three methods for $\varepsilon = 1/5$ and $\varepsilon = 1/10$. In

	Method	Rel. error in ω_1	Rel. error in $[-\varepsilon, \varepsilon]^2$
$\varepsilon = 1/5$	periodic correctors	0.436	1.589
	non-periodic correctors	0.396	0.992
	optimization based coupling	0.119	0.030
$\varepsilon = 1/10$	periodic correctors	0.281	1.076
	non-periodic correctors	0.260	0.720
	optimization based coupling	0.039	0.006

Table 6.1 – Relative error in ω_1 and $[-\varepsilon, \varepsilon]^2$, with $\varepsilon = 1/5$ and $\varepsilon = 1/10$, between the reference solution and the periodic, non-periodic reconstructed solution, and for the optimization based solution.

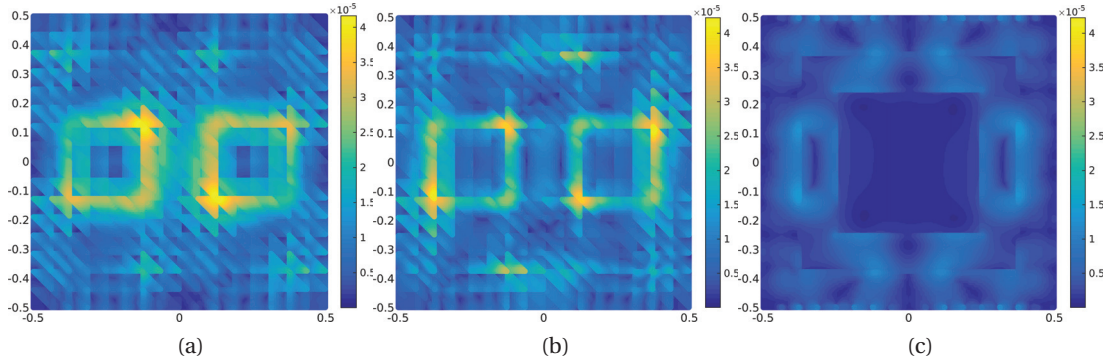


Figure 6.15 – Error in ω_1 between the reference solution and the numerical fine scale solution obtained with periodic correctors (a), non periodic correctors (b), and by the coupling (c).

Figure 6.15, we display the error in ω_1 between the reference solution and the numerical fine scale solutions obtained with the periodic correctors 6.15a, non-periodic correctors 6.15b, and the optimization based method 6.15c. While the errors between the periodic and non-periodic methods are similar in ω_1 , the difference is more important in $[-\varepsilon, \varepsilon]^2$, near the defect. There is however a significant improvement when the optimization based coupling method is used. This is to be expected as a fine scale solver is used in ω_1 and is coupled with a coarse scale solver. The strength of the method is that it produces a good H^1 approximation of the fine scale solution on Ω , but allows for a large mesh size H in $\Omega \setminus \omega_1$. We note that in [31], the same macro and micro number of degrees of freedom were used, with macro mesh size of $1/1000$ leading to a smaller discretization error and a larger difference between the periodic correctors and the non-periodic correctors. Setting H to such a small value is not necessary in our experiments as we only need a fine mesh in ω_1 and want to take full advantage of the homogenization techniques in the region with scale separation.

6.12 Summary

In this chapter, we gave the fully discrete optimization based coupling method, using the FEM for the fine scale region and the DG-FE-HMM in the coarse scale region. To find the numerical solution of the coupling, we derived an optimality system using Lagrange multipliers. The system is written as a saddle point problem, and its well-posedness can be proved using discrete Caccioppoli inequalities and a discrete strong Cauchy–Schwarz inequality. Fully discrete a priori error analysis and error estimates are given; i.e.,

$$\begin{aligned} \|u^\varepsilon - \tilde{u}_{\tilde{h}H}\|_{\tilde{H}^1(\omega^+)} &\leq C_1 \tilde{h}^s |u^\varepsilon|_{H^{s+1}(\omega_1)} + \frac{C_2}{\tau - \tau^+} (\tilde{h}^{s+1} |u^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + e_{\text{HMM}, L^2}), \\ \|u^\varepsilon - \tilde{u}_{\tilde{h}H}\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 \varepsilon^{1/2} + C_2 \left(\frac{h}{\varepsilon} \right) + C_3 H |u_2^0|_{H^2(\omega_2)} \\ &\quad + \frac{C_4}{\tau^+} (\tilde{h}^{s+1} |u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2 |u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

This chapter ends with numerical experiments. We verified that the method gives a good approximation of the fine scale solution in the region ω and compare our method with other global to local methods.

7 Improvement of meshing and coupling strategy

In this chapter, we present numerical improvements to our optimization based method derived in Chapter 6. We propose two numerical improvements to decrease the cost and computational time of the method, which do not affect the global efficiency of the method. This chapter is based on the article [15].

Let u^ε be the heterogeneous solution of

$$-\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) = f, \quad \text{in } \Omega, \quad (7.1)$$

with some boundary conditions on Γ , and where the tensor $a^\varepsilon \in (L^\infty(\Omega))^{d \times d}$ is highly oscillatory, bounded, and uniformly elliptic (6.2) with constants $0 < \lambda \leq \Lambda < +\infty$. Let $\omega \subset \Omega$ be a region without scale separation and ω_0 be the overlapping region. Assume that the boundaries $\Gamma_1 = \partial\omega_1 \setminus \Gamma$ and $\Gamma_2 = \partial\omega_2 \setminus \Gamma$ are Lipschitz continuous boundaries; Figure 7.1 shows a possible domain decomposition with Γ_1 (red) and Γ_2 (black) enhanced. The heterogeneous tensor a^ε of problem (7.1) is decomposed into $a^\varepsilon = a_\omega^\varepsilon + a_2^\varepsilon$, where $a_2^\varepsilon = a^\varepsilon \mathbb{1}_{\omega_2}$ and $a_\omega^\varepsilon = a^\varepsilon \mathbb{1}_\omega$ are tensors with and without scale separation, respectively. The tensor a_2^ε H -converges towards a homogenized tensor a_2^0 [90].

Outline. In Section 7.1, we start by deriving the coupling strategy for a new cost function that allows a reduction of the number of degrees of freedom. In Section 7.2, we state the discrete coupling method used in the numerics, give the algorithm and recall the a priori error estimates for the method. In Section 7.3 we give different meshing strategies and, finally, in Section 7.4, we propose various numerical examples that compare the coupling derived in Chapter 6 with the two novelties of this chapter; i.e,

- a cost function over the boundary $\Gamma_1 \cup \Gamma_2$ of the overlapping region;
- considering an interpolation between the fine and coarse meshes in the overlapping region.

The numerical experiments show that the total number of degrees of freedom and the cost of the method are reduced without affecting the efficiency of the method.

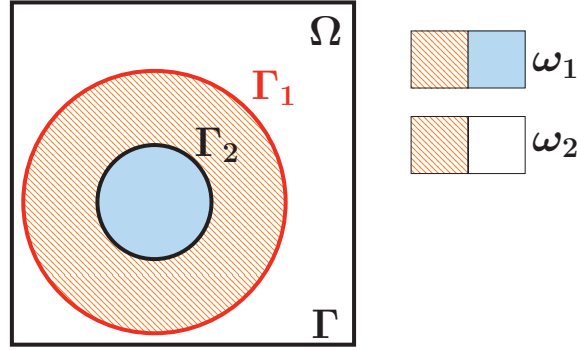


Figure 7.1 – Illustration of a domain decomposition with Γ_1 (red) and Γ_2 (black) enhanced, and where ω is in blue (full) and ω_0 in orange (hatched).

7.1 Optimization based coupling with minimization on $\Gamma_1 \cup \Gamma_2$

The optimization coupling method is based on the minimization of a cost involving the solutions of two partial differential equations over ω_1 and ω_2 . The heterogeneous control method restricted to Dirichlet boundary controls is given by the following problem: find $u_1^\varepsilon \in H^1(\omega_1)$ and $u_2^0 \in H^1(\omega_2)$, such that the cost functional

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2 \quad (7.2)$$

is minimized under the following constraints, for $i = 1, 2$,

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla u_i) &= f, & \text{in } \omega_i, \\ u_i &= \theta_i, & \text{on } \Gamma_i, \\ u_i &= g_D, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n \cdot (a_i(x)\nabla u_i) &= g_N, & \text{on } \partial\omega_i \cap \Gamma_N, \end{aligned} \quad (7.3)$$

where the boundary conditions θ_i , called the virtual controls, are to be determined. We will often use u_i to denote $u_i(\theta_i)$ and we set $a_1 = a^\varepsilon \mathbb{1}_{\omega_1}$, $u_1 = u_1^\varepsilon$, and $u_2 = u_2^0$. The strategy is to solve a minimization problem in the space of admissible controls, and as Dirichlet controls are considered, the space of admissible (Dirichlet) controls is

$$\mathcal{U}_i^D = \{\mu_i \in H^{1/2}(\Gamma_i) \mid \exists u \in H^1(\omega_i), u|_{\Gamma_i} = \mu_i, \text{ in the sense of the trace}\}.$$

For a discussion about admissible spaces and optimal controls, we refer to [85] and Sections 6.2 and 6.3. Again, we split the solutions u_1^ε and u_2^0 into

$$u_1^\varepsilon(\theta_1) = u_{1,0}^\varepsilon + v_1^\varepsilon(\theta_1), \quad u_2^0(\theta_2) = u_{2,0}^0 + v_2^0(\theta_2),$$

7.1. Optimization based coupling with minimization on $\Gamma_1 \cup \Gamma_2$

where $(v_1^\varepsilon(\theta_1), v_2^0(\theta_2))$ are called the state variables and satisfy, for $i = 1, 2$,

$$\begin{aligned} -\operatorname{div}(a_i(x)\nabla v_i) &= 0, & \text{in } \omega_i, \\ v_i &= \theta_i, & \text{on } \Gamma_i, \\ v_i &= 0, & \text{on } \partial\omega_i \cap \Gamma_D, \\ n \cdot (a_i(x)\nabla v_i) &= 0, & \text{on } \partial\omega_i \cap \Gamma_N, \end{aligned} \quad (7.4)$$

where $v_1 = v_1^\varepsilon$, and $v_2 = v_2^0$. The functions $u_{i,0}$ are solutions of problem (7.3) with zero controls on Γ_i , for $i = 1, 2$. We define a Hilbert space for the solutions v_i ,

$$H_D^1(\omega_i) = \{w_i \in H^1(\omega_i) \mid w_i = 0 \text{ on } \partial\omega_i \cup \Gamma_D, \text{ in the sense of the trace}\}.$$

The solutions $u_{1,0}^\varepsilon$ and $u_{2,0}^0$ exist and are unique, thanks to the Lax–Milgram lemma, and the solutions v_1^ε and v_2^0 can be uniquely determined if the controls θ_1 and θ_2 are known. As $u_{1,0}^\varepsilon$ and $u_{2,0}^0$ are independent of the virtual controls (θ_1, θ_2) , they can be computed beforehand.

The well-posedness of the optimization problem is proved following Lions [85] and Section 6.9. The key point consists in proving that the cost function induces a norm over $\mathcal{U} = (\mathcal{U}_1^D, \mathcal{U}_2^D)$. One consider then the completion of \mathcal{U} (still denoted by \mathcal{U}) with respect to the cost induced norm, and the minimization problem admits a unique solution $(\theta_1, \theta_2) \in \mathcal{U}$ satisfying the Euler–Lagrange formulation; i.e.,

$$\int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\theta_1) - v_2^0(\theta_2))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2))ds = - \int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2))(u_{1,0}^\varepsilon - u_{2,0}^0)ds, \quad (7.5)$$

for all $(\mu_1, \mu_2) \in \mathcal{U}$.

From the homogenization theory (H -convergence), we consider a family of problems (7.1) indexed by ε . In what follows, we will often assume $\varepsilon \leq \varepsilon_0$, where ε_0 is a parameter used in the strong Cauchy–Schwarz inequality, Lemma 6.3.3. We assume that $\theta_i \in \mathcal{U}_i^D$ and hence $u_i(\theta_i)$ is in $H^1(\omega_i)$, for $i = 1, 2$.

We write the cost in terms of the state variables v_1^ε and v_2^0 ; i.e.,

$$\begin{aligned} J(\theta_1, \theta_2) &= \frac{1}{2} \|v_1^\varepsilon(\theta_1) - v_2^0(\theta_2)\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2 + \|(v_1^\varepsilon(\theta_1) - v_2^0(\theta_2))(u_{1,0}^\varepsilon - u_{2,0}^0)\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2 \\ &\quad + \frac{1}{2} \|u_{1,0}^\varepsilon - u_{2,0}^0\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2, \end{aligned}$$

and set

$$\pi((\theta_1, \theta_2), (\mu_1, \mu_2)) = \int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\theta_1) - v_2^0(\theta_2))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2))ds. \quad (7.6)$$

The following Lemma proves that the form π is a scalar product over the space \mathcal{U} .

Lemma 7.1.1. *The bilinear form π given in (7.6) is a scalar product over \mathcal{U} .*

Proof. The symmetry and positivity are clear, and it remains to prove that the form is positive definite; $\pi(\theta_1, \theta_2) = 0$ if and only if $\theta_1 = 0$ and $\theta_2 = 0$. We use the short-hand notation $\pi(\theta_1, \theta_2)$ to denote $\pi((\theta_1, \theta_2), (\theta_1, \theta_2))$.

Assuming that θ_1 and θ_2 are zero, the state variables v_1^ε and v_2^0 are solutions of boundary value problems with zero data, thus v_1^ε and v_2^0 are zero over ω_1 and ω_2 respectively. This leads to $\pi(\theta_1, \theta_2) = 0$.

Assume now that $\pi(\theta_1, \theta_2) = 0$. It holds that

$$\int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\theta_1) - v_2^0(\theta_2))^2 ds = \int_{\Gamma_1} (\theta_1 - v_2^0(\theta_2))^2 ds + \int_{\Gamma_2} (v_1^\varepsilon(\theta_1) - \theta_2)^2 ds = 0,$$

and

$$\int_{\Gamma_1} (\theta_1 - v_2^0(\theta_2))^2 ds = 0, \quad \int_{\Gamma_2} (v_1^\varepsilon(\theta_1) - \theta_2)^2 ds = 0.$$

This implies that $v_1^\varepsilon(\theta_1)|_{\Gamma_1} = \theta_1 = v_2^0(\theta_2)|_{\Gamma_1}$ a.e., and $v_2^0(\theta_2)|_{\Gamma_2} = \theta_2 = v_1^\varepsilon(\theta_1)|_{\Gamma_2}$ a.e. As v_1^ε and v_2^0 are H^1 functions on ω_1 and ω_2 respectively, we obtain

$$\|\theta_1 - v_2^0(\theta_2)\|_{H^{1/2}(\Gamma_1)} = 0, \text{ and } \|v_1^\varepsilon(\theta_1) - \theta_2\|_{H^{1/2}(\Gamma_2)} = 0.$$

We now use H -convergence on the tensor a_1^ε , to obtain a homogenized tensor a_1^0 in ω_1 . It holds that v_1^ε converges weakly in H^1 towards v_1^0 the homogenized solution of

$$\begin{aligned} -\operatorname{div}(a_1^0(x) \nabla v_1^0) &= 0, & \text{in } \omega_1, \\ v_1^0 &= \theta_1, & \text{on } \Gamma_1, \\ v_1^0 &= 0, & \text{on } \partial\omega_1 \cap \Gamma_D, \\ n \cdot (a_1^0(x) \nabla v_1^0) &= 0, & \text{on } \partial\omega_1 \cap \Gamma_N. \end{aligned}$$

Using the compact embedding of L^2 in H^1 , the solution v_1^ε converges strongly, up to a subsequence, towards v_1^0 in L^2 , and it holds that

$$\|v_1^0 - v_2^0\|_{L^2(\Gamma_1 \cup \Gamma_2)} = \lim_{\varepsilon \rightarrow 0} \|v_1^\varepsilon - v_2^0\|_{L^2(\Gamma_1 \cup \Gamma_2)} = 0,$$

hence $v_1^0|_{\Gamma_2} = \theta_2 = v_2^0|_{\Gamma_2}$. Consequently it holds

$$\|v_1^0 - v_2^0\|_{H^{1/2}(\Gamma_1 \cup \Gamma_2)} = 0$$

Using the strong Cauchy–Schwarz lemma, Lemma 6.3.3, we obtain

$$\begin{aligned} \|v_1^0 - v_2^0\|_{H^1(\omega_0)}^2 &\geq \|v_1^0 - v_2^0\|_{L^2(\omega_0)}^2 \\ &\geq (1 - C_s)(\|v_1^0\|_{L^2(\omega_0)}^2 + \|v_2^0\|_{L^2(\omega_0)}^2), \end{aligned}$$

where $C_s < 1$ is the strong Cauchy-Schwarz constant. The tensor a_2^0 and a_1^0 are equal in the overlapping region ω_0 , due to the locality of H -convergence, the difference $v_1^0 - v_2^0$ satisfies

$$-\operatorname{div}(a_2^0 \nabla (v_1^0 - v_2^0)) = 0, \text{ in } \omega_0,$$

and one can bound the H^1 norm over ω_0 by the $H^{1/2}$ norm over its boundary $\Gamma_1 \cup \Gamma_2$; i.e.,

$$\|v_1^0 - v_2^0\|_{H^1(\omega_0)} \leq C \|v_1^0 - v_2^0\|_{H^{1/2}(\Gamma_1 \cup \Gamma_2)} = 0.$$

Collecting the results lead to $v_1^0 = 0$ and $v_2^0 = 0$ a.e. in ω_0 . Further from the Caccioppoli inequality, Lemma 6.3.2, it holds $v_1^0 = 0$ a.e. in ω_1 and $v_2^0 = 0$ a.e. in ω_2 . We can conclude that $\theta_i = 0$ a.e. in Γ_i , $i = 1, 2$, by using the trace inequality; i.e.,

$$\|\theta_i\|_{H^{1/2}(\Gamma_i)} \leq C \|v_i^0\|_{H^1(\omega_i)} = 0.$$

□

The norm induced from the scalar product π is given by

$$\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})} := \left(\int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2))^2 ds \right)^{1/2}, \quad \forall (\mu_1, \mu_2) \in \mathcal{U}. \quad (7.7)$$

7.1.1 A priori error analysis

Let u^ε be the solution of the heterogeneous problem (7.1), and let us derive a priori error bounds between u^ε and the solution of the coupling

$$\bar{u} = \begin{cases} u_1^\varepsilon(\theta_1), & \text{in } \omega^+, \\ u_2^{rec}(\theta_2), & \text{in } \Omega \setminus \omega^+, \end{cases} \quad (7.8)$$

where u_2^{rec} is the reconstructed homogenized solution u_2^0 with periodic correctors, and ω^+ is a subdomain of Ω such that $\omega \subset \omega^+ \subset \omega_1$. The term u_2^{rec} is given by

$$u_2^{rec}(x) = u_2^0(x) + \varepsilon \sum_{j=1}^d \chi^j(x, x/\varepsilon) \frac{\partial u_2^0(x)}{\partial x_j}, \quad x \in \Omega \setminus \omega^+,$$

where $u_2^0 = u_2^0(\theta_2)$.

For ε fixed, we define u^0 as the solution of

$$\begin{aligned} -\operatorname{div}(a_2^0(x)\nabla u^0) &= f, & \text{in } \omega_2, \\ u^0 &= \gamma_2(u^\varepsilon), & \text{on } \Gamma_2, \\ u^0 &= g_D, & \text{on } \partial\omega_2 \cap \Gamma_D, \\ n \cdot (a_2^0(x)\nabla u^0) &= g_N, & \text{on } \partial\omega_2 \cap \Gamma_N, \end{aligned} \quad (7.9)$$

where $\gamma_2 : H^1(\omega_2) \rightarrow H^{1/2}(\Gamma_2)$ denotes the trace operator on Γ_2 . Similarly, we define the trace operator γ_1 on Γ_1 . Assuming that the tensor a_2^ε is periodic in the fast variable, i.e., $a_2^\varepsilon(x) = a_2(x, x/\varepsilon) = a_2(x, y)$ is Y -periodic in y , where $Y = (0, 1)^d$, explicit equations are available to compute the homogenized tensor a_2^0

$$a_2^0(x) = \frac{1}{|Y|} \int_Y a_2(x, y) (I + \nabla \chi) dy,$$

where $\nabla \chi = (\nabla \chi^1, \dots, \nabla \chi^d)$ and I is the $d \times d$ identity matrix. The functions $\chi^j \in W_{per}^1(Y)$ are called the first order correctors and, for $j = 1, \dots, d$, χ^j is the unique solution of the cell problem

$$\int_Y a_2(x, y) \nabla \chi^j \cdot \nabla z dy = - \int_Y a_2(x, y) e_j \cdot \nabla z dy, \quad \forall z \in W_{per}^1(Y),$$

with periodic boundary conditions, and where $(e_i)_{i=1}^d$ denotes the canonical basis of \mathbb{R}^d . Assuming sufficient regularity on u^0 and on χ^j , it can be proved that

$$\|u^\varepsilon - u^0\|_{L^2(\omega_2)} \leq C\varepsilon, \quad (7.10)$$

where the constant is independent of ε . For proofs, we refer to [30, 80, 89].

Estimates for the fine solution

Let us define an operator $P : \mathcal{U} \rightarrow H^1(\omega_1) \times H^1(\Omega \setminus \omega_1)$ such that

$$P(\mu_1, \mu_2) \mapsto \begin{cases} u_1^\varepsilon(\mu_1), & \text{in } \omega_1, \\ u_2^0(\mu_2), & \text{in } \Omega \setminus \omega_1. \end{cases}$$

It can be split into $P = Q + U_0$, where $Q : \mathcal{U} \rightarrow H^1(\omega_1) \times H^1(\Omega \setminus \omega_1)$ is defined by

$$Q(\mu_1, \mu_2) \mapsto \begin{cases} v_1^\varepsilon(\mu_1), & \text{in } \omega_1, \\ v_2^0(\mu_2), & \text{in } \Omega \setminus \omega_1, \end{cases}$$

where the state variables v_1^ε and v_2^0 are solutions of (7.4) for $i = 1, 2$ respectively, and where U_0 is given by

$$U_0 = \begin{cases} u_{1,0}^\varepsilon, & \text{in } \omega_1, \\ u_{2,0}^0, & \text{in } \Omega \setminus \omega_1. \end{cases}$$

7.1. Optimization based coupling with minimization on $\Gamma_1 \cup \Gamma_2$

In the a priori error analysis conducted, in Section 6.5, with a cost functional over ω_0 , it has been shown that the operator Q is bounded in the operator norm, i.e.,

$$\|Q\| := \sup_{(\mu_1, \mu_2) \in \mathcal{U}} \frac{\|Q(\mu_1, \mu_2)\|_{L^2(\Omega)}}{\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}} \leq C.$$

Here we assume that Q is bounded for the norm in \mathcal{U} induced by the scalar product (7.7) for the cost function of (7.2).

Theorem 7.1.2. *Let u^ε be the solution of (7.1) and \bar{u} be given by (7.8). Assume that u^0 and χ^j are smooth enough for (7.10) to hold, and that $\|Q\| \leq C$. Then we have*

$$\|u^\varepsilon - \bar{u}\|_{H^1(\omega^+)} \leq C\varepsilon,$$

where the constant C depends on the constant of the Caccioppoli inequality, the bound $\|Q\|$, and the trace constants associated to the trace operators γ_1 and γ_2 on Γ_1 and Γ_2 , respectively.

The proof of Theorem 7.1.2 follows closely the proof for minimization problem with the cost functional over ω_0 given in Chapter 6.

Proof of Theorem 7.1.2. The difference $u^\varepsilon - \bar{u}$ is a_1^ε -harmonic in ω_1 , thus Caccioppoli inequality, Lemma 5.1.1, can be applied,

$$\begin{aligned} \|u^\varepsilon - \bar{u}\|_{H^1(\omega^+)} &\leq C \frac{1}{\tau} \|u^\varepsilon - \bar{u}\|_{L^2(\omega_1)} \\ &= C \frac{1}{\tau} \|P(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - P(\theta_1, \theta_2)\|_{L^2(\omega_1)} \\ &\leq \frac{C}{\tau} \|Q\| \|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})}. \end{aligned} \tag{7.11}$$

We next need to bound $\|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})}$.

Lemma 7.1.3. *Let u^ε and u^0 solve (7.1) and (7.9) respectively, and let $(\theta_1, \theta_2) \in \mathcal{U}$ be the optimal virtual controls. Then*

$$\|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})} \leq \|u^\varepsilon - u^0\|_{L^2(\Gamma_1 \cup \Gamma_2)}.$$

Proof. From the definition, it holds

$$\begin{aligned} &\|(\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)) - (\theta_1, \theta_2)\|_{L^*(\mathcal{U})} = \\ &\sup_{(\mu_1, \mu_2) \in \mathcal{U}} \frac{|\pi((\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)), (\mu_1, \mu_2)) - \pi((\theta_1, \theta_2), (\mu_1, \mu_2))|}{\|(\mu_1, \mu_2)\|_{L^*(\mathcal{U})}}. \end{aligned}$$

We look at the numerator. As the pair (θ_1, θ_2) minimizes the cost function J , the Euler–Lagrange

formulation (7.5) holds and

$$\begin{aligned}
 & \pi((\gamma_1(u^\varepsilon), \gamma_2(u^\varepsilon)), (\mu_1, \mu_2)) - \pi((\theta_1, \theta_2), (\mu_1, \mu_2)) = \\
 &= \int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\gamma_1(u^\varepsilon)) - v_2^0(\gamma_2(u^\varepsilon)))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) ds \\
 & \quad + \int_{\Gamma_1 \cup \Gamma_2} (v_1^\varepsilon(\mu_1) - v_2^0(\mu_2))(u_{1,0}^\varepsilon - u_{2,0}^0) ds \\
 &= \int_{\Gamma_1 \cup \Gamma_2} ((v_1^\varepsilon(\gamma_1(u^\varepsilon)) + u_{1,0}^\varepsilon) - (v_2^0(\gamma_2(u^\varepsilon)) + u_{2,0}^0))(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) ds \\
 &= \int_{\Gamma_1 \cup \Gamma_2} (u^\varepsilon - u^0)(v_1^\varepsilon(\mu_1) - v_2^0(\mu_2)) ds \leq \|u^\varepsilon - u^0\|_{L^2(\Gamma_1 \cup \Gamma_2)} \|(\mu_1, \mu_2)\|_{L^*(\mathcal{W})}.
 \end{aligned}$$

The result follows. \square

We obtain

$$\|u^\varepsilon - \bar{u}\|_{H^1(\omega^+)} \leq \frac{C}{\tau} \|Q\| \|u^\varepsilon - u^0\|_{L^2(\Gamma_1 \cup \Gamma_2)}.$$

The next Lemma gives an upper bound to the norm in Lemma 7.1.3.

Lemma 7.1.4. *Let u^ε and u^0 be the solutions of (7.1) and (7.9) respectively. Assume that u^0 and χ^j have enough regularity for (7.10) to hold. Then*

$$\|u^\varepsilon - u^0\|_{L^2(\Gamma_1 \cup \Gamma_2)} \leq C\varepsilon,$$

where the constant C is independent of ε .

Proof. It holds

$$\|u^\varepsilon - u^0\|_{L^2(\Gamma_1 \cup \Gamma_2)} \leq \|u^\varepsilon - u^0\|_{L^2(\Gamma_1)} + \|u^\varepsilon - u^0\|_{L^2(\Gamma_2)}.$$

Using the continuity of the traces, the first term can be bounded by

$$\|u^\varepsilon - u^0\|_{L^2(\Gamma_1)} \leq C \|u^\varepsilon - u^0\|_{L^2(\omega_2)} \leq C\varepsilon,$$

whereas the second term is zero because $u^0|_{\Gamma_2} = \gamma_2(u^\varepsilon) = u^\varepsilon|_{\Gamma_2}$. This prove the result. \square

The proof of Theorem 7.1.2 follows from (7.11), Lemmas 7.1.3, and 7.1.4.

Estimates for the coarse scale solution

The a priori error estimates to the coarse scale solver follows from Section 6.5 using Lemma 7.1.4. We skip the details, and state the result.

Theorem 7.1.5. *Let u^ε be the solution of (7.1) and $u_2^{rec}(\theta_2)$ be given by (6.25). Let $a_2(x, y) \in \mathcal{C}(\overline{\omega_2}; L_{per}^\infty(Y))$ and $\chi^j \in W_{per}(Y)$, $j = 1, \dots, d$. If in addition, $u^\varepsilon \in H^2(\Omega)$, $u_2^0(\theta_2) \in H^2(\omega_2)$, and $\chi^j \in W^{1,\infty}(Y)$, $j = 1, \dots, d$, it holds*

$$\|u^\varepsilon - u_2^{rec}(\theta_2)\|_{H^1(\Omega \setminus \omega^+)} \leq C\varepsilon^{1/2},$$

where the constant C is independent of ε , but depends on τ , τ^+ , and the ellipticity constants of a_2^ε .

7.2 Fully discrete coupling method

In this section, we describe the fully discrete overlapping coupling method, and perform an a priori error analysis for the minimization problem with the cost function (7.2). The discrete optimization based method couples a fine scale solver over ω_1 (FEM) with a coarse scale solver over ω_2 (FE-HMM or DG-FE-HMM). The fine scale solver over ω_1 requires a partition of size \tilde{h} sufficiently small to resolve the multiscale nature of the tensor. In contrast, the coarse scale solver on ω_2 takes full advantage of the scale separation and allows for a mesh size larger than the fine scale.

As the finite elements of the fine and coarse meshes in ω_0 are different, an interpolation between the two meshes should be considered. One can also chose to use the same finite elements in the overlap, leading to a discontinuity at Γ_1 in the mesh over ω_2 . In that latter situation, the discontinuous Galerkin FE-HMM [5] should be used instead of the FE-HMM.

In what follows, we consider for simplicity the problem (7.1) with homogeneous Dirichlet boundary conditions, i.e., we set $g_D = 0$ and $\Gamma_N = \emptyset$. Further, we assume that the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and its discrete version, Lemma 6.8.4, hold.

7.2.1 Numerical method for the fine scale problem.

Let $\mathcal{T}_{\tilde{h}}$ be a partition of ω_1 , in simplicial or quadrilateral elements, with mesh size $\tilde{h} \ll \varepsilon$ where $\tilde{h} = \max_{K \in \mathcal{T}_{\tilde{h}}} h_K$, and h_K is the diameter of the element K . In addition, we suppose that the family of partitions $\{\mathcal{T}_{\tilde{h}}\}$ is admissible (T1) and shape-regular (T2). For each partition $\mathcal{T}_{\tilde{h}}$ of the family $\{\mathcal{T}_{\tilde{h}}\}$, we define a FE space in ω_1

$$V_D^p(\omega_1, \mathcal{T}_{\tilde{h}}) = \{w \in H_D^1(\omega_1) \mid w|_K \in \mathcal{R}^p(K), \quad \forall K \in \mathcal{T}_{\tilde{h}}\}.$$

Further, $V_0^p(\omega_1, \mathcal{T}_{\tilde{h}})$ denotes the space of functions in $V_D^p(\omega_1, \mathcal{T}_{\tilde{h}})$ that vanish on $\partial\omega_1$.

Let $u_{1,\tilde{h}}$ be the numerical approximation of u_1^ε satisfying problem (7.3) for $i = 1$. We can split $u_{1,\tilde{h}}$ into $u_{1,\tilde{h}} = u_{1,0,\tilde{h}} + v_{1,\tilde{h}}$, where $v_{1,\tilde{h}} \in V_D^p(\omega_1, \mathcal{T}_{\tilde{h}})$ is obtained by the optimization method

and $u_{1,0,\tilde{h}} \in V_0^p(\omega_1, \mathcal{T}_{\tilde{h}})$ satisfies

$$B_1(u_{1,0,\tilde{h}}, w_{1,\tilde{h}}) = \int_{\omega_1} a_1 \nabla u_{1,0,\tilde{h}} \cdot \nabla w_{1,\tilde{h}} dx = F_1(w_{1,\tilde{h}}), \quad \forall w_{1,\tilde{h}} \in V_0^p(\omega_1, \mathcal{T}_{\tilde{h}}),$$

where F_1 is given by

$$F_1(w_{1,\tilde{h}}) = \int_{\omega_1} f w_{1,\tilde{h}} dx.$$

Thanks to the Poincaré inequality, the coercivity and boundedness of the bilinear form B_1 can be proved; the existence and uniqueness of $u_{1,0,\tilde{h}}$ follows.

7.2.2 Numerical method for the coarse scale problem.

Let $\{\mathcal{T}_H\}$ be a family of admissible (T1) and shape-regular (T2) partitions of ω_2 , with mesh size $H = \max_{K \in \mathcal{T}_H} h_K$. For each partition \mathcal{T}_H of the family $\{\mathcal{T}_H\}$, we define a FE space over ω_2

$$V_D^p(\omega_2, \mathcal{T}_H) = \{v \in H_D^1(\omega_2) \mid w|_K \in \mathcal{R}^p(K), \quad \forall K \in \mathcal{T}_H\},$$

and use $V_0^p(\omega_2, \mathcal{T}_H)$ to denote the set of functions of $V_D^p(\omega_2, \mathcal{T}_H)$ that vanish over $\partial\omega_2$.

We consider a macroscopic quadrature formula is given by the pair $\{x_{j,K}, \omega_{j,K}\}$ of quadrature nodes $x_{j,K}$ and weights $\omega_{j,K}$, for $j = 1, \dots, J$. The sampling domain of size δ around each quadrature point is denoted by $K_{\delta_j} = x_{j,K} + \delta[-1/2, 1/2]^d$. We assume that the quadrature formula verifies the necessary assumptions to guarantee that the standard error estimates for a FEM hold, see Section 2.1.

The numerically homogenized tensor $a_2^{0,h}(x_{j,K})$ is obtained using numerical solutions of micro problems defined in K_{δ_j} . In each sampling domain, we consider a partition \mathcal{T}_h in simplicial or quadrilateral elements K with mesh size $h = \max_{K \in \mathcal{T}_h} h_K$ satisfying $h \leq \varepsilon$. The micro FE space is

$$V^q(K_{\delta_j}, \mathcal{T}_h) = \{w^h \in W(K_{\delta_j}) \mid w|_K \in \mathcal{R}^q(K), \quad \forall K \in \mathcal{T}_h\},$$

where the space $W(K_{\delta_j})$ depends on the boundary conditions (2.21) or (2.22). The discrete micro problems read: find $\psi_{K_{\delta_j}}^{i,h} \in V^q(K_{\delta_j}, \mathcal{T}_h)$, $i = 1, \dots, d$, the solution of

$$\int_{K_{\delta_j}} a_2^\varepsilon(x) \nabla \psi_{K_{\delta_j}}^{i,h} \cdot \nabla z^h dx = - \int_{K_{\delta_j}} a_2^\varepsilon(x) e_i \nabla z^h dx, \quad \forall z^h \in S^1(K_{\delta_j}, \mathcal{T}_h).$$

The numerical homogenized solution $u_{2,H}$ is split into $u_{2,H} = u_{2,0,H} + v_{2,H}$, where $v_{2,H} \in V_D^p(\omega_2, \mathcal{T}_H)$ is given by the coupling and $u_{2,0,H} \in V_0^p(\omega_2, \mathcal{T}_H)$ is the solution off

$$B_{2,H}(u_{2,0,H}, w_{2,H}) = F_2(w_{2,H}), \quad \forall w_{2,H} \in V_0^p(\omega_2, \mathcal{T}_H),$$

where $B_{2,H}(\cdot, \cdot) : V_D^p(\omega_2, \mathcal{T}_H) \times V_D^p(\omega_2, \mathcal{T}_H) \rightarrow \mathbb{R}$ is given by

$$B_{2,H}(v_{2,H}, w_{2,H}) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{j,K} a_2^{0,h}(x_{j,K}) \nabla v_{2,H}(x_{j,K}) \cdot \nabla w_{2,H}(x_{j,K}),$$

and $F_2 : V_D^p(\omega_2, \mathcal{T}_H) \rightarrow \mathbb{R}$ given by

$$F_2(w_{2,H}) = \int_{\omega_2} f w_{2,H} dx.$$

7.2.3 Numerical Algorithm

In here, we state the discrete coupling and state the main convergence results. The well-posedness and the proofs of the error estimates follow from Section 6.10 and Section 7.1.

The solution $(u_{1,\tilde{h}}, u_{2,H}) \in V_D^p(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_D^p(\omega_2, \mathcal{T}_H)$ satisfies

$$\min_{\mu_{1,\tilde{h}}, \mu_{2,H}} \frac{1}{2} \|u_{1,\tilde{h}}(\mu_{1,\tilde{h}}) - u_{2,H}(\mu_{2,H})\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2 \text{ subject to } \begin{cases} B_1(u_{1,\tilde{h}}, w_{1,\tilde{h}}) = F_1(w_{1,\tilde{h}}), \\ B_{2,H}(u_{2,H}, w_{2,H}) = F_2(w_{2,H}), \end{cases}$$

for all $w_{1,\tilde{h}} \in V_0^p(\omega_1, \mathcal{T}_{\tilde{h}})$ and $w_{2,H} \in V_0^p(\omega_2, \mathcal{T}_H)$. We introduce discrete Lagrange multipliers for each of the constraint, and obtain a discrete optimality system:

find $(v_{1,\tilde{h}}, \lambda_{1,\tilde{h}}, v_{2,H}, \lambda_{2,H}) \in V_D^p(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_0^p(\omega_1, \mathcal{T}_{\tilde{h}}) \times V_D^p(\omega_2, \mathcal{T}_H) \times V_0^p(\omega_2, \mathcal{T}_H)$ satisfying

$$\int_{\Gamma_1 \cup \Gamma_2} (v_{1,\tilde{h}} - v_{2,H}) w_{1,\tilde{h}} ds - B_1(w_{1,\tilde{h}}, \lambda_{1,\tilde{h}}) = - \int_{\Gamma_1 \cup \Gamma_2} (u_{1,0,\tilde{h}} - u_{2,0,H}) w_{1,\tilde{h}} ds, \quad (7.12)$$

$$B_1(v_{1,\tilde{h}}, \xi_{1,\tilde{h}}) = F_1(\xi_{1,\tilde{h}}) - B_1(u_{1,0,\tilde{h}}, \xi_{1,\tilde{h}}), \quad (7.13)$$

$$\int_{\Gamma_1 \cup \Gamma_2} (v_{2,H} - v_{1,\tilde{h}}) w_{2,H} ds - B_{2,H}(w_{2,H}, \lambda_{2,H}) = \int_{\Gamma_1 \cup \Gamma_2} (u_{1,0,\tilde{h}} - u_{2,0,H}) w_{2,H} ds, \quad (7.14)$$

$$B_{2,H}(v_{2,H}, \xi_{2,H}) = F_2(\xi_{2,H}) - B_{2,H}(u_{2,0,H}, \xi_{2,H}), \quad (7.15)$$

for all $w_{1,\tilde{h}} \in V_D^p(\omega_1, \mathcal{T}_{\tilde{h}})$, $\xi_{1,\tilde{h}} \in V_0^p(\omega_1, \mathcal{T}_{\tilde{h}})$, $w_{2,H} \in V_D^p(\omega_2, \mathcal{T}_H)$, and $\xi_{2,H} \in V_0^p(\omega_2, \mathcal{T}_H)$.

The optimality system (7.12) to (7.15) can be written in matrix form as

$$\begin{pmatrix} M & -B^\top \\ B & 0 \end{pmatrix} U = G, \quad (7.16)$$

where the unknown vector U is given by $U = (v_{1,\tilde{h}}, v_{2,H}, \lambda_{1,\tilde{h}}, \lambda_{2,H})^\top$, and

$$M(\{v_{1,\tilde{h}}, v_{2,H}\}, \{w_{1,\tilde{h}}, w_{2,H}\}) = \begin{pmatrix} \int_{\Gamma_1 \cup \Gamma_2} v_{1,\tilde{h}} w_{1,\tilde{h}} ds & - \int_{\Gamma_1 \cup \Gamma_2} v_{2,H} w_{1,\tilde{h}} ds \\ - \int_{\Gamma_1 \cup \Gamma_2} v_{1,\tilde{h}} w_{2,H} ds & \int_{\Gamma_1 \cup \Gamma_2} v_{2,H} w_{2,H} ds \end{pmatrix},$$

$$B(\{v_{1,\tilde{h}}, v_{2,H}\}, \{\lambda_{1,\tilde{h}}, \lambda_{2,H}\}) = \begin{pmatrix} B_1(v_{1,\tilde{h}}, \lambda_{1,\tilde{h}}) & 0 \\ 0 & B_{2,H}(v_{2,H}, \lambda_{2,H}) \end{pmatrix}.$$

7.2.4 Fully discrete error estimates

The coupling solution, denoted by $\bar{u}_{\tilde{h}H}$, is defined as

$$\bar{u}_{\tilde{h}H} = \begin{cases} u_{1,\tilde{h}}(\theta_{1,\tilde{h}}), & \text{in } \omega^+, \\ u_{2,H}^{rec}(\theta_{2,H}), & \text{in } \Omega \setminus \omega^+, \end{cases} \quad (7.17)$$

where $u_{2,H}^{rec}(\theta_{2,H})$ is a fine scale approximation obtained from the coarse scale solution of the coupling, $u_{2,H}(\theta_{2,H})$, using a post-processing procedure in the following way. We assume that a_2^ε is Y -periodic in y and we restrict the FE spaces to piecewise FE spaces. Periodic coupling is used with sampling domains K_ε of size ε . The reconstructed solution $u_{2,H}^{rec}(\theta_{2,H})$ is given by

$$u_{2,H}^{rec}(x) = u_{2,H}(x) + \sum_{j=1}^d \psi_{K_\varepsilon}^{j,h}(x) \frac{\partial u_{2,H}}{\partial x_j}(x), \quad x \in K,$$

where $\psi_{K_\varepsilon}^{j,h}$ are the micro solutions of (6.32) in the sampling domain K_ε . As the numerical solutions might be discontinuous in ω_2 , we consider a broken H^1 semi-norm,

$$\|v\|_{\tilde{H}^1(\Omega)}^2 := \sum_{K \in \mathcal{T}_h(\omega^+)} \|\nabla v\|_{L^2(K)}^2 + \sum_{K \in \mathcal{T}_H(\Omega \setminus \omega^+)} \|\nabla v\|_{L^2(K)}^2.$$

We next state our main convergence result for the optimization based numerical solution. Let $u^H \in V_0^1(\omega_2, \mathcal{T}_H)$ be the FE-HMM approximation of the homogenized solution u^0 .

Theorem 7.2.1 (A priori error analysis in ω^+). *Let ε_0 be given by the strong Cauchy–Schwarz lemma, Lemma 6.3.3, and consider $\varepsilon \leq \varepsilon_0$. Let u^ε and u^0 be the exact solutions of problems (7.1) and (7.9), respectively, and $\bar{u}_{\tilde{h}H}$ be the numerical solution of the coupling (7.17). Further, let $u^H \in V_0^1(\omega_2, \mathcal{T}_H)$ be the FE-HMM approximation of u^0 . Assume $u^\varepsilon \in H^{s+1}(\Omega)$, with $s \leq 1$, $u^0 \in H^2(\omega_2)$, and assume that (7.10) holds, then*

$$\|u^\varepsilon - \bar{u}_{\tilde{h}H}\|_{\tilde{H}^1(\omega^+)} \leq C_1 \tilde{h}^s |u^\varepsilon|_{H^{s+1}(\omega_1)} + \frac{C_2}{\tau - \tau^+} (\tilde{h}^{s+1} |u^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + e_{HMM, L^2}),$$

where the constants are independent of ε , H , \tilde{h} , and h , and where $e_{HMM, L^2} = \|u^0 - u^H\|_{L^2(\omega_2)}$ is derived in Section 2.5.

7.3. Partitions used in the numerical coupling method

Proof. Follows the proof of Theorem 6.7.4 using a continuous macro FEM (FE-HMM) instead of a discontinuous Galerkin FEM (DG-FE-HMM). \square

Theorem 7.2.2 (Error estimates in $\Omega \setminus \omega^+$). *Let u^ε be the exact solution of problem (7.1) and $\bar{u}_{\tilde{h}H}$ be the numerical solution of the coupling (7.17). Let $a_2^\varepsilon(x) = a_2(x, x/\varepsilon)$, where $a_2(x, y)$ is Y -periodic in y and satisfies $a_2(x, y) \in \mathcal{C}(\bar{\omega}_2; L_{per}^\infty(Y))$. Let $\psi_{K_\varepsilon}^j(x) \in W_{per}^1(K_\varepsilon)$, $j = 1, \dots, d$. If in addition, $u^\varepsilon \in H^2(\Omega)$, $u_2^0(\theta_2) \in H^2(\omega_2)$, $u_1^\varepsilon \in H^{s+1}(\omega_1)$, with $s \leq 1$, and $\psi_{K_\varepsilon}^j(x) \in W^{1,\infty}(K_\varepsilon)$, $j = 1, \dots, d$. It holds,*

$$\begin{aligned} \|u_2^{rec}(\theta_2) - u_{2,H}^{rec}(\theta_2^H)\|_{\tilde{H}^1(\Omega \setminus \omega^+)} &\leq C_1 \varepsilon^{1/2} + C_2 \left(\frac{h}{\varepsilon} \right) + C_3 H |u_2^0|_{H^2(\omega_2)} \\ &\quad + \frac{C_4}{\tau^+} (\tilde{h}^{s+1} |u_1^\varepsilon|_{H^{s+1}(\omega_1)} + \varepsilon + H^2 |u_2^0|_{H^2(\omega_2)}). \end{aligned}$$

where the constants are independent of H, \tilde{h}, h , and ε .

Proof. Follows the lines of Theorem 6.7.5, where DG-FE-HMM is replaced by FE-HMM. \square

7.3 Partitions used in the numerical coupling method

The computational cost of the optimization based method relies on the total number of degrees of freedom in the discretization of the computational domain. The macro to micro coupling of the FE-HMM (or DG-FE-HMM) leads to a good approximation of the effective solution u^0 where the fine scales are needed only in small subdomains located around macro quadrature points. The advantage is that it allows for a macro partition of Ω with a mesh size much larger than the fine scales.

In the optimization based coupling method, the fine and coarse scale solutions u_1^ε and u_2^0 , satisfying problems (6.4), are defined on the overlapping region ω_0 . Let $\mathcal{T}_{\tilde{h}}$ and \mathcal{T}_H be a fine and coarse partitions of ω_1 and ω_2 respectively, and consider two FE spaces $V^p(\omega_1, \mathcal{T}_{\tilde{h}})$ and $V^p(\omega_2, \mathcal{T}_H)$. The coupling then requires

in ω_1 : a partition $\mathcal{T}_{\tilde{h}}$ with $\tilde{h} \leq \varepsilon$;

in ω_2 : a partition \mathcal{T}_H with $H \gg \varepsilon$, a quadrature formula $\{x_{j,K}, \omega_{j,K}\}$ and in each sampling domain K_{δ_j} , a partition \mathcal{T}_h with $h \leq \varepsilon$.

In the overlap ω_0 , we can either consider the same finite elements or use an interpolation between the meshes $\mathcal{T}_{\tilde{h}}$ and \mathcal{T}_H ;

in ω_0 : consider the same finite element K in the partitions $\mathcal{T}_{\tilde{h}}$ and \mathcal{T}_H . In that situation, discontinuous Galerkin FE-HMM is considered to avoid that either the fine partition $\mathcal{T}_{\tilde{h}}$ has elements K with $h_K \geq \varepsilon$, which would be too coarse, or the coarse partition \mathcal{T}_H has elements with $h_K \leq \varepsilon$, which would be unnecessarily small — see Figures 7.2a and 7.2b for illustrations;

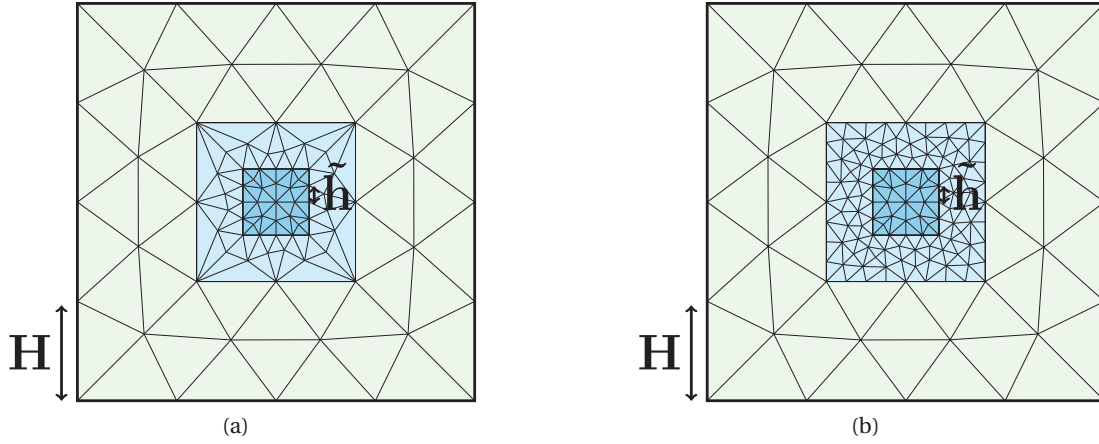


Figure 7.2 – (a) Conform and (b) non-conform discretizations of the computational domain Ω , with ω (in dark blue), ω_0 (in light blue), and $\Omega \setminus \omega_1$ (in light green).

in ω_0 : use an interpolation between $\mathcal{T}_{\tilde{h}}$ and \mathcal{T}_H .

In the numerical experiments, we are interested in the convergence rates between the heterogeneous solution u^ε and the numerical coupling solution $u_{1,\tilde{h}}$ in the domain ω , i.e., where the fine scales are not well separated. Then, the partition $\mathcal{T}_{\tilde{h}}$ in ω_1 is uniformly refined, and, if the FE in ω_0 are identical in the two meshes \mathcal{T}_H and $\mathcal{T}_{\tilde{h}}$, the number of DOF and the cost of the method will increase. Considering an interpolation between the two meshes in ω_0 implies that the number of DOF of the FE-HMM in ω_2 can be fixed which reduces the computational cost of the coupling method.

7.4 Numerical experiments

In this section, we give three numerical experiments that can be seen as a complement of the ones carried in Section 6.11 and in [16], where we focused on a minimization in $L^2(\omega_0)$, with interior subdomains and matching grids in the overlap ω_0 . We compare the convergence rates and computational costs of the minimization problem using the cost over ω_0 or the cost over $\Gamma_1 \cup \Gamma_2$, together with different partitions of the computational domain as explained in Section 7.3. We recall the two costs;

Case 1. Minimization in $L^2(\omega_0)$, with

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)}^2. \quad (\text{case 1})$$

Case 2. Minimization in $L^2(\Gamma_1 \cup \Gamma_2)$, with

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2. \quad (\text{case 2})$$







	matching grids	non-matching grids	Exp. 1
min. in ω_0	  		Exp. ②
min. on $\Gamma_1 \cup \Gamma_2$			Exp. ③

Table 7.1 – Summary of the discrete coupling methods used in the experiments.

Outline. In the first experiment given in 7.4.1, we consider the minimization over $L^2(\omega_0)$ and compare matching and non-matching meshes. The second experiment in 7.4.2 investigates the coupling with the cost function of case 2 over $\Gamma_1 \cup \Gamma_2$, and comparisons with the cost function of case 1 over ω_0 . In the last example given in 7.4.3, we combine non-matching grids and a minimization over the boundary. We observe order of magnitude of saving in computational cost when compared to the method proposed in [16]. Table 7.1 sums up the outline of the section.

7.4.1 Comparison of matching and non-matching grids on the overlap

In here, we use the cost function of case 1; i.e.,

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)}^2.$$

Using FEM and FE-HMM in ω_1 and ω_2 respectively, leads to two main restrictions; the mesh size in ω_1 should be smaller than the fine scale, whereas the mesh size in ω_2 can be larger than the fine scales, in order to take full advantage of the FE-HMM. Since both methods are defined in ω_0 , we can chose to have the same FE in both meshes on the overlap, or one can impose two different meshes. With the first choice, no interpolations must be considered between $\mathcal{T}_{\tilde{h}}$ and \mathcal{T}_H over ω_0 , but \mathcal{T}_H is composed of FE with mesh size as small as the fine scales. In that situation, DG-FE-HMM is chosen instead of FE-HMM due to the discontinuity at the interface Γ_1 . The second choice requires interpolation between the meshes in ω_0 , but \mathcal{T}_H is not restricted by the size of the fine mesh $\mathcal{T}_{\tilde{h}}$. We show that both cases give similar convergence rates, but the computational cost is significantly reduced in the second case.

Experiment 1. Let us consider a Dirichlet elliptic boundary value in $\Omega = [0, 1]^2$,

$$\begin{aligned} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, \text{ in } \Omega, \\ u &= 0, \text{ on } \Gamma, \end{aligned}$$

with $f \equiv 1$ and a^ε is given by

$$a_2^\varepsilon(x_1, x_2) = \frac{1}{6} \left(\frac{1.1 + \sin(2\pi(x_1/\varepsilon)(x_2/\varepsilon))}{1.1 + \sin(2\pi x_2/\varepsilon)} + \sin(4x_1^2 x_2^2) + 2 \right) I_2,$$

$$a_\omega^\varepsilon(x_1, x_2) = 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos(\lfloor 8(ix_2 - x_1/(i+1)) \rfloor + \lfloor 150ix_1 \rfloor + \lfloor 150x_2 \rfloor).$$

We use the notation a_ω^ε to emphasize that a fine scale solver should be used, although the fine scale structure in the tensor are not represented by a small coefficient ε .

Let x_c be the center of Ω , we consider $\omega_1 = x_c + [-1/4, 1/4]I_2$ and $\omega = x_c + [-1/8, 1/8]I_2$. Let $H = 1/8$, $\varepsilon = 1/10$, and a micro mesh size $h = \varepsilon/L$, so that the micro error is negligible. We initialize the fine mesh to $\tilde{h} = 1/16$. We use uniform simplicial meshes in ω_1 and ω_2 , and assume that $\mathcal{T}_{\tilde{h}}$ is obtained from \mathcal{T}_H using a uniform refinement in ω_0 . This allows simplification in the interpolation between the two meshes in the overlap. We couple the FEM over ω_1 with the mesh $\mathcal{T}_{\tilde{h}}(\omega_1)$ with the FE-HMM over ω_2 with mesh $\mathcal{T}_H(\omega_2)$, and compare it with a coupling between FEM over $\mathcal{T}_{\tilde{h}}(\omega_1)$ with DG-FE-HMM over a mesh composed of coarse FE from $\mathcal{T}_H(\omega_2 \setminus \omega_0)$ with small FE from the fine mesh $\mathcal{T}_{\tilde{h}}(\omega_0)$. The reference fine scale solution is computed on a very fine mesh, and we compare the two numerical solutions with the reference one. After three iterations, we plot the numerical approximations of the fine scale solution u_1^ε and coarse scale solution u_2^0 (in transparent), for a coupling with minimization of the cost function of case 1 with non-matching grid (Figure 7.3a) and with matching grids (Figure 7.3b). A zoom of the coarse scale solutions in the overlapping region ω_0 can be seen in Figure 7.3c for the coupling with non-matching grids and in Figure 7.3d (for the sake of readability, we plot the solutions after only one iteration) with matching grids, where the coupling is performed with the cost function of case 1.

We refine either only in ω_1 for the fine scale solver (non-matching grids) or in addition in ω_0 for the coarse scale solver (matching grids). We set $\delta = \varepsilon$ for the sampling domains, and consider a micro mesh size $h = \varepsilon/L$, so that the micro error is negligible. Figure 7.4a shows the H^1 norm in ω with non-matching grids (bullet) and with matching grids (diamond); we see that the errors are similar. We also measured the times, using Matlab timer, to compute the numerical solutions. We see in Figure 7.4b that using non-matching grids is faster as the number of micro problems, that have to be computed with the coarse solver, is smaller and fixed, whereas it increases when matching grids are used, causing a significant time overhead.

The rate of convergence in ω is influenced by H and ε , and when \tilde{h} is refined, we expect a saturation, depending on H and ε , in the convergence. Let $\varepsilon = 1/20$ and initialize the fine mesh to $\tilde{h} = 1/64$. We set $H = 1/8, 1/16$, and $1/32$, and refine \tilde{h} in each iteration. In Figure 7.5, we plot the H^1 norm between the reference and numerical solutions w.r.t the mesh size in ω . We see indeed that the error saturates at a threshold value that depends on H . For the influence of ε in the convergence rates, we refer to experiment 6.11.2.

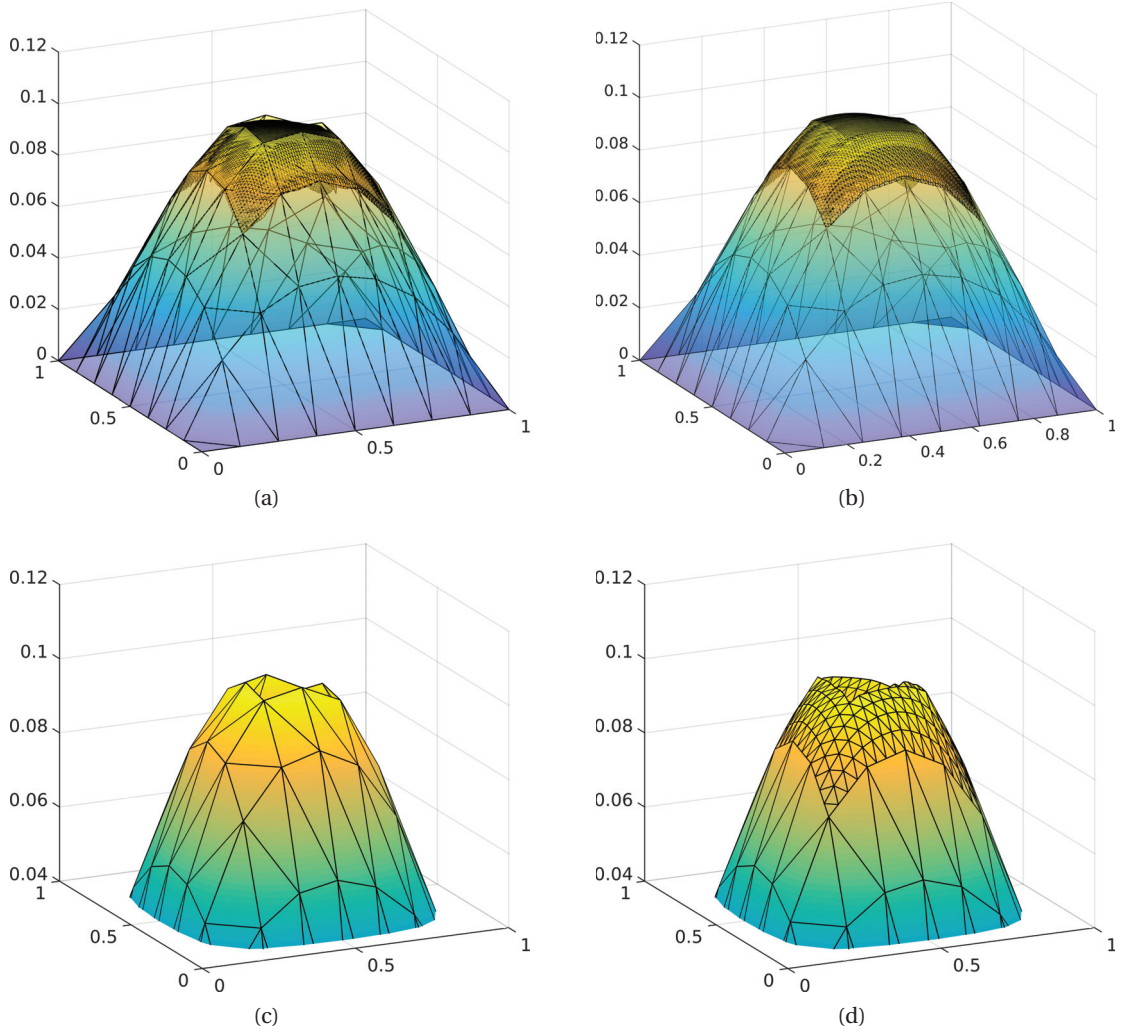


Figure 7.3 – experiment 1: numerical solutions of the coupling with minimization of the cost function of case 1 using non-matching grids (a) and matching grids (b), zoom in ω_0 of the coarse scale solution with the cost function of case 1 and non-matching grids (c) and with matching grids (d).

Experiment 2. Consider now an elliptic problem with a highly heterogeneous tensor a^ε given by

$$a_2^\varepsilon(x_1, x_2) = \frac{1}{6} \left(\frac{1.1 + \sin(2\pi(x_1/\varepsilon))}{1.1 + \sin(2\pi(x_2/\varepsilon))} + \sin(4x_1^2 x_2^2) + 2 \right) I_2,$$

$$a_\omega^\varepsilon(x_1, x_2) = 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos([8(ix_2 - x_1/(i+1))] + [150ix_1] + [150x_2]),$$

where the tensor a_2^ε is locally periodic in Y . We take the same settings as in Experiment 1, with $h/\varepsilon = 1/4$, and compare the couplings, with minimization over ω_0 , with matching grids and

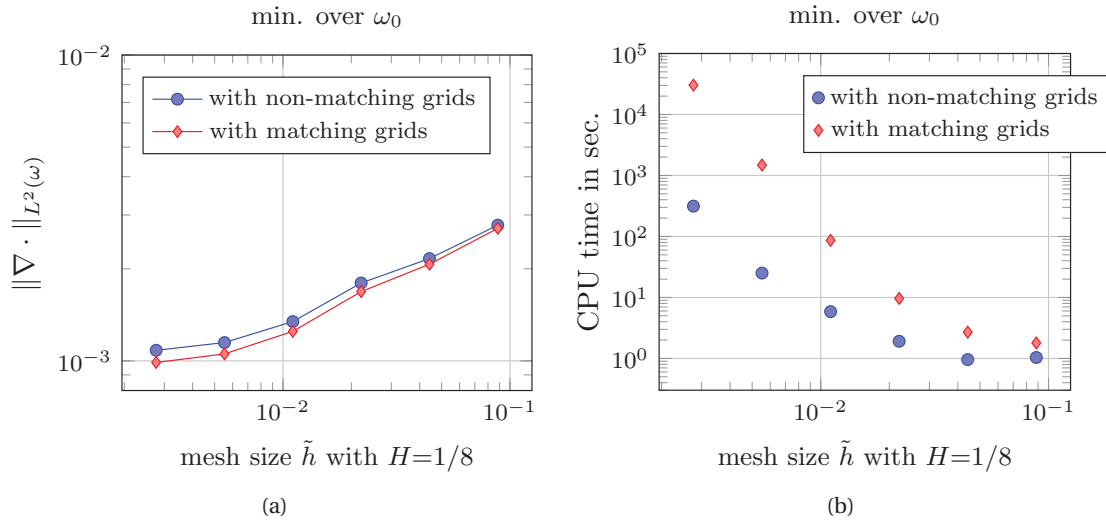


Figure 7.4 – experiment 1: (a) error in ω between the reference solution and the numerical solution of the coupling with minimization of the cost function of case 1 using non-matching grids (bullet, blue) and matching grids (diamond, red), (b) CPU time with the cost function of case 1 using non-matching grids (bullet, blue) and matching grids (diamond, red).

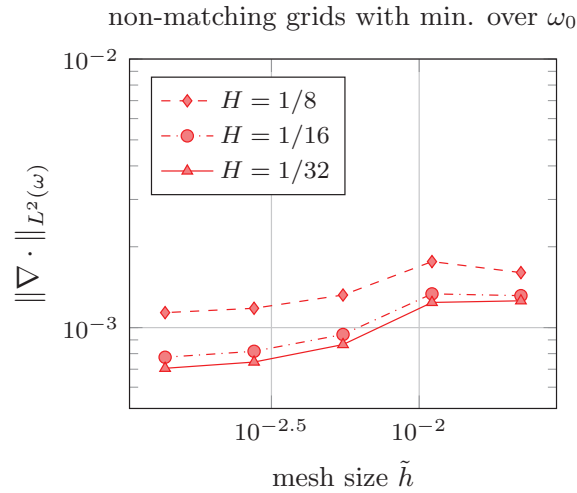
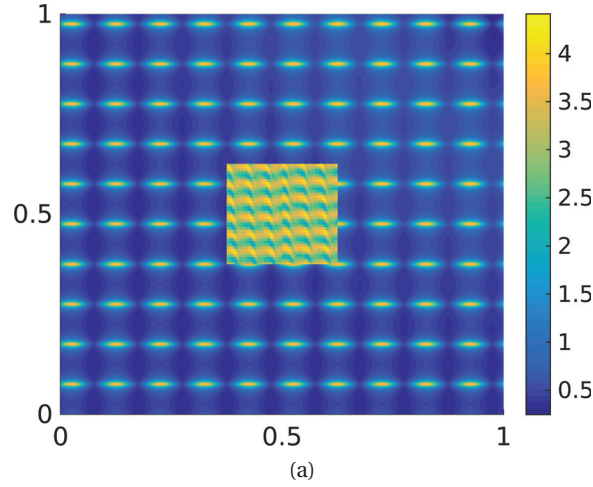


Figure 7.5 – experiment 1: (a) error in ω between the reference solution and the numerical solution of the coupling using non-matching grids and cost function of case 1 for different macro mesh size $H = 1/8$ (dashes, diamond), $H = 1/16$ (dash-dots, bullet), and $H = 1/32$ (full, triangle).

Figure 7.6 – Experiment 2: tensor a^ε for $\varepsilon = 1/10$.

non-matching grids. Further, we compare as well the coupling method with the classical global to local method, where the numerical solution satisfies a heterogeneous elliptic problem over ω_1 with Dirichlet boundary condition on Γ_1 given by a numerical homogenized solution. In Figure 7.6, we plot the tensor a^ε for $\varepsilon = 1/10$. In Figure 7.7a, we can see the H^1 semi-norm over ω for the coupling method with non-matching grids (blue) and matching grids (red) together with the classical coupling (green). The CPU time is plotted in Figure 7.7b.

7.4.2 Minimization with interface controls

For this experiment, we compare the coupling done with the cost function of case 1 and of case 2 on an elliptic problem with $\omega \subseteq \Omega$, i.e., when the boundaries of ω and Ω intersect.

Experiment 3. Let us consider a Dirichlet elliptic boundary value in $\Omega = [0, 1]^2$,

$$\begin{aligned} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, \text{ in } \Omega, \\ u &= 0, \text{ on } \Gamma, \end{aligned}$$

with $f \equiv 1$ and a^ε — plotted in Figure 7.8b — is given by

$$\begin{aligned} a_2^\varepsilon(x_1, x_2) &= (\cos(2\pi x_1/\varepsilon) + 2)I_2, \\ a_\omega^\varepsilon(x_1, x_2) &= 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos([\![8(ix_2 - x_1/(i+1))]\!] + [150ix_1] + [150x_2]). \end{aligned}$$

The tensor a_2^ε in ω_2 has scale separation, is Y -periodic in the fast variable, and the homoge-

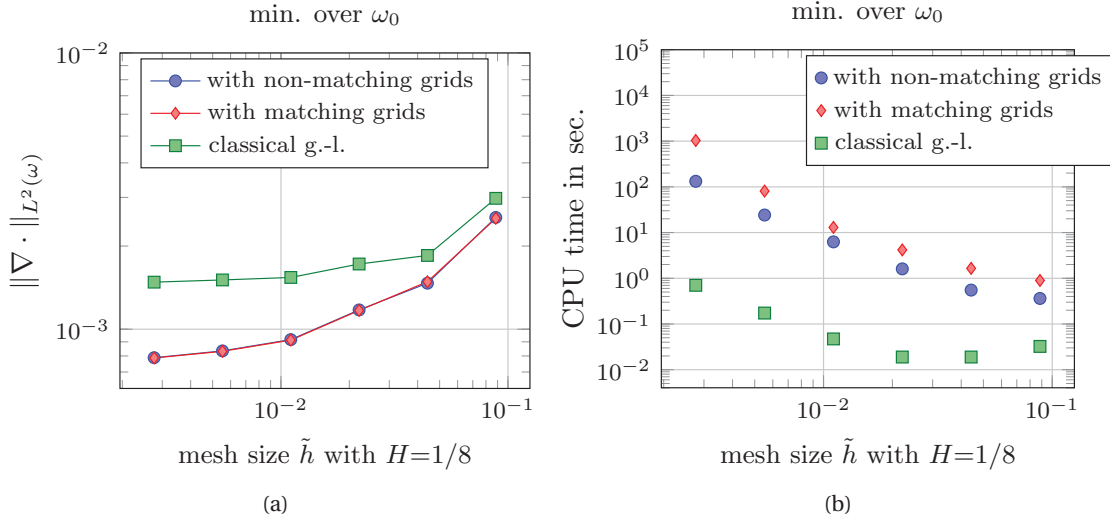


Figure 7.7 – Experiment 2: (a) error in ω between the reference solution and the numerical solution of the coupling with minimization of the cost function of case 1 using non-matching grids (bullet, blue) and matching grids (diamond, red), and together with the classical global to local method (square, green), (b) CPU time with the cost function of case 1 using non-matching grids (bullet, blue) and matching grids (diamond, red), and together with classical global to local method (square, green).

nized tensor a_2^0 can be explicitly derived as

$$a_2^0(x) = \begin{pmatrix} \left(\int_0^1 \frac{1}{a(y_1)} dy_1 \right)^{-1} & 0 \\ 0 & 2 \end{pmatrix}.$$

Let $\omega_1 = [0, 1/2] \times y$ and $\omega = [0, 1/4] \times y$, with $y \in [0, 1]$. An illustration of a numerical solution is given in Figure 7.9a. At first, we consider the cost of case 1,

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\omega_0)}^2.$$

Let $\varepsilon = 1/50$, and $h/\varepsilon = 1/L$ be small enough to neglect the micro error. We initialize the fine mesh to $\tilde{h} = 1/128$. For different macro mesh sizes $H = 1/8, 1/16, 1/32$ and $1/64$, we refine \tilde{h} and monitor the convergence rates between the numerical solution of the coupling and the reference solution. In Figure 7.8a, the H^1 norm is displayed for $H = 1/8$ (dots), $H = 1/16$ (dashes-dots), $H = 1/32$ (dashes) and $H = 1/64$ (full lines). One can see that the error saturate at a value depending on the macro mesh size H .

Now, we compare the cost of case 1 over ω_0 with the cost of case 2 over $\Gamma_1 \cup \Gamma_2$. We fix $\varepsilon = 1/10$, $H = 1/16$, and $h = \varepsilon/L$ small enough in order to neglect the micro error. We initialize the fine mesh to $\tilde{h} = 1/32$ and refine the mesh only in ω_1 . The numerical approximations of u_1^ε and u_2^0 are shown in Figure 7.9a, for the cost of case 1 over ω_0 , and in Figure 7.9b, for the cost of case 2

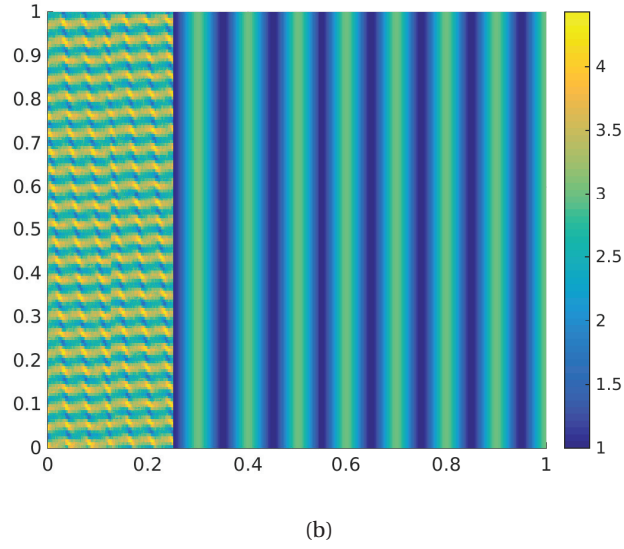
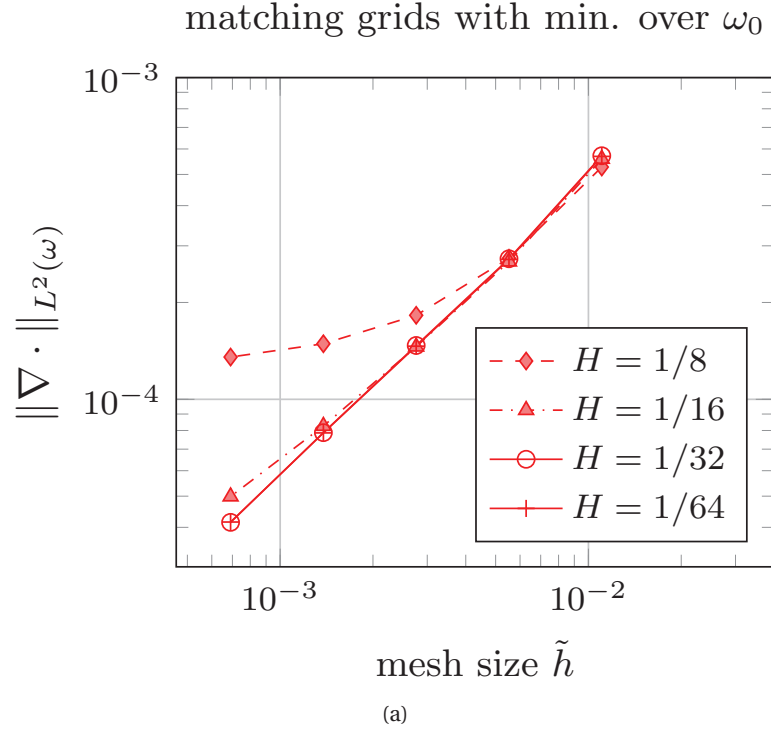


Figure 7.8 – experiment 3: (a) errors in ω between the reference solution and the numerical solution of the coupling using matching grids and the cost function of case 1 with different macro mesh size $H = 1/8$ (star), $H = 1/16$ (diamond), $H = 1/32$ (bullet), and $H = 1/64$ (plus), (b) tensor a^ε over Ω for $\varepsilon = 1/10$.

over $\Gamma_1 \cup \Gamma_2$. The H^1 and L^2 errors between u^H and a reference solution in ω_0 , are shown in Figures 7.9c and 7.9d, respectively. Computational times are compared as well in Figure 7.10, for the cost over ω_0 (diamond, blue) and the cost over $\Gamma_1 \cup \Gamma_2$ (bullet, red). As the number of degrees of freedom of the saddle point problem (7.16) is reduced when minimizing over the boundaries $\Gamma_1 \cup \Gamma_2$, we see that the coupling over ω_0 is more costly than the coupling over $\Gamma_1 \cup \Gamma_2$. Considering an interpolation between the two meshes in the interface ω_0 gives similar results as, due to the periodicity of a_2^ε , we need only to resolve one cell problem to compute the homogenized tensor a_2^0 .

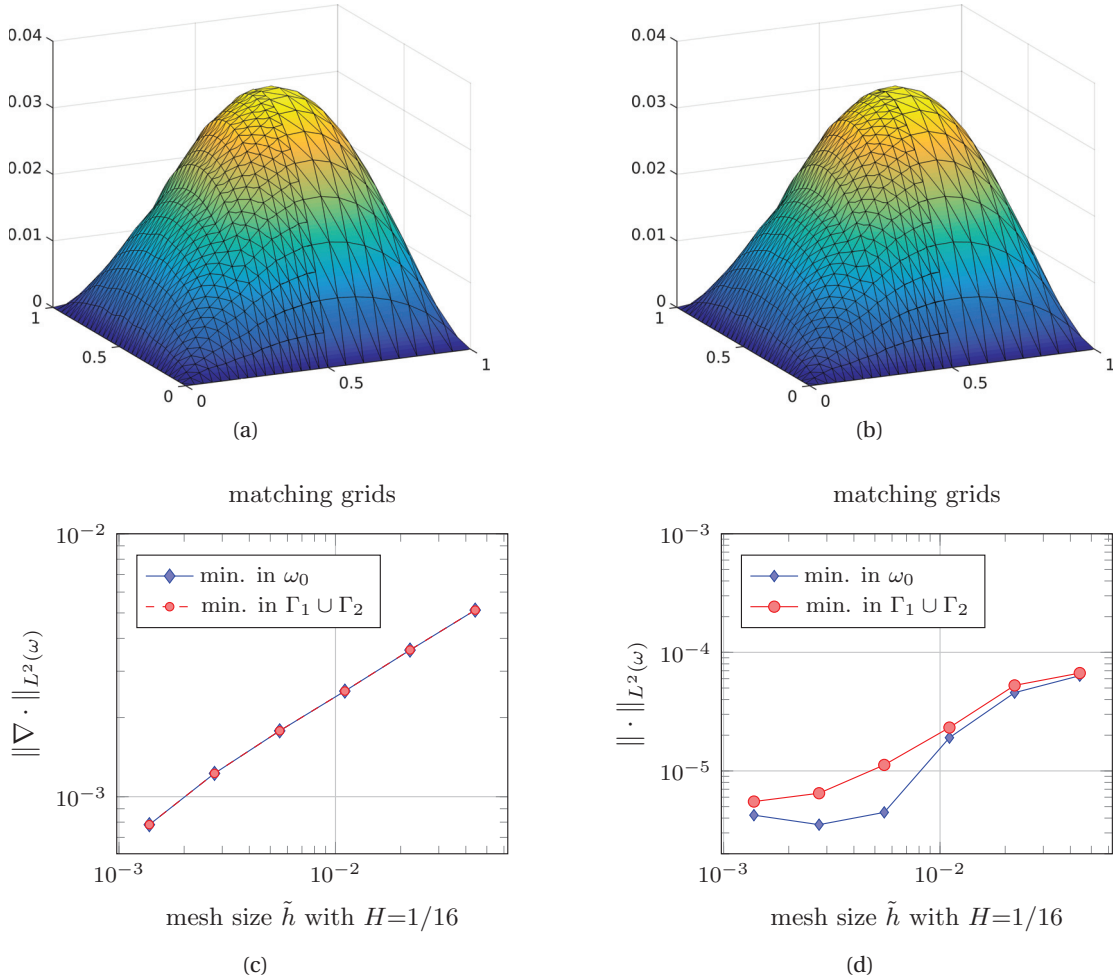


Figure 7.9 – experiment 3: numerical solutions using matching grids and the cost function of case 1 (a) and of case 2 (b), (c) the H^1 semi-norm in ω between the reference solution and the numerical solutions of the coupling using matching grids and the cost function of case 1 (diamond, blue) and the cost function of case 2 (bullet, red), (d) L^2 error between the numerical and reference solutions in ω , using matching grids and the cost function of case 1 (diamond, blue) and the cost function of case 2 (bullet, red).

We next vary the size of the overlap ω_0 , and consider $\omega_1 = [0, 1/4 + mH] \times y$, for $m = 1, 4, 8$,

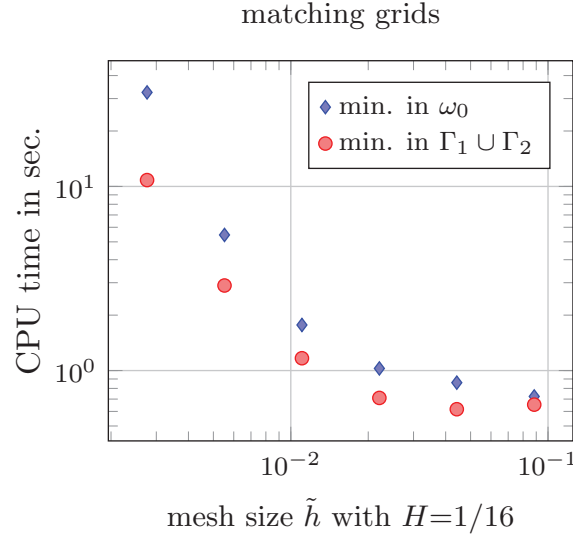


Figure 7.10 – experiment 3: CPU time using matching grids with the cost function of case 1 (diamond, blue) and the cost function of case 2 (bullet, red).

where $H = 1/32$ is the coarse mesh size, and initialize $\tilde{h} = 1/64$. We minimize over the overlap ω_0 . We observe that both couplings are influenced by the size of $\tau = \text{dist}(\Gamma_1 \cup \Gamma_2)$ and this is shown in the H^1 errors in Figure 7.11. The rates deteriorate when τ goes to zero.

7.4.3 Minimization with interface controls on non-matching grids

For the last experiment, we combine the two previous effects. The fastest coupling is obtained by performing the minimization with of the cost of case 2 with interpolation of the two meshes in the overlap, whereas the slowest coupling is obtained by the minimization with the cost function of case 1 using identical meshes in the overlap.

Experiment 4. We consider a Dirichlet elliptic boundary value in $\Omega = [0, 1]^2$,

$$\begin{aligned} -\text{div}(a^\varepsilon(x)\nabla u^\varepsilon) &= f, \text{ in } \Omega, \\ u &= 0, \text{ on } \Gamma \end{aligned}$$

with $f \equiv 1$ and a^ε is given by

$$\begin{aligned} a_2^\varepsilon(x_1, x_2) &= \frac{1}{6} \left(\frac{1.1 + \sin(2\pi(x_1/\varepsilon)(x_2/\varepsilon))}{1.1 + \sin(2\pi x_2/\varepsilon)} + \sin(4x_1^2 x_2^2) + 2 \right) I_2, \\ a_\omega^\varepsilon(x_1, x_2) &= 3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=0}^j \frac{2}{j+1} \cos(\lfloor 8(i x_2 - x_1/(i+1)) \rfloor + \lfloor 150 i x_1 \rfloor + \lfloor 150 x_2 \rfloor). \end{aligned}$$

We set $H = 1/16$ and $\varepsilon = 1/10$. We initialize $\tilde{h} = 1/32$. In Figure 7.12a, we see the H^1 error for the two settings are similar whereas the computational cost using minimization over the

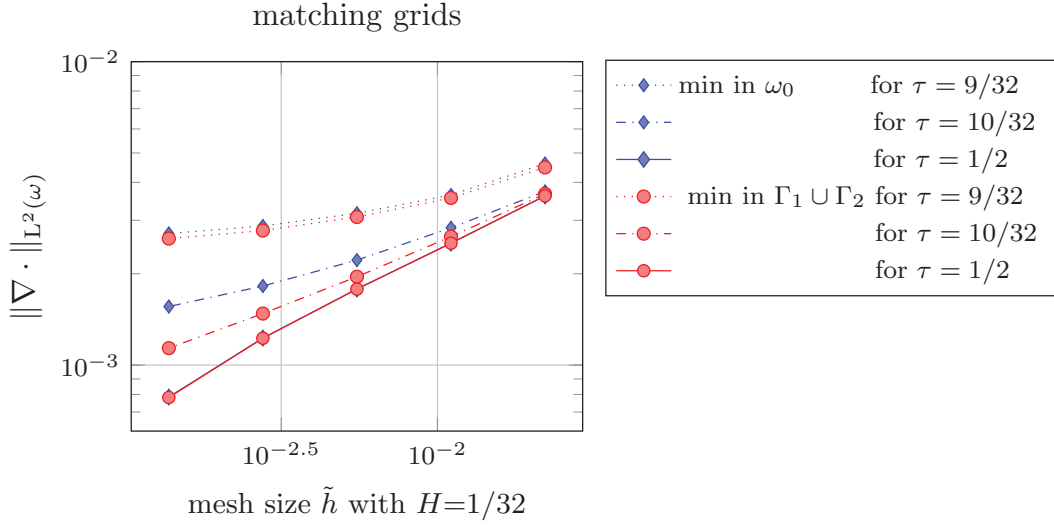


Figure 7.11 – experiment 3: errors in ω between the reference solution and the numerical solution of the coupling with matching grids and the cost function of case 1 (diamond) and the cost function of case 2 (bullet) for $\tau = 9/32$ (dots), $\tau = 10/32$ (dash-dots), and $\tau = 1/2$ (full).

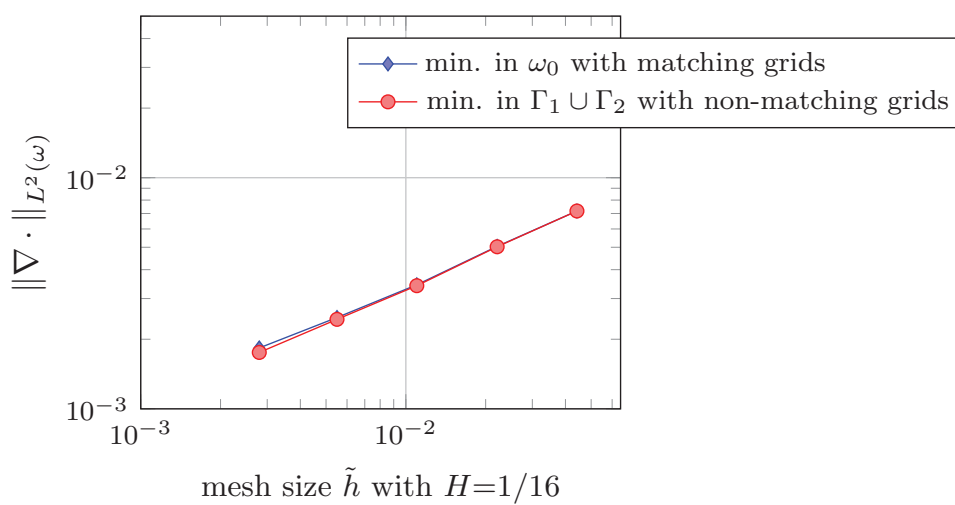
overlap and non-matching grid in ω_0 dramatically decrease (see Figure 7.12b).

7.5 Summary

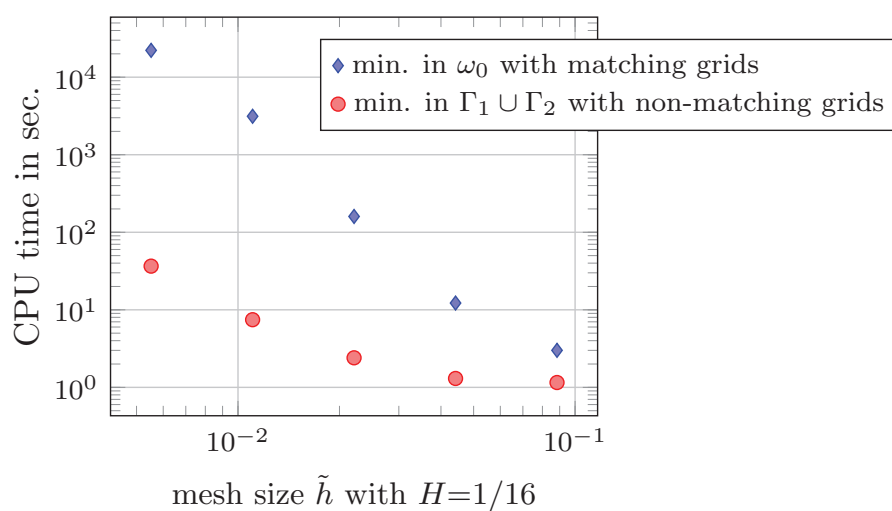
In this chapter, we give numerical improvements to the optimization based coupling method given in Chapter 6. The cost of the method is reduced by considering a minimization problem with a cost functional over the boundary of the overlapping region; i.e.,

$$J(\theta_1, \theta_2) = \frac{1}{2} \|u_1^\varepsilon(\theta_1) - u_2^0(\theta_2)\|_{L^2(\Gamma_1 \cup \Gamma_2)}^2.$$

The well-posedness and a priori error analysis are given and numerical examples are proposed to compare the coupling methods with the different cost functionals and meshing strategies. We show that the efficiency of the method is not affected by the two novelties and that an important saving is made in the computational cost.



(a)



(b)

Figure 7.12 – experiment 4: (a) errors between the numerical and reference solutions with the cost function of case 1 and matching grids (diamond) and with the cost function of case 2 with non-matching grids (bullet), (b) CPU time with the cost function of case 1 and matching grids (diamond) and with the cost function of case 2 with non-matching grids (bullet).

8 Conclusion and outlook of Part II

Conclusion

Multiscale methods have been developed during the past decades to resolve PDEs with several scales. When the scales are well separated, one could use homogenization and averaging techniques to obtain an effective solution of the multiscale problem (and a set of effective equations) without resolving the fine scales everywhere in the computational domain. An efficient numerical method is the FE-HMM, which combines the accuracy of a fine scale method with the efficiency of a coarse scale method. However, the FE-HMM relies on the existence of a characteristic length scale ε , i.e., when the small scales are well separated. Other multiscale methods, for example the LOD or the MsFEM, do not need scale separability but they have the disadvantage of being expensive, as the fine scales are resolved and used everywhere in the computational domain.

Many problems fall in between scale separation and non scale separation, and using the FE-HMM might give inaccurate results whereas other numerical methods, such as the FEM, LOD, or MsFEM, might be too expensive. The idea is to couple a fine scale solver in regions without scale separation, with a coarse scale solver in regions with scale separation.

In Chapter 5, we review two global to local methods which allow for a recovery of the fine scales in regions of interest. The two methods are the L^2 projection method and the goal-oriented method. They couple a fine scale solver with a coarse scale solver, and give good H^1 accuracy in the fine scale regions. However, such methods rely on the existence of a precomputed global solution over the whole computational domain. Obtaining such an effective solution might be computationally expensive, and the quality of the numerical solution will depend on the quality of the effective precomputed solution. These limitations motivate the design of a new numerical method for problems with and without scale separation.

In Chapter 6, we developed a new multiscale method for problems that fall in between scale and non scale separation. The method is based on overlapping domain decompositions and is formulated as a minimization problem under constraints. The constraints are a heterogeneous

and homogenized state equation, and the minimization is done with respect to a cost function. In Chapter 6, we use the L^2 norm of the difference between the heterogeneous and homogenized solutions in the overlapping regions. The discrete method couples the (DG-)FE-HMM in region with scale separation with the FEM in region without scale separation. Reconstruction can be used on the (DG-)FE-HMM solution to obtain H^1 convergence in the computational domain, between the multiscale solution and the numerical coupling solution. We prove that the method is well-posed and give a priori error estimates with explicit convergence rates in terms of the fine and coarse mesh sizes. The analysis of the method relies on Caccioppoli inequalities and on a strong version of the well known Cauchy–Schwarz inequality. Due to the use of Caccioppoli inequalities, the coupling method depends on the width of the overlapping region, and when the width is made smaller, we see that the convergence rates deteriorate.

As opposed to the global to local methods given in Chapter 5, the FE-HMM is not used in region without separation of scales, and through numerical experiments, we compare the goal-oriented method with our coupling method. In practice, we want to avoid computing the numerical homogenized solution at each mesh refinement, as it is expensive to compute, thus we fix the mesh size in $\Omega \setminus \omega_1$ — we recall that ω_1 is the region where the fine scale solver is used — and compute the numerical homogenized solution on a coarse initial mesh. This solution (or its interpolation) is then used as boundary conditions in the goal-oriented method. We refine only in ω_1 and compare the two methods. In that situation, the goal-oriented method reaches a threshold value depending on the accuracy of the precomputed solution, whereas the optimization based method produces better convergence rates; it will eventually reach a threshold value depending on the coarse mesh size.

In Chapter 7, we give numerical improvements to our coupling method. They consists in reducing the number of DOF of the method and through various experiments, we show that the efficiency of the method does not deteriorate when using the two following novelties.

First improvement. We consider the optimization based coupling with a minimization over the boundary of the overlapping region instead of a minimization over the whole overlapping region.

This reduces the number of degrees of freedom of the minimization problem. In the numerical experiments, we compare the coupling method with the minimization over the boundary of the overlapping region with the coupling method where the minimization is done in the overlapping region. We see that the convergence rates are similar, but that the minimization over the boundary is computationally faster than minimizing in the overlapping region.

Second improvement. In order to take advantage of the macro to micro coupling in the FE-HMM, the number of DOF in the region with separation of scale should not be large. On the one hand, the partition used in the FE-HMM should be coarse and have a small number of DOF. On the other hand, the FEM requires a fine partition so that the small scales are fully resolved. An interpolation between the two meshes can be done in the overlapping regions.

One can then contain the DOF of the FE-HMM to a low value, whereas the DOF of the fine scale solver can be large in order to resolve the fine scales. This results in a large saving in computational time. In the numerical experiments, we compare the convergence rates and computational times of the coupling with and without matching grids in the overlapping region. We see that the rates are similar, but that a significant saving can be observed with non-matching grids.

Further, we combine the matching grids with the minimization in the overlapping region with the non-matching grids with minimization over the boundary of the overlapping region. Then again, the convergence rates are similar whereas the computational time dramatically decreases.

Outlook

Several interesting extensions could be pursued following the work developed in this thesis.

At first, the computational cost of the coupling method can be further decreased. For example one can use a reduced basis approach in the region where scale separation is present.

At second, the method can be adapted to more realistic multiscale situations. The method presented in this thesis is derived for a domain decomposition with two overlapping subdomains. It can however be adapted to a family of overlapping domains in order to approximate for example multiple defects present in a medium. One can also consider time dependent problems, and in such situations, an interesting question arises: should the coupling algorithm be computed at each time iteration? If the answer is positive, then the computational cost of the method is very large, and one should use strategies, e.g. using reduced basis, to reduce the cost of the coupling algorithm.

Further, in many situations, the regions where a fine scale solver should be used are unknown. This is for example the case when a crack propagates through a material. The regions of interests can be found before performing the coupling method. By computing the H^1 error between a relatively coarse FEM approximation and the reference fine scale solution, we can select a set of coarse elements where the error is above a given tolerance. The regions of interests and their sizes and shapes can be further determined by conducting the same procedure on a finer partition in the selected coarse elements.

At last, the method can be generalized to multiscale problems of different nature. Indeed, in this thesis, we focused on the coupling of two elliptic problems, however the optimization based method is not restricted to this class of problems. Considering two different state equations offers the possibility to model realistic situations.

Bibliography

- [1] A. ABDULLE, *On a priori error analysis of fully discrete heterogeneous multiscale FEM*, Multiscale Model. Simul., 4 (2005), pp. 447–459.
- [2] ———, *Analysis of a heterogeneous multiscale FEM for problems in elasticity*, Math. Models Methods Appl. Sci., 16 (2006), pp. 615–635.
- [3] ———, *The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs*, in Multiple scales problems in biomathematics, mechanics, physics and numerics, vol. 31 of GAKUTO Internat. Ser. Math. Sci. Appl., Gakkōtoshō, Tokyo, 2009, pp. 133–181.
- [4] ———, *A priori and a posteriori error analysis for numerical homogenization: a unified framework*, Ser. Contemp. Appl. Math. CAM, 16 (2011), pp. 280–305.
- [5] ———, *Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales*, Math. Comp., 81 (2012), pp. 687–713.
- [6] A. ABDULLE AND Y. BAI, *Fully discrete analysis of the heterogeneous multiscale method for elliptic problems with multiple scales*, IMA J. Numer. Anal., 35 (2015), pp. 133–160.
- [7] A. ABDULLE AND O. BUDÁČ, *A reduced basis finite element heterogeneous multiscale method for stokes flow in porous media*, Computer Methods in Applied Mechanics and Engineering, 307 (2016), pp. 1–31.
- [8] A. ABDULLE, W. E, B. ENGQUIST, AND E. VANDEN-EIJNDEN, *The heterogeneous multiscale method*, Acta Numer., 21 (2012), pp. 1–87.
- [9] A. ABDULLE AND M. J. GROTE, *Finite element heterogeneous multiscale method for the wave equation*, Multiscale Model. Simul., 9 (2011), pp. 766–792.
- [10] A. ABDULLE, M. J. GROTE, AND C. STOHRER, *Finite element heterogeneous multiscale method for the wave equation: long-time effects*, Multiscale Model. Simul., 12 (2014), pp. 1230–1257.
- [11] A. ABDULLE AND M. E. HUBER, *Discontinuous Galerkin finite element heterogeneous multiscale method for advection–diffusion problems with multiple scales*, Numer. Math., 126 (2014), pp. 589–633.

Bibliography

- [12] ———, *Finite element heterogeneous multiscale method for nonlinear monotone parabolic homogenization problems*, M2AN Math. Model. Numer. Anal., 50 (2016), pp. 1659–1697.
- [13] A. ABDULLE AND O. JECKER, *An optimization-based, heterogeneous to homogeneous coupling method*, Commun. Math. Sci., 13 (2015), pp. 1639–1648.
- [14] ———, *Numerical experiments for multiscale problems in linear elasticity*, in Numerical Mathematics and Advanced Applications ENUMATH 2015, Springer, 2016, pp. 123–131.
- [15] ———, *On heterogeneous coupling of multiscale methods for problems with and without scale separation*. submitted, 2016.
- [16] A. ABDULLE, O. JECKER, AND A. V. SHAPEEV, *An optimization based coupling method for multiscale problems*, Multiscale Model. Simul., 14 (2016), pp. 1377–1416.
- [17] A. ABDULLE, P. LIN, AND A. V. SHAPEEV, *A priori and a posteriori $W^{1,\infty}$ error analysis of a QC method for complex lattices*, SIAM J. Numer. Anal., 51 (2013), pp. 2357–2379.
- [18] A. ABDULLE AND A. NONNENMACHER, *Adaptive finite element heterogeneous multiscale method for homogenization problems*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 2710–2726.
- [19] A. ABDULLE AND T. POUCHON, *A priori error analysis of the finite element heterogeneous multiscale method for the wave equation over long time*, SIAM J. Numer. Anal., 54 (2016), pp. 1507–1534.
- [20] A. ABDULLE AND C. SCHWAB, *Heterogeneous multiscale FEM for diffusion problems on rough surfaces*, Multiscale Model. Simul., 3 (2005), pp. 195–220.
- [21] G. ALLAIRE, *Homogenization and two-scale convergence*, SIAM J. Math. Anal., 23 (1992), pp. 1482–1518.
- [22] G. ALLAIRE, *Shape Optimization by the Homogenization Method*, Applied Mathematical Sciences, 146, Springer-Verlag New York, 2002.
- [23] G. ALLAIRE AND M. BRIANE, *Multiscale convergence and reiterated homogenisation*, Proc. Roy. Soc. Edinburgh Sect. A, 126 (1996), pp. 297–342.
- [24] G. ALLAIRE AND R. BRIZZI, *A multiscale finite element method for numerical homogenization*, Multiscale Model. Simul., 4 (2005), pp. 790–812 (electronic).
- [25] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.
- [26] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

-
- [27] I. BABUŠKA AND R. LIPTON, *L2-global to local projection an approach to multiscale analysis*, Math. Models and Meth. in Appl. Science, 21 (2011), pp. 2211–2226.
- [28] ———, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Model. Simul., 9 (2011), pp. 373–406.
- [29] F. BASSI AND S. REBAY, *A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier–stokes equations*, Journal of computational physics, 131 (1997), pp. 267–279.
- [30] A. BENSOUSSAN, J.-L. LIONS, AND G. PAPANICOLAOU, *Asymptotic analysis for periodic structures*, North-Holland Publishing Co., Amsterdam, 1978.
- [31] X. BLANC, C. LE BRIS, AND P.-L. LIONS, *A possible homogenization approach for the numerical simulation of periodic microstructures with defects*, Milan J. Math., 80 (2012), pp. 351–367.
- [32] A. BOURGEOIS, M. BOURGET, P. LAILLY, M. POULET, P. RICARTE, AND R. VERSTEEG, *Marmousi, model and data*, The Marmousi experience: Eur. Assn. Expl. Geophys, (1991), pp. 5–16.
- [33] S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.
- [34] F. BREZZI, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, ESAIM Math. Model. Numer. Anal., 8 (1974), pp. 129–151.
- [35] F. BREZZI, G. MANZINI, D. MARINI, P. PIETRA, AND A. RUSSO, *Discontinuous finite elements for diffusion problems*, Atti Convegno in onore di F. Brioschi (Milano 1997), Istituto Lombardo, Accademia di Scienze e Lettere, (1999), pp. 197–217.
- [36] J. Q. BROUGHTON, F. F. ABRAHAM, N. BERNSTEIN, AND E. KAXIRAS, *Concurrent coupling of length scales: methodology and application*, Physical review B, 60 (1999), p. 2391.
- [37] J. M. CARCIONE, S. PICOTTI, F. CAVALLINI, AND J. E. SANTOS, *Numerical test of the schoenberg-muir theory*, Geophysics, 77 (2012), pp. C27–C35.
- [38] Z. CHEN AND H. CHEN, *Pointwise error estimates of discontinuous galerkin methods with penalty for second-order elliptic problems*, SIAM J. Numer. Anal., 42 (2005), pp. 1146–1166(electronic).
- [39] E. T. CHUNG, Y. EFENDIEV, AND S. FU, *Generalized multiscale finite element method for elasticity equations*, GEM-International Journal on Geomathematics, 5 (2014), pp. 225–254.

Bibliography

- [40] E. T. CHUNG, Y. EFENDIEV, AND W. T. LEUNG, *Generalized multiscale finite element methods for wave propagation in heterogeneous media*, Multiscale Modeling & Simulation, 12 (2014), pp. 1691–1721.
- [41] P. G. CIARLET, *The finite element method for elliptic problems*, vol. 4 of Studies in Mathematics and its Applications, North-Holland, 1978.
- [42] P. G. CIARLET AND P. A. RAVIART, *The combined effect of curved boundaries and numerical integration in isoparametric finite element methods*, in The mathematical foundations of the finite element method with applications to partial differential equations, 1972, pp. 409–474.
- [43] D. CIORANESCU, A. DAMLAMIAN, AND G. GRISO, *Periodic unfolding and homogenization*, C. R. Math. Acad. Sci. Paris, 335 (2002), pp. 99–104.
- [44] D. CIORANESCU AND P. DONATO, *An introduction to homogenization*, vol. 17 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, New York, 1999.
- [45] E. DE GIORGI AND S. SPAGNOLO, *Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine*, Boll. Un. Mat. Ital., 4 (1973), pp. 391–411.
- [46] M. D'ELIA, M. PEREGO, P. BOCHEV, AND D. LITTLEWOOD, *A coupling strategy for nonlocal and local diffusion models with mixed volume constraints and boundary conditions*, Computer and Mathematics with Applications, 71 (2016), pp. 2218–2230.
- [47] A. DEMLOW, J. GUZMÁN, AND A. H. SCHATZ, *Local energy estimates for the finite element method on sharply varying grids*, Math. Comp., 80 (2011), pp. 1–9.
- [48] D. A. DI PIETRO AND A. ERN, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer, 2012.
- [49] M. DISCACCIATI, P. GERVASIO, AND A. QUARTERONI, *The interface control domain decomposition (icdd) method for elliptic problems*, SIAM J. Control Optim., 51 (2013), pp. 3434–3458.
- [50] M. DOROBANTU AND B. ENGQUIST, *Wavelet-based numerical homogenization*, SIAM J. Numer. Anal., 35 (1998), pp. 540–559 (electronic).
- [51] J. DOUGLAS AND T. DUPONT, *Interior penalty procedures for elliptic and parabolic galerkin methods*, in Computing methods in applied sciences, Springer, 1976, pp. 207–216.
- [52] G. DUVAUT, *Analyse fonctionnelle et mécanique des milieux continus. applications à l'étude des matériaux composites élastiques à structure périodique, homogénéisation*, Theoretical and Applied Mechanics, (1978).
- [53] W. E, *Principles of multiscale modeling*, Cambridge University Press, Cambridge, 2011.

-
- [54] W. E AND B. ENGQUIST, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.
 - [55] W. E, B. ENGQUIST, X. LI, W. REN, AND E. VANDEN-EIJNDEN, *Heterogeneous multiscale methods: a review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.
 - [56] W. E, P. MING, AND P. ZHANG, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156.
 - [57] Y. EFENDIEV, J. GALVIS, AND T. Y. HOU, *Generalized multiscale finite element methods (gmsfem)*, J. Comput. Physics, 251 (2013), p. 116–135.
 - [58] Y. EFENDIEV AND T. Y. HOU, *Multiscale finite element methods. Theory and applications*, vol. 4 of Surveys and Tutorials in the Applied Mathematical Sciences, Springer, New York, 2009.
 - [59] B. ENGQUIST, H. HOLST, AND O. RUNBORG, *Multi-scale methods for wave propagation in heterogeneous media*, Commun. Math. Sci., 9 (2011).
 - [60] ———, *Multiscale methods for wave propagation in heterogeneous media over long time*, in Numerical analysis of multiscale computations, Springer, 2012, pp. 167–186.
 - [61] Y. EPSHTEYN AND B. RIVIÈRE, *Estimation of penalty parameters for symmetric interior penalty Galerkin methods*, J. Comput. Appl. Math., 206 (2007), pp. 843–872.
 - [62] K. GAO, E. T. CHUNG, R. L. GIBSON JR, S. FU, AND Y. EFENDIEV, *A numerical homogenization method for heterogeneous, anisotropic elastic media based on multiscale theory*, Geophysics, 80 (2015), pp. D385–D401.
 - [63] M. G. D. GEERS, V. G. KOUZNETSOVA, AND W. A. M. BREKELMANS, *Multi-scale computational homogenization: Trends and challenges*, J. Comput. Appl. Math., 234 (2010), pp. 2175–2182.
 - [64] M. G. GERRITSEN AND L. J. DURLOFSKY, *Modeling fluid flow in oil reservoirs*, Annu. Rev. Fluid Mech., 37 (2005), pp. 211–238.
 - [65] P. GERVASIO, J.-L. LIONS, AND A. QUARTERONI, *Heterogeneous coupling by virtual control methods*, Numer. Math., 90 (2000), pp. 241–264.
 - [66] M. GIAQUINTA, *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, Princeton University Press, 1983.
 - [67] M. GIAQUINTA AND L. MARTINAZZI, *An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs*, Scuola Normale Superiore, 2012.
 - [68] A. GLORIA, *Reduction of the resonance error. Part 1: Approximation of homogenized coefficients*, Math. Models Methods Appl. Sci., 21 (2011), pp. 1601–1630.

Bibliography

- [69] ———, *Numerical homogenization: survey, new results, and perspectives*, 37 (2012), pp. 50–116.
- [70] A. GLORIA AND Z. HABIBI, *Reduction of the resonance error - part 2: Approximation of correctors, extrapolation, and spectral theory*. Preprint, hal-00933234, 2014.
- [71] R. GLOWINSKI, J. PÉRIAUX, AND G. TERRASSON, *On the coupling of viscous and inviscid models for compressible fluid flows via domain decomposition*, Domain decomposition methods for partial differential equations, (1990), pp. 64–97.
- [72] J. A. GOFF AND T. H. JORDAN, *Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics*, Journal of Geophysical Research: Solid Earth, 93 (1988), pp. 13589–13608.
- [73] P. HENNING AND A. MÅLQVIST, *Localized orthogonal decomposition techniques for boundary value problems*, SIAM J. Sci. Comput., 36 (2014), pp. A1609–A1634.
- [74] P. HENNING AND D. PETERSEIM, *Oversampling for the Multiscale Finite Element Method*, SIAM Multiscale Model. Simul., 11 (2013), pp. 1149–1175.
- [75] J. S. HESTHAVEN, S. ZHANG, AND X. ZHU, *High-order multiscale finite element method for elliptic problems*, Multiscale Model. Simul., 12 (2014), pp. 650–666.
- [76] T. R. HILL AND W. H. REED, *Triangular mesh methods for the neutron transport equation*, Los Alamos Report LA-UR-73-479, (1973).
- [77] V. H. HOANG AND C. SCHWAB, *High-dimensional finite elements for elliptic problems with multiple scales*, Multiscale Model. Simul., 3 (2005), pp. 168–194.
- [78] T. Y. HOU AND X.-H. WU, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
- [79] T. Y. HOU, X.-H. WU, AND Z. CAI, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.
- [80] V. V. JIKOV, S. M. KOZLOV, AND O. A. OLEINIK, *Homogenization of differential operators and integral functionals*, Springer-Verlag, Berlin, Heidelberg, 1994.
- [81] L. KLIMEŠ, *Correlation functions of random media*, Pure and applied geophysics, 159 (2002), pp. 1811–1831.
- [82] P. LESANT AND P.-A. RAVIART, *On a finite element method for solving the neutron transport equation*, Mathematical aspects of finite elements in partial differential equations, (1974), pp. 89–123.
- [83] P. LIN, *Theoretical and numerical analysis for the quasi-continuum approximation of a material particle model*, Mathematics of computation, 72 (2003), pp. 657–675.

-
- [84] ———, *Convergence analysis of a quasi-continuum approximation for a two-dimensional material without defects*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 313–332.
- [85] J.-L. LIONS, *Optimal Control of systems governed by partial differential equations*, Springer-Verlag, New York, 1971.
- [86] J.-L. LIONS AND E. MAGENES, *Problèmes aux limites non homogènes et applications*, vol. 1 of Travaux et recherches mathématiques, Dunod, Paris, 1968.
- [87] J.-L. LIONS AND O. PIRONNEAU, *Sur le contrôle parallèle des systèmes distribués*, C.R. Acad. Sci., (1998), pp. 993–998.
- [88] A. MÅLQVIST AND D. PETERSEIM, *Localization of elliptic multiscale problems*, Math. Comp., 83 (2014), pp. 2583–2603.
- [89] S. MOSKOW AND M. VOGELIUS, *First-order corrections to the homogenised eigenvalues of a periodic composite medium. a convergence proof*, Proc. Roy. Soc. Edinburgh, 127A (1997), pp. 1263–1299.
- [90] F. MURAT AND L. TARTAR, *H-convergence*, in Topics in the mathematical modelling of composite materials, vol. 31 of Progr. Nonlinear Differential Equations Appl., Birkhäuser Boston, Boston, MA, 1997, pp. 21–43.
- [91] G. NGUETSENG, *A general convergence result for a functional related to the theory of homogenization*, SIAM J. Math. Anal., 20 (1989), pp. 608–623.
- [92] J. A. NITSCHKE AND A. H. SCHATZ, *Interior estimates for ritz-galerkin methods*, Math. Comp., 28 (1974), pp. 937–958.
- [93] A. OBERMANN, T. PLANÈS, E. LAROSE, C. SENS-SCHÖNFELDER, AND M. CAMPILLO, *Depth sensitivity of seismic coda waves to velocity perturbations in an elastic heterogeneous medium*, Geophysical Journal International, 194 (2013), pp. 372–382.
- [94] J. T. ODEN, S. PRUDHOMME, A. ROMKES, AND P. T. BAUMAN, *Multiscale modeling of physical phenomena: adaptive control of models*, SIAM J. Sci. Comput., 28 (2006), pp. 2359–2389.
- [95] J. T. ODEN AND K. S. VEMAGANTI, *Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. I. Error estimates and adaptive algorithms*, J. Comput. Phys., 164 (2000), pp. 22–47.
- [96] O. A. OLEINIK, A. SHAMAEV, AND G. YOSIFIAN, *Mathematical Problems in Elasticity and homogenization*, North-Holland, Amsterdam, 1992.
- [97] D. OLSON, P. BOCHEV, M. LUSKIN, AND A. V. SHAPEEV, *Development of an optimization-based atomistic-to-continuum coupling method*, in International Conference on Large-Scale Scientific Computing, Springer, 2013, pp. 33–44.

Bibliography

- [98] D. OLSON, P. B. BOCHEV, M. LUSKIN, AND A. V. SHAPEEV, *An optimization-based atomistic-to-continuum coupling method*, SIAM J. Numer. Anal., 52 (2014), pp. 2183–2204.
- [99] N. PANASENKO AND N. BAKHVALOV, *Homogenization: Averaging Processes in Periodic Media: Mathematical Problems in the Mechanics of Composite Materials*, Kluwer Academic, 1989.
- [100] G. C. PAPANICOLAOU AND S. R. S. VARADHAN, *Boundary value problems with rapidly oscillating random coefficients*, Random fields, 1 (1979), pp. 835–873.
- [101] A. QUARTERONI, *Numerical Models for Differential Problems*, vol. 2 of Modeling, Simulation & Applications, Springer, 2009.
- [102] E. SÁNCHEZ-PALENCIA, *Non-homogeneous media and vibration theory*, vol. 127 of Lecture Notes in Phys., Springer, 1980.
- [103] A. H. SCHATZ AND L. B. WAHLBIN, *Interior maximum norm estimates for finite element methods*, Math. Comp., 31 (1977), pp. 414–442.
- [104] M. SCHOENBERG AND F. MUIR, *A calculus for finely layered anisotropic media*, Geophysics, 54 (1989), pp. 581–589.
- [105] A. V. SHAPEEV, *Consistent energy-based atomistic/continuum coupling for two-body potentials in one and two dimensions*, Multiscale Modeling & Simulation, 9 (2011), pp. 905–932.
- [106] S. SPAGNOLO, *Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 22 (1968), pp. 571–597.
- [107] E. B. TADMOR, M. ORTIZ, AND R. PHILLIPS, *Quasicontinuum analysis of defects in solids*, Philosophical magazine A, 73 (1996), pp. 1529–1563.
- [108] L. TARTAR, *Cours Peccot*. Collège de France, 1977.
- [109] K. S. VEMAGANTI AND J. T. ODEN, *Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. II. A computational environment for adaptive modeling of heterogeneous elastic solids*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 6089–6124.
- [110] G. J. WAGNER AND W. K. LIU, *Coupling of atomistic and continuum simulations using a bridging scale decomposition*, Journal of Computational Physics, 190 (2003), pp. 249–274.

Orane Jecker

Date of Birth 1st August 1988
Nationality Swiss

Education

- 2012 – 2016 **Ph.D. in Mathematics**, *EPFL*, Lausanne, Switzerland.
Chair of Computational Mathematics and Numerical Analysis (ANMC)
Thesis advisor: Prof. A. Abdulle
- 2010 – 2012 **M.Sc. in Mathematics**, *ETHZ*, Zürich, Switzerland.
M.Sc. thesis advisor: Prof. S. Mishra
- 2007 – 2010 **B.Sc. in Mathematics**, *EPFL*, Lausanne, Switzerland.

Publications

- [1] A. ABDULLE AND O. JECKER, *An optimization-based, heterogeneous to homogeneous coupling method*, Commun. Math. Sci., 13 (2015), pp. 1639–1648.
- [2] ———, *Numerical experiments for multiscale problems in linear elasticity*, in Numerical Mathematics and Advanced Applications ENUMATH 2015, Springer, 2016, pp. 123–131.
- [3] ———, *On heterogeneous coupling of multiscale methods for problems with and without scale separation*. submitted, 2016.
- [4] A. ABDULLE, O. JECKER, AND A. V. SHAPEEV, *An optimization based coupling method for multiscale problems*, Multiscale Model. Simul., 14 (2016), pp. 1377–1416.

Presentations

- 2015 **ENUMATH 2015**, *METU*, Ankara, Turkey.
An optimization based heterogeneous-to-homogeneous coupling method.
Contributed talk.
- 2015 **Swiss Numerics Colloquium**, *UNIGE*, Genève, Switzerland.
An optimization-based coupling method for multiscale problems.
Poster.
- 2015 **MATHICSE retreat**, Leysin, Switzerland.
An optimization-based coupling method for multiscale problems.
Contributed talk.