

# Adaptive Finite Element Methods for Multiscale Partial Differential Equations

THÈSE N° 5097 (2011)

PRÉSENTÉE LE 9 SEPTEMBRE 2011

À LA FACULTÉ 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

**Achim NONNENMACHER**

acceptée sur proposition du jury:

Prof. T. Mountford, président du jury

Prof. A. Abdulle, directeur de thèse

Dr S. Prudhomme, rapporteur

Prof. A. Quarteroni, rapporteur

Prof. E. Süli, rapporteur



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2011



# Abstract

Engineers rely on *efficient* simulations that provide them with *reliable* data in order to make proper engineering *design decisions*. The purpose of this thesis is to design *adaptive* numerical methods for *multiscale problems* in this spirit.

We consider elliptic homogenization problems discretized by the finite element heterogeneous multiscale method (FE-HMM). Unlike standard (single-scale) finite element methods, our *multiscale* discretization scheme relies on coupled macro and micro finite elements. The framework of the HMM allows to design an algorithm that follows the classical finite element structure on the macro level. The fine scales of the multiscale problems are taken into account by replacing the element-wise numerical integration over unknown macroscopic data by a numerical integration over suitably averaged micro solutions. These micro solutions are obtained from micro FE problems on sampling domains within the macro elements.

This thesis is divided into two parts.

In the first part, we discuss a short and versatile FE implementation of the multiscale algorithm. The implementation is flexible, easy to use and to modify and can handle simplicial or quadrilateral FE and various macro-micro coupling conditions for the constrained micro problems. The implementation of time-dependent problems is also discussed. Numerical examples including three dimensional problems are presented and demonstrate the efficiency and the versatility of the computational strategy.

In the second part (the main part of this thesis), we present an *a posteriori* error analysis for the FE-HMM. The *a posteriori* analysis enables us to estimate the accuracy of a numerical solution (and therefore its *reliability*) and further it allows for the design of *adaptive* numerical methods, which are the most *efficient*. The crucial component for the design of an adaptive multiscale method is the introduction of appropriate error indicators. As the error indicators depend on macroscopic data (such as the macroscopic diffusion tensor) that are not readily available, we construct error indicators that only depend on the available macro and micro FE solutions, available from previous computations.

We provide *a posteriori* estimates for the upper and lower bound in the energy norm. The corresponding macroscopic mesh refinement strategy is therefore both reliable and efficient. The microscopic mesh is refined simultaneously and – under appropriate assumptions – optimally with the macroscopic mesh. This means that the strategy reduces the macro and

## Abstract

---

micro error at the same rate. In the case of a uniformly oscillating tensor and exact micro computations, the standard *a posteriori* error estimates for the FEM applied to the homogenized problem are recovered. Numerical experiments confirm the efficiency and reliability of the adaptive multiscale method and demonstrate the optimality of the chosen macro-micro coupling. We extend the adaptive FE-HMM to higher order FE.

We further derive *a posteriori* estimates for the error in quantities of interest that are needed to make certain *design decisions*; the quantity of interest is represented by a linear functional. We derive and analyze a multiscale counterpart to the classical dual-weighted residual method and design a corresponding goal-oriented adaptive multiscale method. The efficiency of the method is shown in numerical experiments.

**Keywords:** Adaptive mesh refinement, *a posteriori* error estimate, finite element method, goal-oriented adaptivity, multiscale method, heterogeneous multiscale method, homogenization.



# Résumé

L'objet de la thèse est la construction de méthodes numériques adaptatives et efficaces pour des problèmes multi-échelles, permettant aux ingénieurs d'obtenir des données fiables pour prendre les décisions d'ingénierie appropriées.

Nous considérons des problèmes d'homogénéisation elliptiques discrétisés par la méthode d'éléments finis hétérogène multi-échelles (FE-HMM). Contrairement à la méthode d'éléments finis standard (une seule échelle), notre schéma de discrétisation multi-échelles repose sur un couplage de méthodes d'éléments finis macro et micro. Le cadre de HMM permet le développement d'un algorithme qui s'appuie sur la structure de la méthode des éléments finis classique au niveau macro. Les échelles fines des problèmes multi-échelles sont prises en compte en remplaçant l'intégration numérique sur chaque élément par une intégration numérique sur des moyennes bien choisies des micro solutions. Ces solutions sont obtenues à partir des micro problèmes d'éléments finis sur les domaines d'échantillonnage au sein des macro éléments.

Cette thèse est divisée en deux parties.

Dans la première partie, nous présentons une implémentation courte et souple de l'algorithme multi-échelles basé sur la méthode des éléments finis. L'implémentation est flexible, facile à utiliser et à modifier et peut supporter des éléments finis simpliciaux ou quadrilatéraux et diverses conditions de couplages macro-micro pour les micro problèmes avec contraintes. L'implémentation de problèmes évoluant en temps est aussi considérée. Des expériences numériques incluant des problèmes tridimensionnels sont présentées et illustrent l'efficacité et la souplesse de la stratégie de calcul.

Dans la deuxième partie (la partie principale de cette thèse), nous présentons une analyse a posteriori de l'erreur de la méthode FE-HMM. L'analyse a posteriori nous permet d'estimer la précision d'une solution numérique (et par conséquent sa fiabilité) et de plus, elle permet la conception de méthodes numériques adaptatives, qui sont les plus efficaces. La composante essentielle pour la conception d'une méthode adaptative multi-échelles est la mise en place d'estimateurs d'erreur appropriés. Comme les estimateurs d'erreur dépendent des données macroscopiques (comme le tenseur de diffusion macroscopique) qui ne sont pas facilement accessibles, nous construisons des estimateurs d'erreur qui dépendent seulement des solutions d'éléments finis macro et micro disponibles, qui proviennent de calculs précédents. Nous fournissons des estimations a posteriori pour les limites supérieures et inférieures dans

## Résumé

---

la norme de l'énergie. La stratégie macroscopique correspondante de raffinement de maillage est donc à la fois fiable et efficace. Le maillage microscopique est raffiné en même temps que le maillage macroscopique et (avec des hypothèses appropriées) de façon optimale. Cela signifie que la stratégie permet de réduire les erreurs macro et micro à la même vitesse. Dans le cas d'un tenseur uniformément oscillant et de calculs micro exacts, on retrouve l'estimateur d'erreur a posteriori standard de la méthode des éléments finis appliquée au problème homogénéisé. Des expériences numériques confirment l'efficacité et la fiabilité de la méthode adaptative multi-échelles et montrent l'optimalité du couplage macro-micro choisi.

Nous obtenons de plus des estimations a posteriori de l'erreur dans des quantités d'intérêts utiles pour prendre certaines décisions d'ingénierie. La quantité d'intérêt est représentée par une fonctionnelle linéaire. Nous obtenons et analysons l'analogie multi-échelles de la méthode classique des résidus à double pondération (DWR) et développons la méthode multi-échelles adaptative. L'efficacité de la méthode est illustrée par des méthodes numériques.

**Mots Clés:** Raffinement adaptatif de maillage, estimation *a posteriori*, méthode des éléments finis, quantité d'intérêt, méthode multi-échelles, heterogeneous multiscale method, homogénéisation.

# Acknowledgments

First of all, I would like to express my deepest gratitude to Prof. Assyr Abdulle for his guidance, availability and insight. Learning from our stimulating discussions, his precious advice, his attention to detail and remarks on style and structure, and his steady encouragement throughout every phase of my doctoral studies was a unique opportunity that has enriched my education both academically and personally.

My acknowledgement goes to the members of the jury, Prof. Endre Süli, Prof. Alfio Quarteroni and Dr. Serge Prudhomme for being part of the jury, for reading my work and for their helpful feedback. I further thank Prof. Thomas Mountford for presiding the jury.

I am grateful to Prof. Christian Lubich for his kind support and his insightful recommendations when choosing where to do my Ph.D.

My special recognition goes to my colleagues at EPFL, Yun, Alex and Gilles for their support, the many scientific and non-scientific discussions, and their friendship. I would additionally like to thank my recently-arrived colleague Martin for carefully reading my thesis draft. Moreover, I highly appreciated Anthea's help not only with the English language. My gratitude also goes to my former colleagues at the University of Edinburgh; I am obliged for the friendly atmosphere, their help and their great sense of humor.

My years in Edinburgh and Lausanne would not have been as enjoyable without the company of the many people with whom I interacted – and especially without my friends who allowed me to see many things from a different perspective. An attempt to mention each of them, while keeping the personal touch I would like to add, would lead to an acknowledgment that would be too long, yet doomed to be incomplete. Nonetheless, I want to thank each and every one of them for their interest, their support and understanding, for inspiring conversations and for answering my questions; I am grateful for getting to know so many different cultures and points of view, and last, but not least, for the fun we had together.

Finally, I want to thank my family for their help, their patience and their continuous support.

*Lausanne, July 2011*

Achim Nonnenmacher



# Contents

|  |             |
|--|-------------|
| <b>Abstract</b>  | <b>iii</b>  |
| <b>Acknowledgements</b>  | <b>vii</b>  |
| <b>Notation</b>  | <b>xiii</b> |
| <b>1 Introduction and Motivation</b>   | <b>1</b>    |
| <b>I The Finite Element Heterogeneous Multiscale Method</b>                    | <b>9</b>    |
| Introduction to Part I . . . . .   | 11          |
| <b>2 The Finite Element Heterogeneous Multiscale Method (FE-HMM)</b>           | <b>13</b>   |
| 2.1 A brief review of Homogenization . . . . .                                 | 13          |
| 2.1.1 Model problem . . . . .  | 13          |
| 2.1.2 Homogenization . . . . .   | 14          |
| 2.2 The basic theory of the Finite Element Method (FEM) . . . . .              | 19          |
| 2.3 Numerical approaches to solve elliptic multiscale PDEs . . . . .           | 21          |
| 2.4 The Finite Element Heterogeneous Multiscale Method (FE-HMM) . . . . .      | 22          |
| 2.4.1 <i>A priori</i> estimates . . . . .                                      | 27          |
| <b>3 Implementation of the FE-HMM and Simulation of Heat Transfer Problems</b> | <b>31</b>   |
| 3.1 Implementation . . . . .   | 32          |
| 3.1.1 Core structure: assembling the macro problem . . . . .                   | 33          |
| 3.1.2 Local assembly and micro solutions on sampling domains . . . . .         | 35          |
| 3.1.3 Parabolic problems . . . . .   | 41          |
| 3.1.4 Three dimensional problems . . . . .                                     | 42          |
| 3.2 Numerical examples . . . . .   | 42          |
| 3.2.1 Problems with non uniformly periodic coefficients . . . . .              | 42          |
| 3.2.2 Problems with random tensor . . . . .                                    | 49          |
| 3.2.3 Parabolic problems . . . . .   | 51          |
| 3.2.4 Three dimensional problem: steady state heat distribution in a heat sink | 52          |
| 3.2.5 Three dimensional problem: heat distribution in a microprocessor . . .   | 58          |
| <b>4 Conclusion of Part I</b>  | <b>63</b>   |

|  |            |
|--|------------|
| <b>II Adaptive Finite Element Heterogeneous Multiscale Methods</b>                     | <b>65</b>  |
| Introduction to Part II . . . . .  | 67         |
| <b>5 A short Review of Adaptive Finite Element Methods</b>                             | <b>73</b>  |
| 5.1 Global error estimators . . . . .  | 74         |
| 5.1.1 Zienkiewicz-Zhu estimator . . . . .  | 74         |
| 5.1.2 Residual-based estimators . . . . .  | 75         |
| 5.2 Goal-oriented error estimators . . . . .   | 76         |
| 5.2.1 Energy-norm based estimates . . . . .  | 80         |
| 5.2.2 Goal-oriented error estimators by Oden and Prudhomme . . . . .                   | 81         |
| 5.3 The Dual Weighted Residual (DWR) method . . . . .                                  | 81         |
| 5.3.1 Choice of $\psi^H$ . . . . .   | 84         |
| 5.3.2 Convergence rates . . . . .  | 84         |
| 5.4 Marking schemes . . . . .  | 84         |
| 5.5 Mesh refinement techniques . . . . .   | 86         |
| 5.6 Adaptivity beyond the macro and micro mesh size . . . . .                          | 87         |
| 5.7 The use higher-order polynomials in non-smooth domains . . . . .                   | 88         |
| 5.8 Summary . . . . .  | 91         |
| <b>6 Adaptive FE-HMM</b>   | <b>93</b>  |
| 6.1 Model Problem . . . . .  | 94         |
| 6.2 Main results . . . . .   | 94         |
| 6.2.1 Algorithm . . . . .  | 96         |
| 6.3 Proof of the main results . . . . .  | 99         |
| 6.3.1 Interpolation, trace and inverse estimates . . . . .                             | 99         |
| 6.3.2 Error representation formula . . . . .   | 100        |
| 6.3.3 Upper bound (Proof of Theorem 33) . . . . .                                      | 102        |
| 6.3.4 Lower bound (Proof of Theorem 34) . . . . .                                      | 103        |
| 6.3.5 Data approximation . . . . .   | 105        |
| 6.4 Numerical experiments . . . . .  | 108        |
| 6.4.1 Uniform refinement test . . . . .  | 110        |
| 6.4.2 Crack problem . . . . .  | 111        |
| 6.4.3 L-shape problem . . . . .  | 117        |
| 6.5 A comparison to the <i>a posteriori</i> FE-HMM obtained in a two-scale framework . | 122        |
| 6.6 Summary . . . . .  | 125        |
| <b>7 Higher order Adaptive FE-HMM</b>  | <b>127</b> |
| 7.1 Model problem . . . . .  | 127        |
| 7.2 Preliminaries . . . . .  | 128        |
| 7.2.1 Quadrature rules and interpolation . . . . .                                     | 128        |
| 7.2.2 Higher order multiscale fluxes and jumps . . . . .                               | 129        |
| 7.2.3 FE-HMM bilinear form . . . . .   | 130        |
| 7.3 Main results . . . . .   | 131        |

|          |  |            |
|----------|--|------------|
| 7.3.1    | Macro-micro refinement . . . . .   | 133        |
| 7.4      | Proof of the main results . . . . .  | 134        |
| 7.4.1    | Upper bound (Proof of Theorem 55) . . . . .                                | 135        |
| 7.4.2    | Lower bound (Proof of Theorem 56) . . . . .                                | 138        |
| 7.5      | Proof of Theorem 57 (Estimation of the data approximation error) . . . . . | 142        |
| 7.5.1    | Macro error . . . . .  | 144        |
| 7.5.2    | Modeling and micro error . . . . .   | 144        |
| 7.6      | Numerical experiments . . . . .  | 146        |
| 7.6.1    | Crack problem . . . . .  | 146        |
| 7.7      | Summary . . . . .  | 150        |
| <b>8</b> | <b>Goal-Oriented Adaptive FE-HMM</b>                                       | <b>151</b> |
| 8.1      | Model problem . . . . .  | 152        |
| 8.2      | Main results . . . . .   | 152        |
| 8.2.1    | Exact error representation . . . . .                                       | 153        |
| 8.3      | Algorithm . . . . .  | 156        |
| 8.3.1    | Macro-micro coupling . . . . .   | 158        |
| 8.3.2    | Approximation of the linear functional . . . . .                           | 159        |
| 8.4      | Proof of Theorem 68 (Exact DWR FE-HMM error representation) . . . . .      | 159        |
| 8.5      | Proof of Theorem 71 (Estimation of the data approximation error) . . . . . | 162        |
| 8.5.1    | Choice of $\psi^H$ . . . . .   | 165        |
| 8.6      | Numerical Experiments . . . . .  | 166        |
| 8.6.1    | Point-wise directional derivative . . . . .                                | 166        |
| 8.7      | Summary . . . . .  | 181        |
| <b>9</b> | <b>Conclusion and Outlook of Part II</b>                                   | <b>183</b> |
| 9.1      | Outlook . . . . .  | 184        |
|          | <b>Bibliography</b>  | <b>187</b> |





# Notation

Throughout the thesis, we will use the following notation (we follow the notation as used in [6, 12]).

|                       |  |
|-----------------------|--|
| $\varepsilon$         | $\varepsilon$ represents a small scale in the considered multiscale problems.  |
| $C$                   | $C > 0$ denotes a generic constant, independent of $\varepsilon$ , whose value can change at any occurrence but depends only on the quantities which are indicated explicitly. |
| $ r $                 | For $r = (r_1, \dots, r_d) \in \mathbb{N}^d$ , we denote $ r  = r_1 + \dots + r_d$ , $D^r = \partial_1^{r_1} \dots \partial_d^{r_d}$ .   |
| $H^1(\Omega)$         | The Sobolev space $H^1(\Omega) = \{u \in L^2(\Omega); D^r u \in L^2(\Omega),  r  \leq 1\}$ .   |
| $\ u\ _{H^1(\Omega)}$ | The norm $\ u\ _{H^1(\Omega)} = (\sum_{ r  \leq 1} \ D^r u\ _{L^2(\Omega)}^2)^{1/2}$ corresponding to the Sobolev space $H^1(\Omega)$ .  |
| $H_0^1(\Omega)$       | $H_0^1(\Omega)$ is the closure of $C_0^\infty(\Omega)$ for the $\ \cdot\ _{H^1(\Omega)}$ norm.   |
| $Y$                   | $Y = (0, 1)^d$ is the unit cube.   |
| $W_{per}^1(Y)$        | For the unit cube $Y = (0, 1)^d$ , we will consider  |

$$W_{per}^1(Y) = \{v \in H_{per}^1(Y); \int_Y v dx = 0\}.$$

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

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

$\|a\|_F$  The Frobenius matrix norm  $\|a\|_F := \sqrt{\sum_i \sum_j |a_{ij}|^2}$ .

**(Q1)** Condition on the quadrature formula for the FE-HMM, see (2.4.4) on page 24.

**(Q2)** Condition on the quadrature formula for the FE-HMM, see (2.4.5) on page 24.



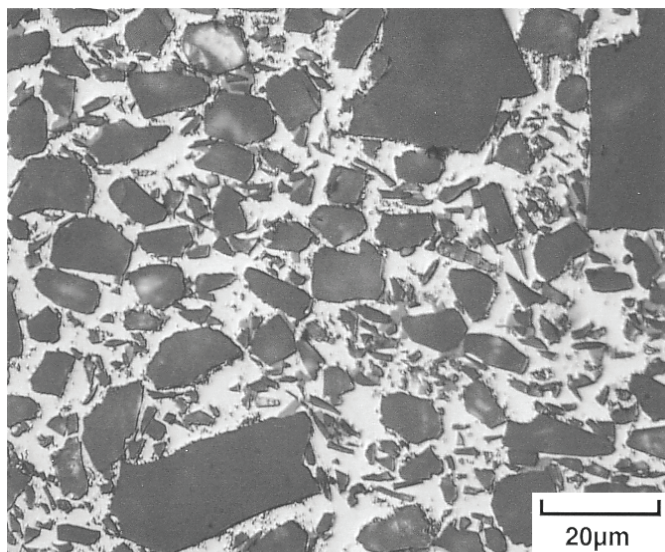
# 1 Introduction and Motivation

What is the purpose of computation? According to Babuška, a pioneer in the field of finite element methods (FEM), the purpose is “*to obtain reliable data leading to the proper engineering decisions*” [27]. It is in this spirit that we pursue in this thesis to efficiently compute the solution of so-called multiscale problems.

From a physics point of view, the (macroscopic) behavior of every object, such as deformation or thermal or electrical conductivity, could be computed solely based upon a detailed model on the atomistic level. Solid-state physicists follow this approach and use simulations on the atomistic and electronic level in order to gain insight on material properties, for example whether a material is an electrical insulator or conductor. This approach, however, will lead to a system of equations so enormous that it is impossible to even compute the behavior of a dust particle on today’s most powerful supercomputer<sup>1</sup>. Engineers, on the other hand, use macroscopic laws and continuum models such as thermodynamics or the Navier-Stokes equations, found empirically over the last centuries [49]. These empirical models allow engineers to make reasonable design decisions. With the increased capabilities of science, such as the invention of nanotechnology, there is a growing need for more realistic simulations that combine the models of the different scales. These multiscale methods predict the behavior or properties of an object or system on the macroscopic scale while also using information and models from microscopic scales. For example, classical mechanics in the macro scale could be coupled with quantum mechanics on the micro scale. Multiscale models are found in many areas, such as the simulation of the pollution of groundwater through infiltration of a fluid in a porous medium, for finding the effective properties of the increasingly used composite materials in engineering or for finding the mechanical properties of heterogeneous tissues such as bones that are important for the understanding of failure or diseases and their cure (see [6, 49]). Even in everyday life, we encounter multiscale phenomena: for example, we measure time in days, months and years to describe phenomena happening over different time scales, ultimately caused by the multiscale nature of our solar system: the rotation of

---

<sup>1</sup>In 2001 a simulation of material failure using approximately 1 billion atoms was performed on the then fastest supercomputer in the world, see [17]. While today’s fastest supercomputer is up to 500 times faster than the fastest supercomputer in 2001 (see [110]), the smallest dust particle visible by the human eye ( $5 \cdot 10^{-5}$ m) consists of roughly  $10^{17}$  atoms, and therefore has  $10^8$  times more atoms than were considered in the simulation of [17].



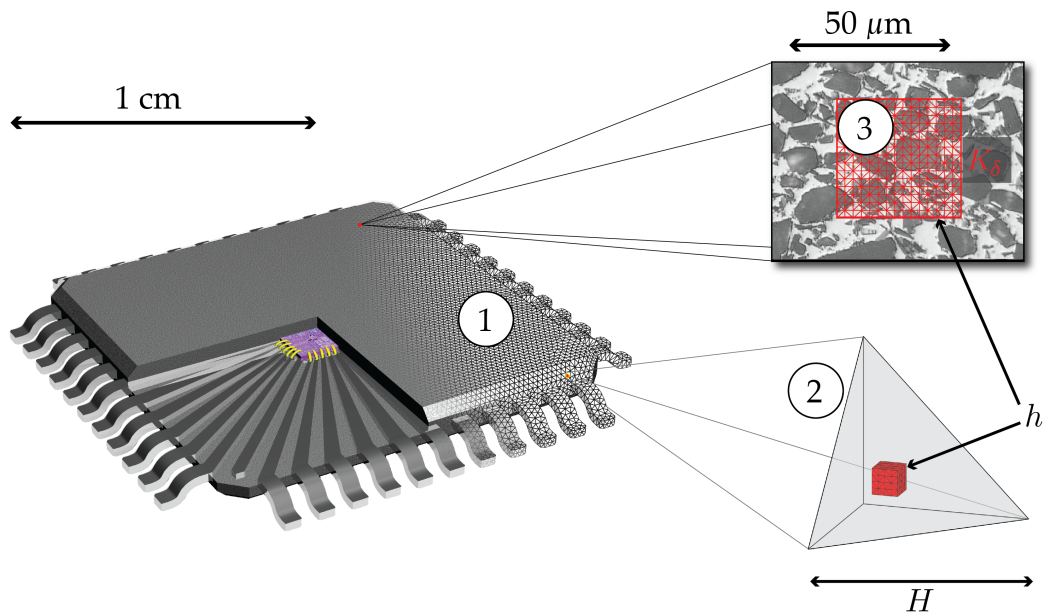
**Figure 1.0.1:** Microscopic view of a metal matrix composite (AlSiC), which is used in thermal management applications such as microprocessors or heat sinks. It consists of an aluminum (Al) matrix (dark) with silicon carbide (SiC) inclusions (white) that can be as small as  $1\ \mu\text{m}$ ; hence the inclusions are approximately 10,000 times smaller than the size of the full microprocessor or heat sink in which it is used. The composite material has properties that neither Al nor SiC have alone. The SiC content can be changed according to some design goal, such as strength or to match the coefficient of thermal expansion to surrounding materials. Picture courtesy of CPS Technologies [108], reproduced with permission.

the earth around itself, the moon around the earth and the earth around the sun [49]. The growing need for multiscale simulations has led to an urge for the development of multiscale mathematics and computational methods.

Throughout this thesis, we consider the class of problems modeled by elliptic multiscale partial differential equations (PDEs), which we solve using a multiscale finite element method (FEM). Elliptic multiscale PDEs are used for example to model the steady state heat distribution in an object made of a composite material. In Figure 1.0.1 we show a microscopic view of a metal matrix composite (AlSiC) that is used in thermal management applications, such as microprocessors or heat sinks. The composite consists of an aluminum (Al) matrix with silicon carbide (SiC) inclusions and features properties that neither Al nor SiC could have on their own. The amount of SiC inclusions can be adapted according to the engineer's design goal, such as giving the composite more strength or matching its coefficient of thermal expansion to the surrounding materials. In Figure 1.0.2 we provide a model of a typical application, where a microprocessor (about 1 cm in length and width) features components made of such a composite material. The SiC inclusions in the metal matrix are as small as  $1\ \mu\text{m}$ , hence the microprocessor will be 10,000 times larger than the inclusion<sup>2</sup>. At the same time, the

---

<sup>2</sup>For scale separation, it is not the size itself that is important, but it is important to notice that the macroscopic properties vary over a lengthscale of approximately 1 mm, whereas the microscopic properties vary over some  $\mu\text{m}$ .



**Figure 1.0.2:** Model of a multiscale setup to simulate the heat distribution in a microprocessor, which is made of a composite material. To obtain the missing data of the macro model (1) with elements of size  $H$ , a micro problem is solved on a *sampling domain* (red cube) in every macro element (2). The mesh of the sampling domain is of size  $h$  and small enough to capture the fine scale properties of the composite material (3). The sampling domain is much smaller than the macro element (see (2) ) and therefore the multiscale computation will be computationally much cheaper than the full fine scale solution. Picture of AlSiC (underneath of (3)) courtesy of CPS Technologies [108], reproduced with permission.

microstructure of the composite will influence the thermal conductivity on the macroscopic length scale. If we try to solve the heat distribution problem of the microprocessor of size 1 cm with a full fine-scale FEM solution, we will face a linear equation with over  $10^{11}$  unknowns. The requirements in memory and computational time would be enormous, even if modern techniques (such as multigrid solvers) and supercomputers were employed. However, often the engineer or scientist is merely interested in the effective, macroscopic behavior (that is influenced by the microstructure) rather than the full fine-scale behavior. This motivates the construction of numerical methods, capable of capturing such effective behavior.

In our example of the microprocessor made of composite material, we have a clear scale separation: the microscopic structure varies over a few  $\mu\text{m}$  whereas the macroscopic structure varies over some mm. As the microscopic structure features a self-similarity, we can use a multiscale framework (the so-called heterogeneous multiscale method) that exploits scale separation: inside every macro element of the macroscopic finite element mesh, we have a sampling domain that suitably represents the microstructure of the composite material within the entire macro element. At the same time, the sampling domain is much smaller than the macro element, thus reducing the total computational cost. We solve an appropriate microscopic problem, which is modeled such that it gives us the effective (homogenized)

properties inside every macro element, here the effective heat conductivity and heat flux, respectively. We then use these effective properties in the macroscopic model (see Figure 1.0.2 for an illustration). In this way, we can obtain the effective macroscopic solution at a very low computational cost.

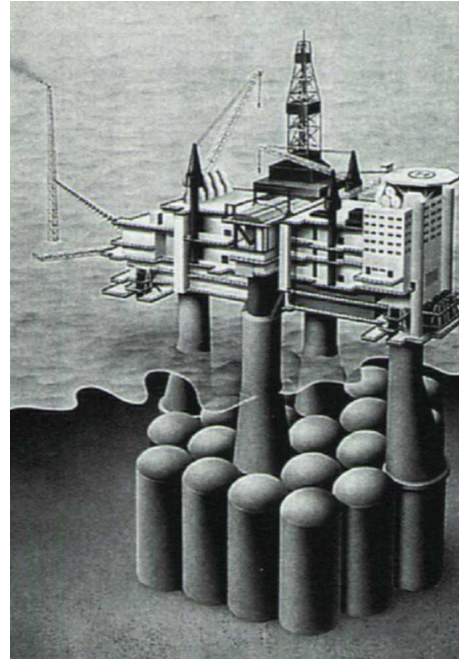
**Multiscale Methods.** In recent years a considerable amount of effort has been put into the design of multiscale methods for elliptic PDEs with multiple scales. In this thesis we will focus on homogenization problems. Babuška and Osborn [25, 24] developed the pioneering work for multiscale FEM for elliptic problems using multiscale basis functions. There has been a vast amount of literature and strategies to solve such problems, we mention – without attempting to be exhaustive – the multiscale finite element method (MsFEM) developed by Hou et al. [65] (see also the book by Efendiev and Hou [54]), the two-scale FEM proposed by Matache, Babuška and Schwab [76], the variational multiscale method by Hughes et al. [68], and the sparse FEM introduced by Hoang and Schwab [63]. Furthermore, there is the multigrid homogenization method by Neuss, Jäger and Wittum [83] and the heterogeneous multiscale method (HMM) proposed by E and Engquist [50, 51, 52]. In the engineering and structural mechanics literature concerning micro-macro methods for multiscale PDEs we mention Yu and Fish [117], Terada and Kukuchi [109], Kouznetsova et al. [71] and Miehe et al. [78]. The heat conduction in heterogeneous solids was studied by Özdemir et al. [93].

In this thesis we will use the so-called *finite element heterogeneous multiscale method* (FE-HMM). For elliptic problems, a semi-discrete *a priori* analysis for the FE-HMM was given in [53, 14] and a fully discrete analysis was obtained in [2, 9, 4]. Furthermore, discontinuous Galerkin FE-HMM was developed in [5, 7] and problems in elasticity were studied in [3]. See also [6] for a thorough overview.

**The Importance of Predictive and Reliable Numerical Methods.** Today, many closed black-box (single-scale) finite element solvers exist, which provide the user with a finite element solution and a graphically appealing visualization thereof. But are the results reliable? Is the accuracy of the finite element solution adequate to the goal of the engineer's analysis? Sometimes it is not.

The worst-case consequences of insufficient accuracy could be seen at the *Sleipner A* offshore oil and gas platform in the North Sea, which was going to operate in 82 m of water. The foundation of the platform is the concrete gravity base structure (GBS), which consists of 24 caisson cells; 4 caisson cells are extended to shafts that support the platform deck, see Figure 1.0.3. The GBS has a total base area of about 16,000 m<sup>2</sup>. Due to design constraints (the walls had to be thin enough so the platform could float, but strong enough to resist an increased hydrostatic pressure during the assembly), the safety factor for the design had to be very low. On August 23, 1991 the base structure was lowered into a Fjord as part of the deck mounting procedure of the platform, where the hydrostatic pressure reached critical values. During this controlled operation one base structure sprang a leak and sank. In an investigation following the accident it was found that a cell wall failed, which caused a

severe crack and leakage that the pumps were unable to cope with. Furthermore, it was found that the numerically computed stresses that were essential for the dimensioning of the walls had an error of 47%. This inaccuracy lead to an inappropriate design of the walls and, hence, to the wall failure. While there were other deficiencies, the main reason for the failure was found to be an insufficient accuracy of the FEM solution with respect to the engineers' design goal [27]. Two years later, a new platform was built and successfully mounted on the deck. The failure of the Sleipner base structure lead to a total loss of about \$700 million and could have been avoided if reliable estimates of the error in the shear stresses had been used [27]. See [69] and [27] for more information on the Sleipner accident.



**Figure 1.0.3:** An artist's view of the *Sleipner A* platform. Taken from [69, Figure 1]; reproduced with permission.

In the pioneering work of Babuška and Rheinboldt [26] it was shown that one indeed can quantitatively estimate the errors of a finite element solution *a posteriori*. This means that we can find an error bound of an actual solution based solely on the computed finite element solution and data. But the goal of *a posteriori error estimates* is not only to offer a criterion that indicates whether a certain prescribed accuracy is met, but also to give *local error indicators*, which can be used to drive an adaptive mesh refinement. They indicate where the error is the largest and thus, where the mesh should be refined. As many problems exhibit local variations or singularities (for example due to reentrant corners) adaptive methods are crucial, because they can equi-distribute the approximation error among the elements and therefore minimize the total computational effort. A vast amount of literature on *a posteriori* error analysis of elliptic PDEs using the FEM is available, we refer to the work of Ainsworth and Oden [19], Verfürth [113] and Babuška and Strouboulis [27] and to the references therein. Until the mid 1990s, the majority of *a posteriori* error estimators concentrated on *global* error estimates in the energy norm.

Generally, however, scientists and engineers are not interested in global errors in the energy norm, but rather in errors in a certain quantity of interest that is needed for an *engineering design decision*. In the above example of the oil platform, the quantity of interest might be the average shear stress over a certain domain or at a certain point. In other applications it might for example be the average heat flux through a boundary. Therefore, starting from the late 1990s, *goal-oriented* error estimates in *quantities of interest* were developed. The basic idea is

that many quantities of interest such as stresses, heat transfer or deformations of a component can be obtained by applying a linear functional to the (numerical) solution. The goal-oriented adaptive techniques usually involve the use of a dual solution. Among the different strategies we mention the work of Prudhomme and Oden [96] and [88] and Bangerth and Rannacher [28] and Becker and Rannacher [31] who developed the so-called dual-weighted residual (DWR) method (see also the review by Grätsch et al. [59] and the references therein). A general overview of adjoint methods in *a posteriori* analysis can be found in Giles and Süli [58].

**Adaptive Multiscale Methods.** As seen previously, *reliable data* needed for proper design decisions require *a posteriori* estimates<sup>3</sup>. However, despite the booming interest around the design and control of multiscale solvers, there have only been very few approaches to do a rigorous analysis of adaptive procedures for multiscale finite element methods. The success of multiscale methods in science and engineering applications will crucially depend on the efficiency and reliability of the adaptive strategies. We will see throughout this work that the gain in computational efficiency that is obtained when adding adaptivity to multiscale methods can be substantially higher than the gain that can be obtained when adding adaptivity to single-scale methods. The large gain in computational efficiency for multiscale methods is possible, because on the micro level, expensive micro computations can be avoided or re-used through the adaptive cycle.

When combining multiscale finite element methods with adaptivity, new challenges arise, namely the quantification of the interplay between the micro and macro solutions and further the design and analysis of an adaptive strategy that takes into account both micro and macro errors. In other words: How does the solution in the microscopic problem influence which macroscopic element we should refine? If we refine a macroscopic element, *how* do we need to refine the microscopic mesh of the corresponding sampling domain?

Among the few existing adaptive multiscale methods, we mention one approach based on the variational multiscale method by Hughes et al. [67, 68], which decomposes the solution into a coarse-scale and a fine-scale part. The solution of the fine-scale equation is formulated in dependence of the residual of the macro solution. *A posteriori* error estimates in the energy norm (only upper bounds) were derived by Larson and Målqvist in [74] and duality-based error estimates in [73].

First *a posteriori* error estimates for the FE-HMM have been obtained by Ohlberger [92]. These estimates are based on a reformulation of the FE-HMM in a two-scale framework [84]. In this framework, the microscopic variable is added as a supplementary variable, which results in doubling the size of the limiting problem. A tensor product FEM with quadrature in the slow variable is introduced to reformulate the FE-HMM within this framework. Piecewise linear elements are used in the macro and micro FE spaces. The *a posteriori* upper and lower

---

<sup>3</sup>In science and engineering applications, the reliability of the mathematical model of the physical behavior should also be validated, but throughout the thesis we only verify the reliability of the approximate solution with respect to the exact solution of the mathematical model.



---

bounds in [92] are therefore obtained in terms of the corresponding two-scale norm over  $\Omega \times Y$  (where  $\Omega$  is the physical (macro) domain and  $Y$  is the microscopic domain) and not in a norm that is related to the physical domain. It is not straightforward to derive optimal *a priori* or *a posteriori* error estimates in the energy norm for the physical domain  $\Omega$  based on the results of [92], and no such results have been obtained so far. Furthermore, the analysis of [92] is restricted to homogenization problems where the tensor is given in an explicit two-scale form, i.e.,  $a^\varepsilon = a(x, x/\varepsilon)$ , where the fast variable  $\frac{x}{\varepsilon}$  is periodic and  $\varepsilon$  assumed to be known (see [92, Section 3]). We note that *a posteriori* upper bounds in the  $L^2(\Omega)$  norm were obtained by Henning and Ohlberger in [61] (in the context of perforated domains) by using the two-scale techniques as in [92] and subsequently employing duality arguments.

### **Main Contributions of this Thesis.**

Despite the activity and interest in computational homogenization problems, only very few open-box computer codes are available in this field. Therefore, we contribute a short finite element implementation for multiscale elliptic (and parabolic) problems to the scientific community. We construct an algorithm with a corresponding Matlab implementation that is versatile, easy to use and to modify, yet capable of solving challenging 3D problems, such as the complex heat conduction problem in microelectronics mentioned above (see Figure 1.0.2). Indeed, such a code will prove to be accessible in many scientific and engineering contexts, as it can be easily and quickly tested, as well as modified by users to encompass features specific to their application.

The main contribution of this thesis consists of the proposition of adaptive FE heterogeneous multiscale methods. We derive *a posteriori* error estimates for the upper and lower bounds for the FE-HMM in the energy norm of the physical domain. Our estimates are obtained using a completely different approach than that proposed by Ohlberger [92]. We use a technique much closer to the standard techniques for residual-based *a posteriori* error analysis [20, 113]. The crucial components of our technique are the so-called *multiscale bilinear forms, fluxes and jumps*. We show that it is possible to relate our multiscale strategy to adaptive strategies for single-scale adaptive FEM.

Our estimates are not dependent on an explicit decomposition of the tensor and our analysis also holds for general (non-periodic) tensors, although in this case it involves a data approximation error that can only be estimated explicitly if additional spatial assumptions on the fast variable are made (such as periodicity or random homogeneity). In the case of a uniformly oscillating tensor and exact micro computations, the standard *a posteriori* error estimates for the FEM applied to the homogenized problem are recovered.

Our *a posteriori* estimates build the foundation for the adaptive FE-HMM algorithm. The error estimator gives information on how to refine the macro mesh. As singularities in the micro problems could only arise in the micro scale of the conductivity tensor, we use a uniform refinement for the micro problems (we emphasize that standard *a posteriori* techniques could be used to refine the micro mesh adaptively). Our analysis reveals *how* –

## Chapter 1. Introduction and Motivation

---

under appropriate assumptions – the refinement of the mesh of the micro sampling domain has to be coupled with the refinement of the corresponding macro element in order to obtain optimal convergence. The optimal coupling rates are confirmed in numerical experiments. We further extend our adaptive, residual-based FE-HMM to support higher order FE in the macro and micro spaces. The extension to higher order FE is a crucial component for the development of our goal-oriented adaptive FE-HMM, where higher-order so-called dual-solutions are required.

Based on the framework we developed for the adaptive, residual-based FE-HMM, we derive *a posteriori* error estimates for the adaptive FE-HMM in quantities of interest, expressed by a linear functional. A multiscale counterpart of the classical dual-weighted residual method by Bangerth and Rannacher [28] is derived and analyzed. We design the corresponding adaptive *goal-oriented* FE-HMM algorithm and demonstrate its performance on numerical experiments.

### Thesis Outline

The thesis is organized as follows.

**Part I** is a review of the Finite Element Heterogeneous Multiscale Method (FE-HMM). In **Chapter 2** we review homogenization, the Finite Element Method (FEM) and FE-HMM. In **Chapter 3** we present a short and versatile implementation of the FE-HMM in Matlab. Several numerical experiments illustrate what problems the FE-HMM is capable of solving. The work of Chapter 3 was published in [11]. We end Part I with a conclusion in **Chapter 4**.

**Part II** deals with *adaptive* Finite Element Heterogeneous Multiscale Methods. In **Chapter 5** we review adaptive (single-scale) Finite Element Methods. In **Chapter 6** we present *a posteriori* estimates for the FE-HMM in the physical energy norm for piecewise linear macro elements and use these estimates to construct an adaptive FE-HMM. We end this chapter with a comparison of our work to that by Ohlberger [92]. The results of Chapter 6 were published in [12, 10]. In **Chapter 7** we generalize the work of Chapter 6 to higher order finite elements. Based on the results obtained in Chapter 7, in **Chapter 8** we derive *a posteriori* error estimates for errors in *quantities of interest* and design a multiscale counterpart to the dual-weighted residual FEM.

We conclude in **Chapter 9** with an outlook on future work.

# **The Finite Element Heterogeneous Multiscale Method**

## **Part I**



## Introduction to Part I.

In this first Part of the thesis, we review the finite element heterogeneous multiscale method (FE-HMM) [50, 51, 52, 53, 2, 5], a computational method for solving elliptic multiscale partial differential equations (PDEs). The multiscale coefficients (such as the heat conductivity) of the PDE originate from some fine-scale structure, see Chapter 1. We follow the presentation of [6] and write the elliptic multiscale PDE in an abstract form as

$$\mathcal{L}^\varepsilon(u^\varepsilon) = f^\varepsilon, \quad (1.0.1)$$

where  $\mathcal{L}^\varepsilon$  refers to a differential operator,  $u^\varepsilon$  refers to the solution of the PDE and  $f^\varepsilon$  denotes some data of the problem. Here and throughout this thesis  $\varepsilon$  stands for a microscopic scale, where we assume scale separation, i.e., we assume the variations on the macro scale to be much larger than the variations on the micro scale. We further assume that a macroscopic description exists.

The introduction of a microscopic scale in the mathematical equation (1.0.1) might for example be motivated by small heterogeneities in the thermal conductivity tensor of the underlying physical problem we want to model. We suppose that the conductivity features a scale separation between a fast and a slow scale. We further suppose that the conductivity is either periodic, locally periodic or *statistically homogeneous* (i.e., the stochastic processes are described by joint probability distributions that are invariant under a translation of the spatial origin, see [111, Chapter 2]). Often scientists or engineers are not interested in the full fine-scale behavior, but rather in the effective, macroscopic behavior, which is influenced by the micro structure. This motivates the use of mathematical tools such as *homogenization* that describe the effective properties of a multiscale problem.

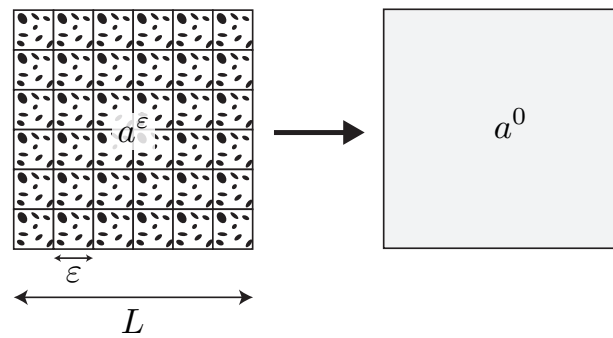
*From a mathematical point of view*, in homogenization we aim at finding the limit solution  $u^0$  for  $u^\varepsilon$  when  $\varepsilon \rightarrow 0$  and a corresponding equation

$$\mathcal{L}^0(u^0) = f^0. \quad (1.0.2)$$

Here, the solution  $u^0$  is given by the limit of  $u^\varepsilon$  for  $\varepsilon \rightarrow 0$  and is independent of  $\varepsilon$ . We call it the homogenized solution.

*From a physical point of view*, homogenization can be interpreted as follows: instead of looking at a composite material consisting of small heterogeneities with phases of different thermal conductivity, we assume an ideal composite material that has perfectly mixed phases and therefore is homogeneous. See Figure 1.0.4 for an illustration.

*From a computational point of view*, the computational cost of an exact solution of problem (1.0.1) is too expensive (see the example in Chapter 1). At the same time, the equations (1.0.2)



**Figure 1.0.4:** The left picture shows a stylized two-phase heterogeneous material with periodic micro-structure. When the macroscopic length scale  $L$  is much bigger than the microscopic length scale  $\epsilon$ , the heterogeneous material (left) with a conductivity of  $a^\epsilon$  can be treated as a homogeneous material (right) with an effective conductivity of  $a^0$ .

for the homogenized problem are in general not available in a closed form. We therefore follow the framework of the heterogeneous multiscale method (HMM) [50, 51, 52], which consists of the following algorithm:

**Step 1:** A macroscopic discretization is defined, which uses *macroscopic* parameters that are unknown *a priori*.

**Step 2:** The scales are coupled simultaneously:

**Macro-to-micro:** when macro parameters need to be evaluated, micro problems on sampling domains are computed. The micro problems are constrained by the unknown<sup>4</sup> macro state.

**Micro-to-macro:** the unknown macro parameters are then obtained based on the micro solutions by a suitable averaging.

**Step 3:** The macroscopic solution is computed.

**Outline of Part I.**

In Chapter 2 we provide a short review of the homogenization, classical FEM and introduce the FE-HMM. In Chapter 3 we give a detailed explanation of the FE-HMM algorithm and present a short implementation in Matlab. Several numerical experiments in two and three dimensions illustrate the capabilities of this short FE-HMM implementation.

---

<sup>4</sup>Note that the unknown macro state is expanded in a macro basis; while the coefficients of this expansion are unknown (solution of the macro problem), the basis functions themselves are known and used for the constraints. See Chapter 3 or [11].

## 2 The Finite Element Heterogeneous Multiscale Method (FE-HMM)

In Section 2.1 we give a quick introduction to the mathematical homogenization theory. In Section 2.2 we review classical (single-scale) FEM to set the notation for the later work. We present some finite element approaches for solving multiscale PDEs in Section 2.3 and finally review the heterogeneous multiscale finite element method (FE-HMM) in Section 2.4.

The presentation of this Chapter is taken in part from [11] and [12].

### 2.1 A brief review of Homogenization

#### 2.1.1 Model problem

We want to solve elliptic multiscale PDEs with coefficients originating from some fine scale structure. We consider a domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$  with a Lipschitz continuous boundary  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$  where Dirichlet conditions are imposed on  $\partial\Omega_D$  and Neumann conditions on  $\partial\Omega_N$ . We assume that  $\partial\Omega_D$  has positive measure. Given  $f \in L^2(\Omega)$ ,  $g_D \in H^{1/2}(\partial\Omega_D)$  and  $g_N \in L^2(\partial\Omega_N)$ <sup>1</sup>, we consider the second-order elliptic equation

$$\begin{aligned} -\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) &= f & \text{in } \Omega, \\ u^\varepsilon &= g_D & \text{on } \partial\Omega_D, \\ n \cdot (a^\varepsilon \nabla u^\varepsilon) &= g_N & \text{on } \partial\Omega_N, \end{aligned} \tag{2.1.1}$$

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

$$\exists \lambda, \Lambda > 0 \text{ such that } \lambda |\xi|^2 \leq a^\varepsilon(x) \xi \cdot \xi \leq \Lambda |\xi|^2, \forall \xi \in \mathbb{R}^d \text{ and } \forall \varepsilon. \tag{2.1.2}$$

Here and throughout this thesis,  $\varepsilon$  represents a small scale in the problem that characterizes the multiscale nature of the tensor  $a^\varepsilon(x)$ . An application of the Lax-Milgram theorem (see e.g. [39]) gives us a family of solutions that are bounded in  $H_D^1(\Omega)$  independently of  $\varepsilon$ . The

---

<sup>1</sup>We could also consider more general Neumann data, e.g.,  $g_N \in (H_{00}^{1/2}(\partial\Omega_N))'$ , see [105, Section 1.4.3] for the variational formulation. For simplicity with respect to the implementation in Chapter 3 we assume  $g_N \in L^2(\partial\Omega_N)$  here; we can thus write the corresponding duality pairing as  $\int_{\partial\Omega_N} g_N v \, ds$ .

variational form of (2.1.1) is given as follows: Find  $u^\varepsilon \in H_D^1(\Omega)$  such that

$$B^\varepsilon(u^\varepsilon, v) := \int_{\Omega} a^\varepsilon \nabla u^\varepsilon \cdot \nabla v \, dx = \int_{\Omega} f v \, dx + \int_{\partial\Omega_N} g_N v \, ds - \int_{\Omega} a^\varepsilon \nabla g_D \cdot \nabla v \, dx =: l^\varepsilon(v), \quad (2.1.3)$$

for all  $v \in H_D^1(\Omega)$ , where  $H_D^1(\Omega) := \{v \in H^1(\Omega); v = 0 \text{ on } \partial\Omega_D\}$ .

There is however a major practical issue: solving (2.1.3) using a standard FEM requires to resolve the smallest scale present in the problem, denoted here by  $\varepsilon$ . If we have a  $d$ -dimensional macroscopic geometry of size  $\mathcal{O}(1)$  this means that we need  $\mathcal{O}(\varepsilon^{-d})$  degrees of freedom to resolve the fine scale properties. For small  $\varepsilon$ , the cost associated with the FEM to solve (2.1.3) will be prohibitive. This fundamental problem associated with the complexity of discretizing multiscale problems such as (2.1.1) has been among the motivations to develop homogenization theory that is described in what follows.

### 2.1.2 Homogenization

We introduce a model problem, derive the homogenized equations, and give convergence results and error estimates of the homogenized solution.

For simplicity, we will in this Section 2.1.2 consider a model problem (2.1.1) with homogeneous Dirichlet boundary conditions. The model problem then reads

$$\begin{aligned} -\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) &= f \quad \text{in } \Omega, \\ u^\varepsilon &= 0 \quad \text{on } \partial\Omega, \end{aligned} \quad (2.1.4)$$

where  $a^\varepsilon$  is symmetric, satisfies  $a^\varepsilon(x) \in (L^\infty(\Omega))^{d \times d}$  and is uniformly elliptic and bounded. In the weak form problem (2.1.4) is given as: Find  $u^\varepsilon \in H_0^1(\Omega)$  s.t.

$$B^\varepsilon(u^\varepsilon, v) := \int_{\Omega} a^\varepsilon \nabla u^\varepsilon \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H_0^1(\Omega). \quad (2.1.5)$$

We further assume that the tensor is  $Y$ -periodic (as defined below), and of the form

$$a^\varepsilon(x) = a\left(\frac{x}{\varepsilon}\right) \quad \text{a.e. in } \Omega.$$

**Definition 1** ( $Y$ -periodicity). (See [45, Definition 2.1]). Let the reference period  $Y$  denote an interval in  $\mathbb{R}^d$  defined by  $Y = (0, \ell_1) \times (0, \ell_2) \times \dots \times (0, \ell_d)$ , where  $\ell_1, \dots, \ell_d$  are given positive numbers. Then  $a^\varepsilon(x)$  is  $Y$ -periodic iff

$$a_{ij}^\varepsilon(x + k\ell_m \mathbf{e}_m) = a_{ij}^\varepsilon(x), \quad \forall k \in \mathbb{Z} \text{ and } \forall m \in \{1, \dots, d\},$$

where  $\{\mathbf{e}_m\}_{m=1}^d$  is the canonical basis of  $\mathbb{R}^d$ .

The goal of homogenization w.r.t. problem (2.1.4) is to find the so-called *homogenized*



tensor  $a^0(x)$  and the homogenized equation

$$\begin{aligned} -\nabla \cdot (a^0 \nabla u^0) &= f \quad \text{in } \Omega, \\ u^0 &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.1.6}$$

by studying the limiting behavior of  $u^\varepsilon$  which converges (usually in a weak sense) to the so-called homogenized solution  $u^0$  as  $\varepsilon \rightarrow 0$ . We will frequently use the weak formulation of problem (2.1.6): Find  $u^0 \in H_0^1(\Omega)$  such that

$$B_0(u^0, v) := \int_{\Omega} a^0 \nabla u^0 \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H_0^1(\Omega), \tag{2.1.7}$$

where  $B_0(\cdot, \cdot)$  is the so-called homogenized bilinear form.

In what follows we show that the homogenized solution  $u^0$  satisfies a boundary value problem (which will turn out to be the homogenized equations above) and that  $u^0$  is uniquely determined. We follow the presentation of [45] and refer the reader to [32, 70, 94] for more details.

**Weak convergence.**

From the Lax-Milgram Theorem [39, 45], we know that (2.1.5) has a unique solution and we also obtain that

$$\|u^\varepsilon\|_{H_0^1(\Omega)} \leq \frac{1}{\lambda} \|f\|_{H^{-1}(\Omega)}, \tag{2.1.8}$$

which together with the following theorem ([45, Theorem 1.18]) will allow us to deduce weak convergence of  $u^\varepsilon$ .

**Theorem 2.** *Let the Banach space  $E$  be reflexive and let  $\{x_n\}$  be a bounded sequence in  $E$ . Then there exists a subsequence  $\{x_{n_k}\}$  of  $\{x_n\}$  and  $x \in E$  s.t. as  $k \rightarrow \infty$ , we have  $x_{n_k} \rightharpoonup x$  weakly in  $E$ , i.e.,*

$$\forall x' \in E' \quad \langle x', x_{n_k} \rangle_{E', E} \rightarrow \langle x', x \rangle_{E', E}.$$

*Remark 3.* Theorem 2 is a special case of the Banach–Alaoglu Theorem, where we assume that  $E$  is reflexive. We refer the reader to [116, Theorem III.3.7] and [107, Section 1.9] for details.

We further introduce weak\* convergence.

**Definition 4** (Weak\* convergence). (See [45, Definition 1.20]). Let  $F$  be a Banach space and set  $E = F'$ . A sequence  $\{x_n\}$  in  $E$  is said to converge weakly\* to  $x$  iff

$$\langle x_n, x' \rangle_{F', F} \rightarrow \langle x, x' \rangle_{F', F}, \quad \forall x' \in F.$$

From Theorem 2 and the bound (2.1.8) follows that there exists a subsequence  $\{u^{\varepsilon'}\}$  and

an element  $u^0 \in H_0^1(\Omega)$  s.t.

$$u^{\varepsilon'} \rightharpoonup u^0 \quad \text{weakly in } H_0^1(\Omega). \quad (2.1.9)$$

We furthermore know [45] that the function  $a^\varepsilon(x)$  weakly converges to the mean value of  $a(\frac{x}{\varepsilon}) = a(y)$  with  $y = \frac{x}{\varepsilon}$ , that means that we have

$$a^\varepsilon(x) \rightharpoonup \frac{1}{|Y|} \int_Y a(y) dy \quad \text{weakly* in } (L^\infty(\Omega))^{d \times d}.$$

The challenge is now to find the limit of the product  $a^\varepsilon \nabla u^\varepsilon$ , as in general the product of two weakly convergent sequences does not converge to the product of the weak limits (see [45]). The main difficulty of homogenization is to find the homogenized coefficients  $a^0$  and the homogenized equation for the limiting solution  $u^0$ . We will now use the method of asymptotic expansion in order to derive the homogenized equation.

**Formal asymptotic expansion.**

We consider a two-scale model for  $u^\varepsilon(x)$  with a macroscopic variable  $x$  and a microscopic variable  $\frac{x}{\varepsilon}$ , which means that we look for an asymptotic expansion of  $u^\varepsilon$  of the form

$$u^\varepsilon(x) = u^0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u^1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u^2\left(x, \frac{x}{\varepsilon}\right) + \dots \quad (2.1.10)$$

where the  $u^j(x, y)$  for  $j = 1, 2, 3, \dots$  satisfy that  $u^j(x, y)$  is defined for  $x \in \Omega$  and  $y \in Y$  and  $u^j(\cdot, y)$  is  $Y$ -periodic.

For  $\varepsilon \ll 1$  the microscopic variable  $y$  will change much more rapidly than  $x$  and when looking at the problem from a microscopic scale, we can assume  $x$  to be a constant parameter to the problem. This way we can exploit scale separation and treat  $x$  and  $y$  independently of each other.

Differentiating a function  $\phi^\varepsilon(x) := \phi\left(x, \frac{x}{\varepsilon}\right)$  w.r.t.  $x$  means that

$$\nabla \rightarrow \nabla_x + \frac{1}{\varepsilon} \nabla_y,$$

so that the total derivative of  $\phi\left(x, \frac{x}{\varepsilon}\right)$  can be expressed as

$$\nabla \phi^\varepsilon(x) = \nabla_x \phi\left(x, y\right) \Big|_{y=\frac{x}{\varepsilon}} + \frac{1}{\varepsilon} \nabla_y \phi\left(x, y\right) \Big|_{y=\frac{x}{\varepsilon}}.$$

Thus the differential operator

$$\mathcal{A}^\varepsilon = -\nabla \cdot (a^\varepsilon \nabla) \quad (2.1.11)$$

becomes

$$\mathcal{A}^\varepsilon = \frac{1}{\varepsilon^2} \mathcal{A}^0 + \frac{1}{\varepsilon} \mathcal{A}^1 + \mathcal{A}^2, \quad (2.1.12)$$

where

$$\mathcal{A}^0 = -\nabla_y \cdot (a(y) \nabla_y) \quad (2.1.13a)$$

$$\mathcal{A}^1 = -\nabla_y \cdot (a(y) \nabla_x) - \nabla_x \cdot (a(y) \nabla_y) \quad (2.1.13b)$$

$$\mathcal{A}^2 = -\nabla_x \cdot (a(y) \nabla_x). \quad (2.1.13c)$$

Inserting (2.1.12) into (2.1.4) we obtain

$$\left( \frac{1}{\varepsilon^2} \mathcal{A}^0 + \frac{1}{\varepsilon} \mathcal{A}^1 + \mathcal{A}^2 \right) u^\varepsilon = f. \quad (2.1.14)$$

We substitute (2.1.10) into (2.1.14) and obtain

$$\frac{1}{\varepsilon^2} \mathcal{A}^0 u^0 + \frac{1}{\varepsilon} (\mathcal{A}^0 u^1 + \mathcal{A}^1 u^0) + (\mathcal{A}^0 u^2 + \mathcal{A}^1 u^1 + \mathcal{A}^2 u^0) + \mathcal{O}(\varepsilon) = f.$$

Equating coefficients of equal powers of  $\varepsilon$  and ignoring the  $\mathcal{O}(\varepsilon)$ -term leads to the following set of equations

$$\mathcal{O}\left(\frac{1}{\varepsilon^2}\right): \quad \mathcal{A}^0 u^0 = 0 \quad (2.1.15a)$$

$$\mathcal{O}\left(\frac{1}{\varepsilon}\right): \quad \mathcal{A}^0 u^1 = -\mathcal{A}^1 u^0 \quad (2.1.15b)$$

$$\mathcal{O}(1): \quad \mathcal{A}^0 u^2 = -\mathcal{A}^1 u^1 - \mathcal{A}^2 u^0 + f. \quad (2.1.15c)$$

From equations (2.1.15) we can derive the following results.

$\mathcal{O}\left(\frac{1}{\varepsilon^2}\right)$ : From equation (2.1.15a) follows that  $u^0(x, \frac{x}{\varepsilon}) = u^0(x)$ , which indicates that the obtained variable  $u^0$  must be the ‘‘homogenized solution’’ that is independent of  $y$ .

$\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ : From equation (2.1.15b) we can derive the cell-problem (2.1.18) and find that  $u^1(y) = \sum_{j=1}^d \chi^j(y) \frac{\partial u^0(x)}{\partial x_j}$  as given in (2.1.21), where  $\chi^j$  is the so-called first order corrector defined in Theorem 5.

$\mathcal{O}(1)$ : Finally, the homogenized equation  $-\nabla \cdot (a^0 \nabla u^0) = f$  follows from (2.1.15c).

We refer the reader to [45, Chapter 7] for details.

### Convergence results

**Theorem 5.** See [45, Theorem 6.1]. Let  $f \in H^{-1}(\Omega)$  and  $u^\varepsilon$  be the solution of (2.1.4), where  $a^\varepsilon$  is symmetric, satisfies  $a^\varepsilon(x) \in (L^\infty(\Omega))^{d \times d}$ , is uniformly elliptic and bounded. We further assume that the tensor is  $Y$ -periodic and of the form

$$a^\varepsilon(x) = a\left(\frac{x}{\varepsilon}\right) \quad \text{a.e. in } \Omega.$$

Then

- i)  $u^\varepsilon \rightharpoonup u^0$  weakly in  $H_0^1(\Omega)$ ,
- ii)  $a^\varepsilon \nabla u^\varepsilon \rightharpoonup a^0 \nabla u^0$  weakly in  $(L^2(\Omega))^d$ ,

where  $u^0 \in H_0^1(\Omega)$  is the unique solution of (2.1.6)

$$\begin{aligned} -\nabla \cdot (a^0 \nabla u^0) &= f \quad \text{in } \Omega, \\ u^0 &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{2.1.16}$$

The homogenized tensor  $a^0 = (a^0)_{ij, 1 \leq i, j \leq d}$  is constant, elliptic and given by

$$a_{ij}^0 = \frac{1}{|Y|} \int_Y \left( a_{ij}(y) + \sum_{k=1}^d a_{ik}(y) \frac{\partial \chi^j}{\partial y_k}(y) \right) dy, \tag{2.1.17}$$

with a so-called first order corrector  $\chi^j \in W_{per}^1(Y)$ . The first order corrector  $\chi^j(y)$ ,  $j = 1, \dots, d$ , is defined to be the unique solution of the so-called cell problems

$$\int_Y \nabla \chi^j(y) \cdot a(y) \nabla v(y) dy = - \int_Y (a(y) \mathbf{e}_j)^T \nabla v(y) dy, \quad \forall v(y) \in W_{per}^1(Y), \tag{2.1.18}$$

where  $\{\mathbf{e}_j\}_{j=1}^d$  is the canonical basis in  $\mathbb{R}^d$ .

*Remark 6.* Theorem 5 holds also for a two-scale tensor with a slow variable  $x$  and a fast variable  $y = \frac{x}{\varepsilon}$ , i.e. for  $a^\varepsilon(x) = a(x, y) \in C(\Omega; L_{per}^\infty(Y))^{d \times d}$ , where the  $x$  in  $a(x, y)$  merely acts as a parameter, see [32].

**Error estimates.** How well can the homogenized solution  $u^0(x)$  approximate the original fine-scale solution  $u^\varepsilon(x)$ ? In the  $L^2$  norm, strong error estimates exist [70, Sect. 1.4]

$$\|u^\varepsilon - u^0\|_{L^2(\Omega)} \leq C\varepsilon. \tag{2.1.19}$$

In general, we cannot obtain strong errors in the  $H^1$  norm, because the oscillations of  $u^\varepsilon$  introduce  $\mathcal{O}(1)$  perturbations in the gradient which are not captured by  $u^0$ . If we look for a development of  $u^\varepsilon$  of the form

$$u^\varepsilon(x) = u^0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u^1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u^2\left(x, \frac{x}{\varepsilon}\right) + \dots \tag{2.1.20}$$

where  $u^j(x, y)$  are  $Y$ -periodic in the second variable  $y$ , then we can add to the homogenized solution  $u^0$  a corrector which has information of the fine scale and is given by

$$u^1\left(x, \frac{x}{\varepsilon}\right) = \sum_{j=1}^d \chi^j\left(\frac{x}{\varepsilon}\right) \frac{\partial u^0(x)}{\partial x_j}. \tag{2.1.21}$$

We have for the corrected solution that [70, Sect. 1.4]

$$\left\| u^\varepsilon - \left( u^0(x) + \varepsilon u^1\left(x, \frac{x}{\varepsilon}\right) \right) \right\|_{H^1(\Omega)} \leq C\sqrt{\varepsilon}. \quad (2.1.22)$$

The  $\sqrt{\varepsilon}$  boundary layer term (instead of an  $\varepsilon$  term that we might expect) is caused because  $u^1$  does not satisfy the boundary conditions of problem (2.1.1). We remark that for estimates (2.1.19) and (2.1.22) to hold, some regularity on  $u^0$  and  $\chi^j$  is required and we refer the reader to [70, Sect. 1.4] for details.

### Properties of the homogenized tensor.

*Remark 7.* We have the following properties for the homogenized tensor (see [45, 94]):

- The homogenized tensor again is elliptic and bounded, i.e.

$$\exists \lambda, \Lambda > 0 \text{ such that } \lambda|\xi|^2 \leq a^0(x)\xi \cdot \xi \leq \Lambda|\xi|^2, \quad \forall \xi \in \mathbb{R}^d. \quad (2.1.23)$$

- Symmetry of  $a^\varepsilon(x)$  implies symmetry of  $a^0(x)$ .
- If  $a^\varepsilon(x)$  is diagonal,  $a^0(x)$  is in general not diagonal. This means that an isotropic composite material can, as the microstructure gets smaller and smaller in the limit of  $\varepsilon \rightarrow 0$ , behave like an anisotropic material.

## 2.2 The basic theory of the Finite Element Method (FEM)

We briefly sketch the basic theory of the FEM that we will need for the FE heterogeneous multiscale method and refer the reader for more details and generalizations to [39, 44, 98]. We consider the non-oscillatory (single-scale) version of problem (2.1.1), i.e. an elliptic partial differential equation on a domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$  with a Lipschitz continuous boundary  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ , where Dirichlet conditions are imposed on  $\partial\Omega_D$  and Neumann conditions on  $\partial\Omega_N$ . We assume that  $\partial\Omega_D$  has positive measure. Given  $f \in L^2(\Omega)$ ,  $g_D \in H^{1/2}(\partial\Omega_D)$ ,  $g_N \in H^{-1/2}(\partial\Omega_N)$  we consider the second-order elliptic equation

$$\begin{aligned} -\nabla \cdot (a\nabla u) &= f && \text{in } \Omega, \\ u &= g_D && \text{on } \partial\Omega_D, \\ n \cdot (a\nabla u) &= g_N && \text{on } \partial\Omega_N, \end{aligned} \quad (2.2.1)$$

where  $a$  is symmetric, satisfies  $a(x) \in (L^\infty(\Omega))^{d \times d}$  and is uniformly elliptic and bounded. Using Lax-Milgram theorem (see e.g. [39]) leads to a family of solutions which are bounded in  $H_D^1(\Omega)$ . The variational form of (2.2.1) is to find  $u \in H_D^1(\Omega)$  such that

$$B_{\text{classic}}(u, v) := \int_{\Omega} a\nabla u \cdot \nabla v dx = \int_{\Omega} f v dx + \int_{\partial\Omega_N} g_N v dx - \int_{\Omega} a\nabla g_D \cdot \nabla v dx =: l(v), \quad (2.2.2)$$

## Chapter 2. The Finite Element Heterogeneous Multiscale Method (FE-HMM)

---

for all  $v \in H_D^1(\Omega)$ , where  $H_D^1(\Omega) := \{v \in H^1(\Omega); v = 0 \text{ on } \partial\Omega_D\}$ . We use the subscript *classic* on  $B_{\text{classic}}(\cdot, \cdot)$  to emphasize that the bilinear form  $B_{\text{classic}}(\cdot, \cdot)$  corresponds to a single-scale (classic) FEM problem. The associated norm is the *energy norm*

$$\|v\| := \sqrt{B_{\text{classic}}(v, v)}.$$

A finite element discretization of this variational problem is standard. To avoid technicality with curved boundary and regularity issues with re-entrant corners, we assume in what follows that  $\Omega$  is a convex polygonal domain in  $\mathbb{R}^d$ . Let  $\mathcal{T}_H$  be a partition of  $\Omega$  in simplicial or quadrilateral elements  $K$  of diameter  $H_K$  and denote  $H = \max_{K \in \mathcal{T}_H} H_K$ . For this partition we define a finite dimensional subspace of  $H_D^1(\Omega)$  by

$$V_D^p(\Omega, \mathcal{T}_H) = \{v^H \in H_D^1(\Omega); u^H|_K \in \mathcal{R}^p(K), \forall K \in \mathcal{T}_H\}, \quad (2.2.3)$$

with elements  $K \in \mathcal{T}_H$ , where  $\mathcal{T}_H$  is assumed to be shape regular. Here  $\mathcal{R}^p = \mathcal{P}^p$  is the space of piecewise polynomials on the element  $K$  of *total* degree  $p$ , if  $K$  is a simplicial element (triangle if  $d = 2$ , tetrahedron if  $d = 3$ ). If  $K$  is a quadrilateral element (quadrilateral if  $d = 2$ , hexahedron if  $d = 3$ ), then  $\mathcal{R}^p = \mathcal{Q}^p$  is the space of piecewise polynomials on the element  $K$  of degree  $p$  in *each variable*.

The solution of the discretized problem (2.2.2) reads: find  $u^H \in V_D^p(\Omega, \mathcal{T}_H)$  such that

$$B_{\text{classic}}(u^H, v^H) = l(v^H) \quad \forall v^H \in V_D^p(\Omega, \mathcal{T}_H). \quad (2.2.4)$$

**A priori estimates.** *A priori estimates* provide information about the asymptotic behavior of the FE solution  $u^H$ . For elliptic bilinear forms *Céa's Lemma* (see [44, 40, 39]) holds,

$$\|u - u^H\|_{H^1(\Omega)} \leq \frac{C}{\lambda} \inf_{v^H \in V_D^p(\Omega, \mathcal{T}_H)} \|u - v^H\|_{H^1(\Omega)}$$

which states that the error  $e^H := u - u^H$  between the finite element solution  $u^H \in V_D^p(\Omega, \mathcal{T}_H)$  of (2.2.4) and the exact solution  $u \in H_D^1(\Omega)$  of (2.2.2) is of the same order of magnitude as the interpolation error, i.e. the quality of the approximation essentially depends on the choice of  $V_D^p(\Omega, \mathcal{T}_H)$ .

Provided that the solution  $u$  of problem (2.2.1) has regularity  $u \in H^{p+1}(\Omega)$ , we have that there exists a constant  $C$  such that the error in the  $H^1$ -norm is given by

$$\|u - u^H\|_{H^1(\Omega)} \leq CH^p |u|_{p+1, \Omega}. \quad (2.2.5)$$

Furthermore, we can apply duality arguments (Aubin-Nitsche duality argument, see [39]) to obtain estimates for the  $L^2$ -error

$$\|u - u^H\|_{L^2(\Omega)} \leq CH^{p+1} |u|_{p+1, \Omega}. \quad (2.2.6)$$

## 2.3 Numerical approaches to solve elliptic multiscale PDEs

We want to briefly mention several numerical approaches in the literature that aim at solving multiscale problems of the form (2.1.1). For an overview of different methods, we refer the reader to [49].

**The variational multiscale method (VMM)** was proposed by Hughes and co-workers [67, 68]. For the VMM in context of adaptive methods, see [74, 73]. The idea is to choose a coarse and a fine space,  $V_c \subset V$  and  $V_f \subset V$ , respectively such that

$$V = V_c \oplus V_f$$

and to decompose the solution  $u \in V$  into a coarse-scale and a fine-scale component  $u = u_c + u_f$ . The key of the method is to decouple the resulting coupled equations for the coarse and fine scale components by using localization techniques. One such technique involves the use of localized Dirichlet problems on subdomains, where the fine-scale problem is approximated by a set of localized, decoupled problems. The choice of these subdomains is crucial and their size can be chosen adaptively, see [74, 73].

**The multiscale finite element method (MsFEM)** is based on the ideas of Babuška and Osborn [24, 25] who proposed to modify the finite element space and use special multiscale basis functions which take the microstructure of the problem into account. The method was later extended by Hou and Wu [64]. We refer the reader to the book by Efendiev and Hou [54] for a detailed description of the method.

The MsFEM follows a different approach than the VMM. The basic idea is to find a set of local multiscale basis functions in such a way that they resemble the correct local fine-scale characteristics of the multiscale solution. A global numerical formulation is then used in order to find a solution of the global multiscale problem within the space spanned by these multiscale basis functions. The choice of the boundary conditions for the local problems that are used to find the multiscale basis functions is crucial (see [54, Section 2.3]). The complexity of the MsFEM is similar to the complexity of the full fine-scale problem, but necessary computations over the fine grid can be re-used e.g. for different right-hand sides (as typically wanted in e.g. fluid-flow applications).

In their general formulation, both VMM and MsFEM aim at obtaining the full fine scale solution, therefore the total degrees of freedom required for finding the fine scale solution is of order  $\mathcal{O}(\varepsilon^{-d})$ . Other methods have been developed, which assume and exploit special features of the problem, such as scale separation and self-similarity of the conductivity tensor in the microscopic scale. These methods aim at reducing the total degrees of freedom of the multiscale problem, ideally making it independent of  $\varepsilon$ . One such method used in the engineering world uses representative volume elements (RVE), another such method is the finite element heterogeneous multiscale method (FE-HMM) that we use throughout this thesis.

**Upscaling methods based on representative volume elements (RVE).** A well-established approach for treating multiscale problems in the engineering and structural mechanics community is upscaling based on representative volume elements (RVEs)<sup>2</sup>. In an RVE, one simulates the microstructure of a material. The result of this micro simulation is then used in order to obtain effective macroscopic properties at a macroscopic point. Since we assume scale separation, the size of the RVE is chosen to be much smaller than the variations on the macro scale, but at the same time big enough to represent the characteristic micro structure sufficiently well. The approach typically follows the structure [93, 57, 71]:

1. At a macroscopic point(s), the micro-structural RVE is defined.
2. Macro-to-micro transition: appropriate boundary conditions involving *previously computed*<sup>3</sup> macroscopic quantities are formulated on the microscopic boundary value problem, in order to impose consistency with the macroscopic scale. (Step I in Figure 2.3.1).
3. The boundary value problem on the RVE is solved (e.g. with a FEM). (Step II in Figure 2.3.1).
4. Micro-to-macro transition: a consistent volume averaging scheme is used in order to extract the desired macroscopic quantities in terms of the microscopic quantities in the RVE. (Step III in Figure 2.3.1).

For a review of computational homogenization methods in the context of solid mechanics we refer to the lecture notes of Geers and Kouznetsova [57] and references therein. More details and references in the context of heat transfer can be found in [93, 56].

## 2.4 The Finite Element Heterogeneous Multiscale Method (FE-HMM)

This Section is an extended overview of the FE-HMM based on [11] and [12]. We again consider the elliptic problem (2.1.1) with oscillatory coefficients

$$\begin{aligned}
 -\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) &= f && \text{in } \Omega, \\
 u^\varepsilon &= g_D && \text{on } \partial\Omega_D, \\
 n \cdot (a^\varepsilon \nabla u^\varepsilon) &= g_N && \text{on } \partial\Omega_N,
 \end{aligned} \tag{2.4.1}$$

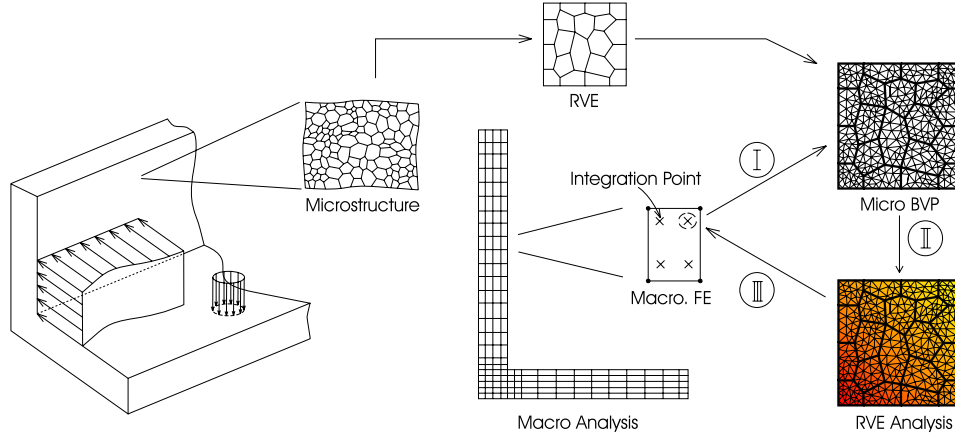
as introduced in Section 2.1.1. The corresponding homogenized elliptic problem is given by

---

<sup>2</sup>In some communities also called representative elementary volume (REV).

<sup>3</sup>In contrast, the FE-HMM uses an on-the-fly computation of required unknown macroscopic quantities.





**Figure 2.3.1:** Schematic of an RVE-based computational homogenization scheme. Figure taken from [93] with permission from John Wiley & Sons, Ltd.

$$\begin{aligned}
 -\nabla \cdot (a^0 \nabla u^0) &= f & \text{in } \Omega, \\
 u^0 &= g_D & \text{on } \partial\Omega_D, \\
 n \cdot (a^0 \nabla u^0) &= g_N & \text{on } \partial\Omega_N,
 \end{aligned} \tag{2.4.2}$$

as introduced in Section 2.1.2 and where the homogenized tensor  $a^0(x)$  is symmetric and uniformly elliptic and bounded, see Remark 7.

The finite element heterogeneous multiscale method (FE-HMM) aims at capturing the homogenized (coarse) solution  $u^0$  of (2.4.2) (with appropriately changed boundary conditions) without computing  $a^0(x)$  explicitly and by relying only on input data given by (2.4.1). Furthermore, the knowledge of the oscillatory data of (2.4.1) is only needed on sampling domains. This is important since in practice, for many applications, the fine scale structure of the problem can be obtained only on part of the computational domain. We mention problems in material science, where the fine scale structure may only be accessible locally by modern scanning and microscopy techniques or in geosciences where required knowledge of a landscape may only be sparsely available. The basic FE-HMM can be described as follows.

**Macro finite element space.** We consider the macro finite element space

$$V_D^p(\Omega, \mathcal{T}_H) = \{v^H \in H_D^1(\Omega); u^H|_K \in \mathcal{R}^p(K), \forall K \in \mathcal{T}_H\}, \tag{2.4.3}$$

with macro elements  $K \in \mathcal{T}_H$ , where  $\mathcal{T}_H$  is assumed to be shape regular. Here  $\mathcal{R}^p$  is defined as for (2.2.3). We emphasize that  $H$ , the size of the macro triangulation, is allowed to be much larger than  $\varepsilon$ .

**Quadrature formulas.** We follow the definition of the quadrature formula as in [15] and consider for each macro element  $K \in \mathcal{T}_H$  a  $C^1$ -diffeomorphism  $F_K$  such that  $K = F(\hat{K})$ , where  $\hat{K}$  is the (simplicial or rectangular) reference element. We consider a given quadrature formula  $\{\hat{x}_\ell, \hat{\omega}_\ell\}_{\ell=1}^{\mathcal{L}}$  on the reference element  $\hat{K}$ ; the corresponding integration points on an element  $K \in \mathcal{T}_H$  are given by  $x_{K_\ell} = F_K(\hat{x}_\ell)$ ,  $\ell = 1, \dots, \mathcal{L}$  and the corresponding quadrature weights on  $K$  are given by  $\omega_{K_\ell} = \hat{\omega}_\ell |\det(\partial F_K)|$ ,  $\ell = 1, \dots, \mathcal{L}$ . We make the following assumptions on the quadrature formulas:

**(Q1)** For  $\hat{\omega}_\ell > 0$ ,  $\ell = 1, \dots, \mathcal{L}$  and  $\hat{\lambda} > 0$  we have that

$$\sum_{\ell=1}^{\mathcal{L}} \hat{\omega}_\ell |\nabla \hat{q}(\hat{x}_\ell)|^2 \geq \hat{\lambda} \|\nabla \hat{q}\|_{L^2(\hat{K})}^2 \quad \forall \hat{q}(\hat{x}) \in \mathcal{R}^p(\hat{K}). \quad (2.4.4)$$

**(Q2)** We have that

$$\int_{\hat{K}} \hat{q}(\hat{x}) d\hat{x} = \sum_{\ell=1}^{\mathcal{L}} \hat{\omega}_\ell \hat{q}(\hat{x}_\ell) \quad \forall \hat{q}(\hat{x}) \in \mathcal{R}^\sigma(\hat{K}), \quad (2.4.5)$$

where  $\sigma = \max(2p-2, p)$  if  $\hat{K}$  is a simplicial element and where  $\sigma = \max(2p-1, p+1)$  if  $\hat{K}$  is a rectangular element.

We remark that there exist some well-known quadrature formulae that satisfy conditions **(Q1)** and **(Q2)**. For example, for piecewise linear elements  $\mathcal{L} = 1$ ,  $\omega_{K_\ell} = |K|$  and  $x_{K_\ell}$  is chosen to be located at the barycenter of the simplicial  $K$ . For bilinear elements,  $\mathcal{L} = 4$  and  $\{\omega_{K_\ell}, x_{K_\ell}\}_{\ell=1}^4$  is the two points Gauss quadrature rule given by  $\omega_{K_\ell} = |K|/4$ ,  $x_{K_\ell} = F_K(1/2 \pm \sqrt{3}/6, 1/2 \pm \sqrt{3}/6)$ . We refer to [6] for details.

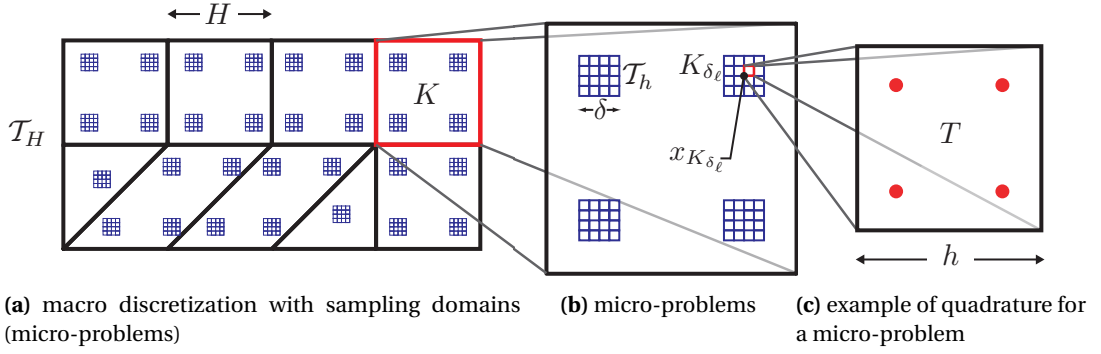
*Remark 8.* We will revisit quadrature rules in the context of higher order FE-HMM in Section 7.2.

Figure 2.4.1 illustrates the quadrature formulas and the coupling between the macro and micro problem which we describe in the following.

**Macro bilinear form.** For a discretization in the macro FE space (2.4.3) we need to modify the single-scale FEM bilinear form (2.2.4). For  $v^H, w^H \in V_D^p(\Omega, \mathcal{T}_H)$  we define

$$B(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_{\delta_\ell}}^h \cdot \nabla w_{K_{\delta_\ell}}^h dx, \quad (2.4.6)$$

## 2.4 The Finite Element Heterogeneous Multiscale Method (FE-HMM)



**Figure 2.4.1:** In every element of the FE-HMM, the contribution to the stiffness matrix of the macro elements (a) is given by the solutions of the micro problems (b), which are computed using numerical quadrature on every micro element (c).

where  $v_{K_\ell}^h, w_{K_\ell}^h$  are appropriate micro functions defined on sampling domains  $K_{\delta_\ell}$  (see below) and the factor  $|K_{\delta_\ell}|$  gives the appropriate weight for the contribution of the integrals defined on  $K_{\delta_\ell}$  instead of  $K$ .

*Remark 9.* The FE-HMM bilinear form (2.4.6) relies on numerical quadrature. Therefore, in our a posteriori analysis in Part II of this thesis, we need to specially take care of the corresponding variational crimes.

**Micro solver.** For every macro element  $K$  we compute the additive contribution to the macro stiffness matrix by computing micro functions  $v_{K_\ell}^h$  (and  $w_{K_\ell}^h$ ) obtained by solving micro-problems on sampling domains  $K_{\delta_\ell}$ ,  $\ell = 1, \dots, \mathcal{L}$ . The micro-problems read as follows: find  $v_{K_\ell}^h$  such that  $(v_{K_\ell}^h - v_{\text{lin}, K_\ell}^H) \in S^q(K_{\delta_\ell}, \mathcal{T}_h)$  and

$$\int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_\ell}^h \cdot \nabla z^h dx = 0 \quad \forall z^h \in S^q(K_{\delta_\ell}, \mathcal{T}_h), \quad (2.4.7)$$

where

$$v_{\text{lin}, K_\ell}^H(x) = v^H(x_{K_\ell}) + (x - x_{K_\ell}) \cdot \nabla v^H(x_{K_\ell}) \quad (2.4.8)$$

is a linearization of the macro function  $v^H$  at the integration point  $x_{K_\ell}$  (see [53, 6] for details) and

$$S^q(K_{\delta_\ell}, \mathcal{T}_h) = \{z^h \in W(K_{\delta_\ell}); z^h|_T \in \mathcal{R}^q(T), T \in \mathcal{T}_h\}, \quad (2.4.9)$$

where  $W(K_{\delta_\ell})$  determines the coupling condition or boundary conditions used for computing the micro functions  $v_{K_\ell}^h$  (or  $w_{K_\ell}^h$ ).

*Remark 10.* The choice of (2.4.7) is motivated as follows. In the periodic case, the micro problem (2.4.7) resembles the (discrete version of the) cell problem (2.1.18) if we choose  $\nabla v_{\text{lin}, i}^H = \mathbf{e}_i$ ,  $i = 1, \dots, d$ . Then the solution  $v_i^h$  of (2.4.7) resembles the FE solution of the cell

problem  $\chi^{h,i}$  as  $v_i^h - v_{\text{lin},i}^H = \chi^{h,i}$ . We will further see in Section 6.3.2 that

$$\frac{|K|}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_\ell}^h \cdot \nabla w_{K_\ell}^h dx = \int_K a_{K_\ell}^0 \nabla v_{\text{lin},K_\ell}^H \cdot \nabla w_{\text{lin},K_\ell}^H dx,$$

where  $a_{K_\ell}^0$  refers to the numerically homogenized tensor.

We will again set  $q = 1$  and since the sampling domain has a simple geometry, we will use quadrilateral FE. Notice that for piecewise linear functions  $v_{\text{lin},K_\ell}^H = v^H$ . Several choices are possible for the coupling condition we will consider

- the *periodic* coupling

$$W(K_{\delta_\ell}) = W_{\text{per}}^1(K_{\delta_\ell}) \tag{2.4.10}$$

and we denote  $S^q(K_{\delta_\ell}, \mathcal{T}_h)$  by  $S_P^q(K_{\delta_\ell}, \mathcal{T}_h)$ ; this coupling condition will be referred in what follows as (P),

- the *Dirichlet* coupling

$$W(K_{\delta_\ell}) = H_0^1(K_{\delta_\ell}) \tag{2.4.11}$$

and we denote  $S^q(K_{\delta_\ell}, \mathcal{T}_h)$  by  $S_D^q(K_{\delta_\ell}, \mathcal{T}_h)$ ; this coupling condition will be referred in what follows as (D),

where  $W_{\text{per}}^1(K_{\delta_\ell}) = \left\{ v \in H_{\text{per}}^1(K_{\delta_\ell}); \int_{K_{\delta_\ell}} v dx = 0 \right\}$ , and  $H_{\text{per}}^1(K_{\delta_\ell})$  is defined as the closure of  $\mathcal{C}_{\text{per}}^\infty(K_{\delta_\ell})$  (the subset of  $\mathcal{C}^\infty(\mathbb{R}^n)$  of periodic functions in  $K_{\delta_\ell}$ ).

*Remark 11.* For a tensor  $a^\varepsilon(x) = a(x, x/\varepsilon)$  with explicit scale separation, it is preferable to collocate the slow variable at the integration points  $a(x_{K_\ell}, x/\varepsilon)$  in both the macro and micro bilinear forms (2.4.6) and (2.4.7). In the periodic case, choosing  $\delta$  as an integer multiple of  $\varepsilon$  gives robust, i.e. independent of  $\varepsilon$ , convergence results (see [14, App. A]).

**Variational problem.** The macro solution of the FE-HMM is defined by the following variational problem: find  $u^H \in V^p(\Omega, \mathcal{T}_H)$  such that

$$B(u^H, v^H) = \int_\Omega f v^H dx + \int_{\partial\Omega_N} g_N v^H dx - B(g_D, v^H) \quad \forall v^H \in V^p(\Omega, \mathcal{T}_H). \tag{2.4.12}$$

The primary goal of this method is to capture the effective solution  $u^0$  of (2.4.2), i.e.,  $u^H$  converges to  $u^0$  as  $H$  goes to zero. Remember that the method also depends on a micro mesh, thus  $h$  going to zero is also necessary for the above convergence (see Section 2.4.1).

**Ellipticity, boundedness and well-posedness.** We have that the HMM bilinear form (2.4.6) is elliptic and bounded, i.e. it satisfies

$$B(v^H, v^H) \geq C \|v^H\|_{H^1(\Omega)}^2$$

and

$$|B(v^H, w^H)| \leq C \|v^H\|_{H^1(\Omega)} \|w^H\|_{H^1(\Omega)}$$

for all  $v^H, w^H \in V^p(\Omega, \mathcal{T}_H)$  with a constant  $C$  that only depends on the quadrature formula and the domain  $\Omega$ . The proof can be found in [6, Lemma 5] and relies on the inequality

$$\|\nabla v_{\text{lin}}^H\|_{L^2(K_{\delta_\ell})} \leq \|\nabla v^h\|_{L^2(K_{\delta_\ell})} \leq C \|\nabla v_{\text{lin}}^H\|_{L^2(K_{\delta_\ell})}$$

and quadrature formula **(Q1)** given in (2.4.4).

With the Lax-Milgram Theorem, there immediately follows the existence and uniqueness of the solution  $u^H$  of problem (2.4.12).

**Post-processing procedure.** While  $u^H$  converges to  $u^\varepsilon$  in the  $L^2$  norm, a convergence of  $u^H$  to  $u^\varepsilon$  in the energy ( $H^1$ ) norm is not guaranteed in general. Indeed, the oscillations of  $u^\varepsilon$  introduce  $\mathcal{O}(1)$  perturbations in the gradient which are not captured by  $u^H$ . An approximation of  $u^\varepsilon$  in the energy norm can nevertheless be obtained from the macro solution  $u^H$  by using a post-processing procedure. The known small scale solution in the sampling domain (2.4.7) computed during the assembly of the FE-HMM can be extended locally on the macro element  $K$  and added to  $u^H$ . This is done as follows (for simplicial elements): define  $u^{H,\varepsilon} : \Omega \rightarrow \mathbb{R}$  given by its restriction on each macro element  $K$  by

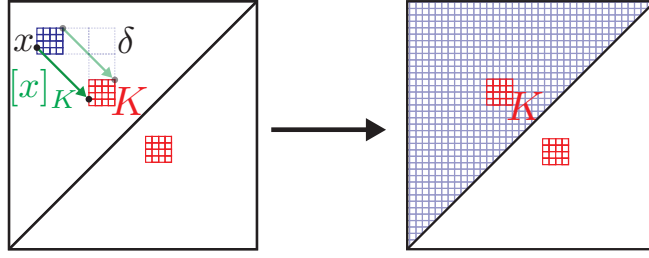
$$u^{H,\varepsilon}(x)|_K = u^H(x) + \left(u^h - u^H\right)(x - [x]_{K_\ell}) \text{ for } x \in K \in \mathcal{T}_H, \quad (2.4.13)$$

where for  $x \in \mathbb{R}^d$ ,  $[x]_{K_\ell}$  denotes the unique combination  $\delta \sum_{i=1}^d b_i e_i$ , where  $b_i \in \mathbb{Z}$  and  $(e_i)_{i=1}^d$  is the canonical basis of  $\mathbb{R}^d$ , such that  $(x - [x]_{K_\ell}) \in K_{\delta_\ell}$ , see Figure 2.4.2 for an illustration. Notice that  $u^{H,\varepsilon}(x)$  can be discontinuous across the macro elements  $K$ .

### 2.4.1 *A priori* estimates

We now list all the sources of errors of the numerical FE-HMM scheme, which can be obtained from an *a priori* estimate [6, 53, 2]. We follow the framework of [6] and decompose the error into three parts:

$$\|u^0 - u^H\| \leq \underbrace{\|u^0 - u^{0,H}\|}_{err_{mac}} + \underbrace{\|u^{0,H} - \bar{u}^H\|}_{err_{mod}} + \underbrace{\|\bar{u}^H - u^H\|}_{err_{mic}}, \quad (2.4.14)$$



**Figure 2.4.2:** Post-processing procedure: the small scale solution of the sampling domain is extended locally on the macro element  $K$  as given in eq. (2.4.13).

where  $err_{mac}$ ,  $err_{mod}$  and  $err_{mic}$  are the macro, modeling and micro error, respectively. Here,  $u^0$  is the solution of (2.1.7),  $u^H$  is the FE-HMM solution of (2.4.12),  $u^{0,H}$  is the FEM solution of (2.1.7) in the space  $V^p(\Omega, \mathcal{T}_H)$  and  $\bar{u}^H$  is the FE-HMM solution of (2.4.12) with exact micro functions (in  $W(K_{\delta_\ell})$ ).

The various components can be estimated as follows:

$err_{mac}$  is the macroscopic error from a standard finite element approximation  $u^{0,H}$  of (2.4.2) in  $V^p(\Omega, \mathcal{T}_H)$  with numerical quadrature (see e.g. [44]). The error can be estimated in the  $H^1$ -norm as

$$err_{mac} = \|u^0 - u^{0,H}\|_{H^1(\Omega)} \leq CH^p$$

and in the  $L^2$ -norm as

$$err_{mac} = \|u^0 - u^{0,H}\|_{L^2(\Omega)} \leq CH^{p+1}.$$

$err_{mic}$  is the error caused by the FE approximation of the micro problems in  $S^q(K_{\delta_\ell}, \mathcal{T}_h)$ . The micro error adds to the total error through the macro bilinear form. Under appropriate assumptions on  $a^\varepsilon$  (e.g. regularity, local periodicity) and appropriate boundary conditions we get

$$err_{mic} = \|\bar{u}^H - u^H\|_{H^1(\Omega)} \leq C \left(\frac{h}{\varepsilon}\right)^{2q}, \quad err_{mic} = \|\bar{u}^H - u^H\|_{L^2(\Omega)} \leq C \left(\frac{h}{\varepsilon}\right)^{2q},$$

for both the  $H^1$  and the  $L^2$ -norm. We refer to [6] for details.

$err_{mod}$  is the error introduced by the modeling of the multiscale scheme, i.e. the upscaling procedure and the coupling between the macro and micro FEMs. While the modeling error depends on the structure of the tensor  $a^\varepsilon$ , the boundary conditions of the micro problem (i.e., the coupling of the micro to the macro FE spaces) and the sampling domain size  $\delta$ , it is independent of the macro and micro mesh size  $H$  and  $h$ , respectively. Under the assumption of a two-scale tensor of form  $a^\varepsilon(x) = a(x, \frac{x}{\varepsilon})$ , the general form of

the modeling error for piecewise linear macro FE is

$$err_{mod} \leq C \sup_{K \in \mathcal{T}_H, x_\ell \in K} \|a^0(x_{K_\ell}) - \bar{a}_K^0\|_F,$$

where the exact homogenized tensor  $a^0(x)$  (evaluated at the quadrature node  $x_{K_\ell}$ ) is defined in Theorem 5 and  $\bar{a}_K^0$  is an appropriate numerically upscaled tensor, obtained from a procedure similar to the FE-HMM (see (6.3.14) for a proper definition). If we have a periodic tensor  $a^\varepsilon = a(x, x/\varepsilon)$ , periodic coupling, sampling domains covering an integer number of periods, i.e.  $\delta = n\varepsilon$ ,  $n \in \mathbb{N}$  and collocation of the slow variable, i.e.  $a(x_{K_\ell}, x/\varepsilon)$  we get that the modeling error is  $err_{mod} = 0$ . We refer to Section 6.3.5.2 and [53, 6] and references therein for a more detailed description of the modeling error.

The general form of the *a priori* error estimate in the  $H^1$ -norm under the appropriate assumptions of the two-scale tensor  $a^\varepsilon(x) = a(x, \frac{x}{\varepsilon})$  is therefore

$$\|u^0 - u^H\|_{H^1(\Omega)} \leq C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} \right) + err_{mod}. \quad (2.4.15)$$

**Finescale error.** In order to estimate  $\|u^\varepsilon - u^0\|_{L^2(\Omega)} \leq \|u^\varepsilon - u^H\|_{L^2(\Omega)} + \|u^H - u^0\|_{L^2(\Omega)}$  in the  $L^2$ -norm, we use the estimate (2.1.19) and the triangle inequality and we obtain the following additional term in (2.4.14)

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

In the  $H^1$ -norm,  $u^0$  does not converge to  $u^\varepsilon$  and therefore  $u^H$  does not converge to  $u^\varepsilon$ . Indeed,  $u^0$  (or  $u^H$ ) can not capture the micro oscillation of  $u^\varepsilon$ , which results in an  $\mathcal{O}(1)$  mismatch when gradients are estimated. One can however create a post-processed solution  $u^{H,\varepsilon}$  based on the macro solution  $u^H$  and the micro solutions  $u_K^h$  (available in the sampling domains  $K_{\delta_\ell} \subset K$ ) using ideas from the construction of correctors in homogenization theory [32]. An estimate for the post-processed FE-HMM solution  $u^{H,\varepsilon}$  (illustrated in Figure 2.4.2) does exist and, under appropriate assumptions and for piecewise linear macro and micro FE, satisfies [2, 5]

$$\|u^\varepsilon - u^{H,\varepsilon}\|_{\bar{H}^1(\Omega)} \leq C \left( H + \sqrt{\varepsilon} + \frac{h}{\varepsilon} \right), \quad (2.4.16)$$

where  $\bar{H}^1(\Omega)$  is a broken norm defined by  $\|u\|_{\bar{H}^1(\Omega)} = \left( \sum_{K \in \mathcal{T}_H} \|\nabla u\|_{L^2(K)}^2 \right)^{1/2}$  (as the reconstructed solution may have jumps on the interfaces of two adjacent elements). The term  $\sqrt{\varepsilon}$  in (2.4.16) originates from a boundary layer; this boundary layer is caused as the corrector does not satisfy the appropriate boundary conditions on  $\partial\Omega$ . Notice that in (2.4.16) we have a term  $\frac{h}{\varepsilon}$ , while in the estimate  $\|u^0 - u^H\|_{H^1(\Omega)}$  we have for piecewise linear macro and micro FE the micro error can be estimated as  $err_{mic} < C \left( \frac{h}{\varepsilon} \right)^2$ .





## 3 Implementation of the FE-HMM and Simulation of Heat Transfer Problems

Despite the flourishing activity on computational homogenization problems or related multiscale problems (see the references in Chapter 1), very few open-box computer codes are currently available in this area. The goal of this chapter is to present a short finite element (FE) implementation for multiscale elliptic or parabolic problems. For convenience, we propose a MATLAB implementation, which allows for a concise coding. A FORTRAN or C/C++ version could be easily obtained following the structure of the code given in this paper.

The algorithm proposed in this Chapter is short (less than 200 lines) and flexible, easy to use and to modify. At the same time, it is capable to handle challenging 3D problems as can be seen in Section 3.2, where we describe the heat conduction in a microprocessor package made of composite materials. Generalization for elastic problems can be done following the implementation described here. We believe that such a code can be useful for many practical applications since it can easily be modified by users to account for specific features of their problems. Our multiscale code can handle

- simplicial and quadrilateral FE (at the macroscopic level, i.e., for meshing the physical domain),
- various coupling conditions between macroscopic and microscopic solvers,
- time dependent (parabolic) and 3D problems.

As we will see, a parallelization of the algorithm is straightforward. We propose an implementation for problems with two scales. This could however be generalized to problems with several separated scales. For simplicity, we also limit ourselves to piecewise (bi)linear FE approximations. The proposed code can easily be extended to higher order FE approximation and the algorithm discussed in the first part of the paper is written in this general setting. The code and implementation discussed is available at <http://anmc.epfl.ch>.

The outline of this Chapter is as follows. We present the FE algorithm and implementation in detail in Section 3.1. Various numerical examples in two and three dimensions (including parabolic problems) are given in Section 3.2, where we use different types of elements and both deterministic and random oscillating coefficients. There we present an application of our code to the steady-state heat distribution in a heat sink and a microprocessor featuring a

multiscale heat conductivity tensor; this simulation resembles the motivating engineering problem we stated in the introduction of this thesis in Chapter 1, see also Figure 1.0.2.

The results in this Chapter were published in [11] and Sections 3.1 and 3.2 are reprinted from [11, Section 3 and 4], with permission from Elsevier. The introduction of this Chapter is based on [11, Section 1]. The references and citations were updated in order to fit into the framework of this thesis.

### 3.1 Implementation

For simplicity of notations, we describe the implementation for two-dimensional problems, and comment on the subtleties which occur for three-dimensional problems when appropriate. We emphasize that our code can handle both two and three-dimensional problems. We also restrict ourselves to a piecewise linear or bilinear macro FE approximation and emphasize again that modification of our solver to accommodate higher order FEM can easily be obtained.

**Data representation of the macro triangulation.** We recall that a data-representation of a triangulation  $\mathcal{T}_h$  of a domain  $G$  requires the following input data

- `Coordinates.dat` containing in line  $i$ :  $x$ -coordinate,  $y$ -coordinate of node number  $i$ ,
- `Elements3.dat` and `Elements4.dat` containing in line  $i$  the node numbers of the  $i^{th}$  triangular or quadrilateral elements,  
node1 node2 node3 (for a triangle) and  
node1 node2 node3 node4 (for a quadrilateral),  
respectively.
- `Dirichlet.dat` and `Neumann.dat` containing the node numbers of boundary nodes.

If  $G = \Omega$  we will use the above notation, if  $G = K_\delta$  (sampling domain) we will use the prefix `Micro...`. The data representation used is standard unless explicitly mentioned and we refer for example to [22] for details.

**Macro and micro basis functions.** In what follows we denote by

- $\{\varphi_m^H\}_{m=1}^{M_{mac}}$  the basis of the macro FE space  $V^1(\Omega, \mathcal{T}_H)$  defined in (2.4.3), where  $M_{mac}$  is the number of discretization points in the macro domain; the elements (simplicial or quadrilateral) of the macro triangulation  $\mathcal{T}_H$  are denoted by  $K$ ,
- $\{\psi_{m,K_\ell}^h\}_{m=1}^{M_{mic}}$  the basis of the micro FE space  $S^1(K_{\delta_\ell}, \mathcal{T}_h)$  defined in (2.4.9) with  $q = 1$ , where  $M_{mic} = (N_{mic})^d$  ( $d = 2, 3$ ) and  $N_{mic}$  is the number of discretization points in each direction of the  $d$ -dimensional micro sampling domain; the elements (simplicial or quadrilateral) of the micro triangulation  $\mathcal{T}_h$  are denoted by  $T$ .

### 3.1.1 Core structure: assembling the macro problem

Following a standard assembly process for the macro bilinear form (2.4.6), we compute the contribution  $A_K$  to the stiffness matrix associated with the macro element  $K$ . We identify the macro basis functions with non-zero support in  $K$  denoted by  $\varphi_i^H$ ,  $i = 1, \dots, \mu_K$  and determine  $A_K$  based on contributions computed on sampling domains  $K_{\delta_\ell} \subset K$ . The process reads as follows

$$A_K := B \left( \varphi_i^H, \varphi_j^H \right)_{i,j=1}^{\mu_K} = \left( \sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla \varphi_{i,K_\ell}^h \cdot \nabla \varphi_{j,K_\ell}^h dx \right)_{i,j=1}^{\mu_K} \quad (3.1.1)$$

$$= \underbrace{\sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K_{\delta_\ell}|} \left( \mathcal{A}_{K_\ell}^T A_{mic,K_\ell} \mathcal{A}_{K_\ell} \right)}_{\text{hmm\_stima\_type\_of\_K}}, \quad (3.1.2)$$

where `type_of_K` refers to the type of the macro element  $K$ , i.e., `type_of_K=tri` for triangular elements or `type_of_K=quad` for quadrilateral elements. Notice that  $A_{mic,K_\ell}$  is a  $M_{mic} \times M_{mic}$  matrix and  $\mathcal{A}_{K_\ell}$  a  $M_{mic} \times \mu_K$  matrix. Furthermore,  $\varphi_{i,K_{\delta_\ell}}^h$  is the micro function solution of (2.4.7) such that  $(\varphi_{i,K_{\delta_\ell}}^h - \varphi_{lin,i,K_{\delta_\ell}}^H) \in S^1(K_{\delta_\ell}, \mathcal{T}_h)$  and  $\varphi_{lin,i,K_{\delta_\ell}}^H$  is the linearization (2.4.8) of the basis function  $\varphi_i^H$ .

We then compute the contribution to the right-hand side for the triangle  $K$  as

$$b_K = \int_K f \varphi_i^H dx \simeq f(x_{K_b}) \int_K \varphi_i^H dx, \quad (3.1.3)$$

where  $x_{K_b} \in K$  is an integration point located at the barycenter of  $K$  and  $\int_K \varphi_i^H dx$  can be computed explicitly (its value is for example  $|K|/6$  if  $K$  is a triangle or  $|K|/4$  if  $K$  is a quadrilateral). Upon a suitable numbering of the macro nodes, the system corresponding to the macro solution can be written as

$$\underbrace{\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}}_A \begin{pmatrix} U \\ U_D \end{pmatrix} = \underbrace{\begin{pmatrix} b_I \\ b_D \end{pmatrix}}_b, \quad (3.1.4)$$

where  $U$  are the free macro nodes to be computed and  $U_D$  the values (known *a priori*) of the nodes at the portion of the boundary enforced with Dirichlet conditions. The first block of the above system of equations gives the desired macro solution

$$A_{11}U = b_I - A_{12}U_D.$$

The above procedure as well as the incorporation of the boundary conditions for the macro problem are standard and we follow the structure of the MATLAB implementation described in [22]. A flow chart of the algorithm is shown in Figure 3.1.1.

While for classical FEM, `stima_tri` or `stima_quad` are given by a suitable element-wise integration, for the FE-HMM, `stima_tri` or `stima_quad` are given by a suitable contribution

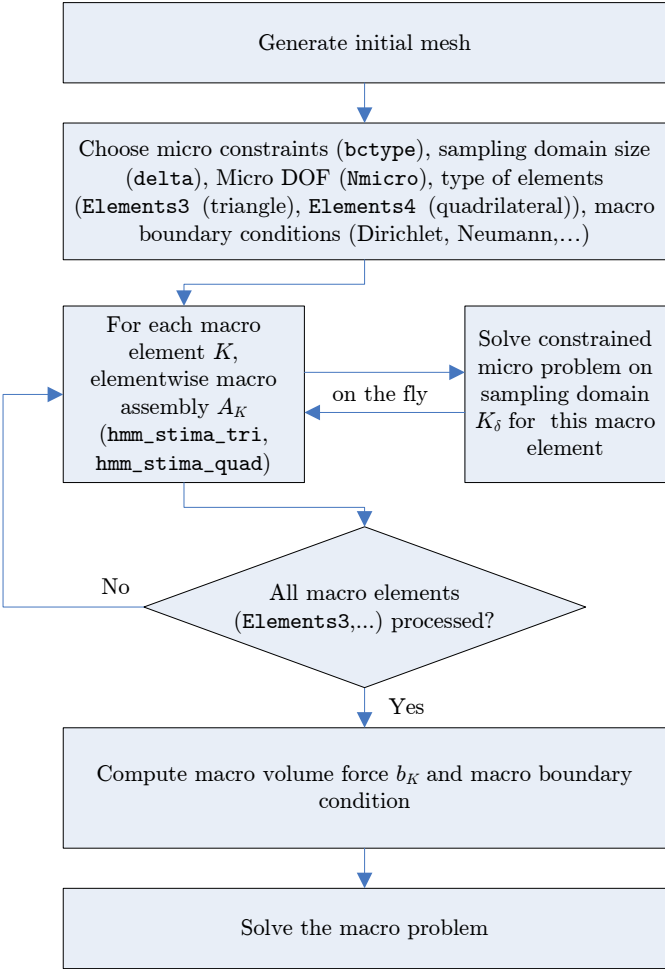


Figure 3.1.1: Flow chart for FE-HMM algorithm.

of micro functions computed on sampling domains, constrained by the basis functions associated with the element  $K$ . This is explained in detail in the next section.

We note that the solution of the constrained micro problem on one macro element is independent of the constrained micro problem on a different macro element. The macro assembly is thus well-suited for parallel implementation.

#### 3.1.2 Local assembly and micro solutions on sampling domains

In order to compute the matrix  $A$  of (3.1.4) we have to compute for each macro element  $K$  the local macro contribution  $A_K$  (see (3.1.2)) to the stiffness matrix. This assembly is based upon suitable solutions of micro problems on sampling domains  $K_{\delta_\ell} \subset K$ .

**Data representation of the micro triangulation.** Since the sampling domains  $K_{\delta_\ell}$  have a simple shape (square or cube) we choose a quadrilateral mesh  $\mathcal{T}_h$  with  $M_{mic} = (N_{mic})^d$  ( $d = 2, 3$ ) grid points. The `MicroCoordinates` must be computed around the integration points of each macro element and is implemented in `micromesh_coords`. The `MicroElements` do not depend on the `MicroCoordinates` and are computed once at the initialization of the code, which is done in `micromesh_elements`.

**Structure of the micro assembly process.** To have the flexibility to incorporate various micro boundary conditions we will use Lagrange multipliers to compute  $\varphi^h$ . Given a sampling domain  $K_{\delta_\ell} \subset K$  and an associated linearized macro basis function  $\varphi_{lin,i,K_{\delta_\ell}}^H$  we consider instead of (2.4.7) the minimization problem

$$\varphi_{i,K_{\delta_\ell}}^h = \operatorname{argmin} \int_{K_{\delta_\ell}} a(x_{K_{\delta_\ell}}, \frac{x}{\varepsilon}) \nabla w^h \cdot \nabla w^h dx \quad (3.1.5)$$

over all functions  $w^h \in S^1(K_{\delta_\ell}, \mathcal{T}_h) = \{z^h \in H^1(K_{\delta_\ell}); z^h|_T \in \mathcal{R}(T), \forall T \in \mathcal{T}_h\}$ , such that

- $w^h - \varphi_{lin,i,K_{\delta_\ell}}^H \in S_D^1(K_{\delta_\ell}, \mathcal{T}_h)$  (periodic coupling) or
- $w^h - \varphi_{lin,i,K_{\delta_\ell}}^H \in S_D^1(K_{\delta_\ell}, \mathcal{T}_h)$  (Dirichlet coupling).

*Remark 12.* If we do not have explicit scale separation, i.e.  $a^\varepsilon(x)$  instead of  $a(x_{K_{\delta_\ell}}, \frac{x}{\varepsilon})$ , we can insert a dummy parameter for  $x_{K_{\delta_\ell}}$ , such as  $x_{K_{\delta_\ell}} = 0$  and only use the second parameter  $\frac{x}{\varepsilon}$  in the tensor function of the MATLAB code.

To compute (3.1.5) we expand  $\varphi_{lin,i,K_{\delta_\ell}}^H$  in the basis of  $S^1(K_{\delta_\ell}, \mathcal{T}_h)$

$$\varphi_{lin,i,K_{\delta_\ell}}^H = \sum_{m=1}^{M_{mic}} \beta_{i,m} \psi_{m,K_{\delta_\ell}}^h(x) \quad (3.1.6)$$

### Chapter 3. Implementation of the FE-HMM and Simulation of Heat Transfer Problems

and the above minimization problem leads, by introducing Lagrange multipliers, to a saddle point problem given in linear form by (see [14, Sect. 5])

$$\begin{aligned} A_{mic,K_{\delta_\ell}} \boldsymbol{\alpha}_{i,K_{\delta_\ell}} + D^T \boldsymbol{\lambda} &= 0 \\ D(\boldsymbol{\alpha}_{i,K_{\delta_\ell}} - \boldsymbol{\beta}_{i,K_{\delta_\ell}}) &= 0, \end{aligned} \quad (3.1.7)$$

where  $\boldsymbol{\alpha}_{i,K_{\delta_\ell}} = (\alpha_{i,1}, \dots, \alpha_{i,M_{mic}})^T$ ,  $\boldsymbol{\beta}_{i,K_{\delta_\ell}} = (\beta_{i,1}, \dots, \beta_{i,M_{mic}})^T$ , and  $\boldsymbol{\lambda}$  is the Lagrange multiplier. We then have

$$\boldsymbol{\varphi}_{i,K_{\delta_\ell}}^h = \sum_{m=1}^{M_{mic}} \alpha_m \boldsymbol{\psi}_{m,K_{\delta_\ell}}^h. \quad (3.1.8)$$

$A_{mic,K_{\delta_\ell}}$  is the micro stiffness matrix with entries given by

$$\left( A_{mic,K_{\delta_\ell}} \right)_{mn} = \int_{K_{\delta_\ell}} a(x_{K_{\delta_\ell}}, \frac{x}{\varepsilon}) \nabla \boldsymbol{\psi}_{m,K_{\delta_\ell}}^h \cdot \nabla \boldsymbol{\psi}_{n,K_{\delta_\ell}}^h dx, \quad (3.1.9)$$

and  $D$  is the matrix of the constraints detailed below.

**Assemble and solve the constrained micro problem.** The problem (3.1.7) is solved for each basis function  $\varphi_{lin,i,K_{\delta_\ell}}^H$ ,  $i = 1, \dots, \mu_K$  and we define the  $M_{mic} \times \mu_K$  matrix  $\mathcal{A}_{K_\ell}$  with columns given by  $\boldsymbol{\alpha}_{i,K_{\delta_\ell}}$ , i.e.,  $\mathcal{A}_{K_\ell} = (\boldsymbol{\alpha}_{1,K_{\delta_\ell}}, \dots, \boldsymbol{\alpha}_{\mu_K,K_{\delta_\ell}})$ . This is done for each quadrature node of the macro element and we obtain

$$A_K = \underbrace{\sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K_{\delta_\ell}|} \left( \mathcal{A}_{K_\ell}^T A_{mic,K_\ell} \mathcal{A}_{K_\ell} \right)}_{\text{hmm\_stima\_type\_of\_K}}$$

as defined in (3.1.2). The corresponding MATLAB code is given here for quadrilateral macro elements (hence the factor 0.25), i.e., type\_of\_K=quad,

```

68 for node_no=1:4
69     % select current quadrature nodes
...
105     % Assemble matrix AConstr for constrained system and corresponding rhs
106     AConstr=[A ConstraintMat'];...
107     ConstraintMat sparse(size(ConstraintMat,1),size(ConstraintMat,1));
108     RhsConstr=[zeros(NoOfNodes,4);ConstraintRhs];
109
110     % solve constrained microproblem
111     x=AConstr\RhsConstr;
112     alpha=x(1:size(MicroCoordinates,1),:);
...
118     %The contribution to the macro stiffness matrix
119     A_macro=A_macro+ .25 * K_macro/K_micro *alpha'*A*alpha;
120 end

```

where `A_macro` and `A` in the above code are denoted  $A_K$  and  $A_{mic}$ , respectively, in the above discussion.

The main steps of `hmm_stima_type_of_K` are

- to assemble the micro stiffness matrix (see Section 3.1.2.1),
- to incorporate the micro coupling constraints (P) or (D) (see Section 3.1.2.2).

These steps are explained in the two following sections. We will mainly concentrate on a given sampling domain and skip the index corresponding to the particular macro integration point. We will thus write  $K_\delta, \varphi_i^h, \varphi_{lin,i}^H, \psi_m^h, x_K$  instead of  $K_{\delta_l}, \varphi_{i,K_{\delta_\ell}}^h, \varphi_{lin,i,K_{\delta_\ell}}^H, \psi_{m,K_{\delta_\ell}}^h, x_{K_{\delta_\ell}}$ . When no confusion can occur, we will also skip the underscore index for the micro functions  $\varphi_i^h, i = 1, \dots, \mu_K$  and simply write  $\varphi^h$  for a given micro function and  $\varphi^H$  for its corresponding macro function.

#### 3.1.2.1 Assembly of the micro stiffness matrix

Gauss quadrature rules are used to evaluate the integral (3.1.9) numerically. This is done in a standard way in `hmm_micro_stima_quad`

$$(A_{mic})_{\ell m} = \sum_{T \in \mathcal{T}_h} \int_T a\left(x_k, \frac{x}{\varepsilon}\right) \nabla \psi_\ell^h \cdot \nabla \psi_m^h dx.$$

For each micro element  $T$  the contribution to the micro stiffness matrix is evaluated via

$$\sum_{i=1}^4 \frac{1}{4} \int_{\hat{T}} a\left(x_k, \frac{F_T(\xi)}{\varepsilon}\right) \left(D_{E,T}^{-T} \nabla \hat{\psi}_\ell^h(\xi)\right) \cdot \left(D_{E,T}^{-T} \nabla \hat{\psi}_m^h(\xi)\right) |\det D_{E,T}| d(\xi),$$

where  $F_T(\xi) = x_0 + D_{E,T} \xi$  is the affine mapping (here we only allow parallelograms for simplicity) from the reference element  $\hat{T} = [0, 1]^2$  onto  $T \in \mathcal{T}_h$  with  $D_{E,T}$  given by `Map`,  $\xi_i$  are determined by the two points Gauss quadrature rule on the reference element defined in `quadnodes`, and  $\hat{\psi}_l, \hat{\psi}_m$  are the usual bilinear shape functions on the reference square  $\hat{T}$  with derivatives computed in `phi_prime_hat`. We note that the slow variable  $x_K$  in  $a\left(x_k, \frac{x}{\varepsilon}\right)$  is kept fixed in the micro problem. The tensor  $a\left(x_k, \frac{x_i}{\varepsilon}\right)$  is evaluated in the function `tensor_a.m`. The MATLAB code for the micro stiffness matrix assembly is given by

```

1 function M = hmm_microstima_quad(vertices, ...
2     MacroQuadNode, epsilon)
...
47 % map from reference element
48 Map=[vertices(2,:) - vertices(1,:); vertices(4,:) - vertices(1,:)];
49
50 % gauss quadrature at the 4 quadrature nodes, reference element and mapped
51 % element
52 quadnodes_ref=[.5-sqrt(3)/6, .5+sqrt(3)/6, .5-sqrt(3)/6, .5+sqrt(3)/6; ...
53     .5-sqrt(3)/6, .5-sqrt(3)/6, .5+sqrt(3)/6, .5+sqrt(3)/6];
54 quadnodes= Map*quadnodes_ref+repmat(vertices(1,:),1,4);
55
56 M=zeros(4,4);
57 for node=1:4

```

```

58     quadnode=quadnodes(:,node);
59     x=quadnodes_ref(1,node);
60     y=quadnodes_ref(2,node);
61
62     phi_prime_hat=[y-1, x-1; 1-y, -x; y, x; -y, 1-x];
63     phi_prime_invD=(phi_prime_hat/Map);
64
65     % evaluate coefficient tensor at the specific nodes
66     a=tensor_a(MacroQuadNode, quadnode/epsilon);
67     M=M+(det(Map))*phi_prime_invD*a*phi_prime_invD';
68 end
69 M=M/4;
70
71 end

```

### 3.1.2.2 Micro boundary conditions and coupling constraints

The suitable constraints between macro and micro solvers (or boundary conditions for the micro problem) are encoded in the matrix  $D$  (see (3.1.7)). In our MATLAB implementation, we provide both Dirichlet or periodic boundary conditions for the micro problem. Other boundary conditions could easily be incorporated using the structure explained below. They can be selected by setting the variable `bctype` in the main file `hmmfem2d.m` to either `dirichlet` or `periodic`. The chosen boundary conditions are then selected in `hmm_stima_type_of_K`.

**Periodic micro constraints.** If  $(\varphi^h - \varphi_{lin}^H) \in S_P(K_\delta, \mathcal{T}_h)$  (see (2.4.9)) we must have

- (i)  $\int_{K_\delta} (\varphi^h - \varphi_{lin}^H) dx = 0$ ,
- (ii)  $(\varphi^h - \varphi_{lin}^H)(p) = (\varphi^h - \varphi_{lin}^H)(p')$ , for all  $L$  couples  $(p, p')$  of boundary nodes on opposite edges avoiding redundant couples at the corner.

In this situation, the matrix  $D$  has the form

$$D = \begin{pmatrix} b_1 & \cdots & b_{M_{mic}} \\ & \tilde{D} & \end{pmatrix}, \quad (3.1.10)$$

where the first row corresponds to the condition (i) and the  $L \times M_{mic}$  matrix  $\tilde{D}$  corresponds to the condition (ii). The first condition simply fixes the constant of the periodic boundary value problem and can be chosen arbitrarily. A convenient choice is  $\int_{K_\delta} \varphi^h dx = 0$ , i.e.,

$$\sum_{m=1}^{M_{mic}} \alpha_m \int_{K_\delta} \psi_m^h dx = \sum_{T \in \mathcal{T}_h} \sum_{m=1}^{M_{mic}} \alpha_m \underbrace{\int_T \psi_m^h dx}_{b_m} = \sum_{T \in \mathcal{T}_h} b^T \alpha = 0. \quad (3.1.11)$$

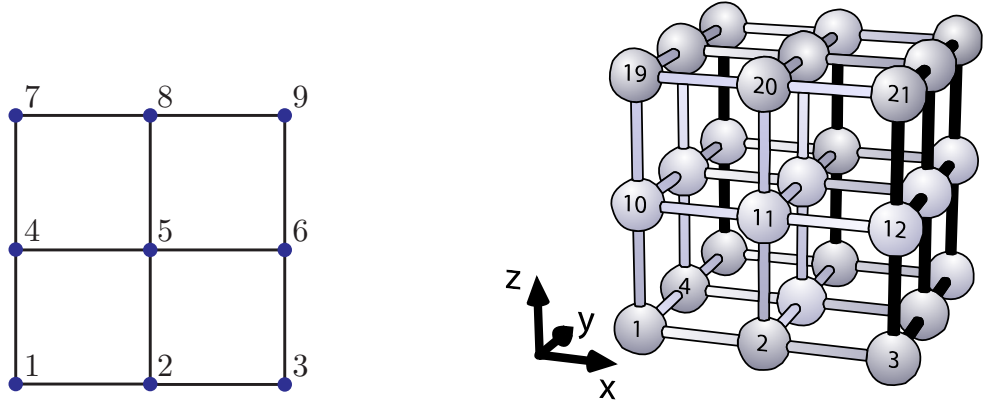
This is implemented in the MATLAB function `micro_constraints_periodic.m`:

```

98 NoOfNodes = size(MicroCoordinates,1);
99 b=sparse(NoOfNodes,1);
100
101 for j=1:size(MicroElements,1)

```





(a) Example of micro mesh. Marked couples of nodes are (1,7), (2,8), (3,9), (1,3), (4,6). Periodicity constraints for the couple (7,9) is redundant as it is implicitly enforced by (1,7), (1,3), (3,9).

(b) Example of a three dimensional micro mesh, black lines correspond to boundary node identifications that are redundant. For example periodicity for the couple (3,21) is implicitly enforced by (1,3), (1,19) and (19,21).

**Figure 3.1.2:** Micro domain and periodic boundary conditions in 2D and 3D

```

102     b(MicroElements(j,:))=b(MicroElements(j,:))+...
103         det([1,1,1; MicroCoordinates(MicroElements(j,1:3),:)]')/4;
104 end
    
```

To treat the periodicity conditions we select the  $L$  non-redundant couples at opposite sides of  $K_\delta$  as  $(p, p') = (p_{\mu(\ell)}, p_{\nu(\ell)})$ ,  $\ell = 1, \dots, L$ , where  $\mu(\ell)$  and  $\nu(\ell)$  are the node numbers of the marked couples. Using (3.1.8) we can express the periodicity condition for the  $\ell$ th constraint as

$$\varphi^h(p_{\mu(\ell)}) - \varphi^h(p_{\nu(\ell)}) = \sum_{m=1}^{M_{mic}} \tilde{D}_{\ell m} \alpha_m = \varphi_{lin}^H(p_{\mu(\ell)}) - \varphi_{lin}^H(p_{\nu(\ell)}) \quad (3.1.12)$$

where  $\varphi^h(p_{\mu(j)}) = \alpha_j$ , and the matrix  $\tilde{D} \in \mathbb{R}^{L \times M_{mic}}$  has entry  $\tilde{D}_{\ell j} = 1$  if  $j = \mu(\ell)$ ,  $\tilde{D}_{\ell j} = -1$  if  $j = \nu(\ell)$  and 0 otherwise. For the example sketched in Figure 3.1.2 (left picture) we get

$$\tilde{D} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}.$$

Notice that the nodes can be marked once at the beginning of the code. This is done in `make_constraints.m`

```

54     Constraints(:,1)=[1:N 1:N:N*N-N];
55     Constraints(:,2)=[(N-1)*N+1:N*N N:N:N*N-N];
    
```

and the computation of the matrix  $\tilde{D}$  is implemented in `micro_constraints_periodic.m` as follows:

```

70 NoOfPerNodes=size(Constraints,1);
71 ConstraintMat=sparse(repmat(1:NoOfPerNodes,1,2), ...
72     [Constraints(:,1) Constraints(:,2)], ...
73     [ones(NoOfPerNodes,1); -ones(NoOfPerNodes,1)]);

```

where we took advantage of the MATLAB function `repmat` which creates a large  $m \times n$  tiled matrix of which each tile is the replicated original matrix. Thus here, with  $m = 1$  and  $n = 2$  we get a vector  $(1, 2, \dots, \text{NoOfPerNodes}, 1, 2, \dots, \text{NoOfPerNodes})^T$ .

We calculate the right-hand side of (3.1.12) for all macro basis functions  $\varphi_j^H$  at the same time via  $\varphi_j^H(x) = \hat{\varphi}_j^H(F_K^{-1}(x))$  where  $F_T(\xi) = x_0 + D_{F,T}\xi$  is the affine mapping from the reference triangle  $\hat{K}$  onto  $K \in \mathcal{T}_h$ , where  $D_{F,T}$  is given by `MacroMap` and  $\hat{\varphi}_j^H$  is the macro reference-element shape function. The corresponding MATLAB code is

```

75 % coordinates of periodic boundary conditions
76 % x_node1 and x_node2 are corresponding periodic nodes
77 x_node1= MicroCoordinates(Constraints(:,1), :);
78 x_node2= MicroCoordinates(Constraints(:,2), :);
79
80 % coordinates in the reference quadrilateral corresponding to the nodes with
81 % periodic boundary conditions
82 x_node_ref1=(MacroMap\(x_node1-repmat(MacroVertices(1,:),NoOfPerNodes,1))');
83 x_node_ref2=(MacroMap\(x_node2-repmat(MacroVertices(1,:),NoOfPerNodes,1))');
84
85 % corresponding rhs
86 if (MacroNodes==3) % we have macro triangles
87     ConstraintRhs=shapefunction_tri(x_node_ref1)-...
88         shapefunction_tri(x_node_ref2);
89 end
90 if (MacroNodes==4) % we have macro quadrilaterals
91     ConstraintRhs=shapefunction_quad_lin(x_node_ref1, refquadnode)-...
92         shapefunction_quad_lin(x_node_ref2, refquadnode);
93 end

```

The macro basis functions  $\hat{\varphi}_j^H$  are evaluated in the functions `shapefunction_tri` and `shapefunction_quad_lin` for triangles and quadrilaterals, respectively. Finally, we merge the constraint matrix  $b$  from (3.1.11) and  $\tilde{D}$  from (3.1.12) to assemble the matrix  $D$  given in (3.1.10) as well as the right-hand side.

```

108 % merge both to one unified matrix / vector
109 ConstraintMat=[b';ConstraintMat];
110 ConstraintRhs=[zeros(1,MacroNodes);ConstraintRhs];

```

**Dirichlet micro constraints.** Other constraints for the micro solver, such as constraints through Dirichlet boundary conditions, can be implemented similarly as explained before. In our code we can switch from periodic to Dirichlet boundary constraints. For Dirichlet boundary constraints we have  $(\varphi^h - \varphi_{lin}^H) \in S_D(K_\delta, \mathcal{T}_h)$  (see (2.4.9)), and thus

$$(\varphi^h - \varphi_{lin}^H)(p_{\mu(\ell)}) = 0 \quad (3.1.13)$$

for all boundary nodes  $p_{\mu(1)}, \dots, p_{\mu(L)}$ . We get for the  $\ell$ th constraint

$$\sum_{m=1}^M D_{\ell m} \alpha_m = \varphi_{lin}^H(p_{\mu(\ell)}) \quad (3.1.14)$$

where the entries of  $D \in \mathbb{R}^{L \times M_{mic}}$  are given by  $D_{\ell m} = 1$  if  $m = \mu(\ell)$  and 0 otherwise. The micro boundary nodes  $p_{\mu(\ell)}$  are determined in `make_constraints.m` at the beginning of the code

```
60 Constraints=[1:N N*N-N+1:N*N 1+N:N:N*N-2*N+1 2*N:N:N*N-N];
```

and the computation of the matrix  $D$  and the right-hand side of (3.1.14) is performed similarly as in the periodic case in `micro_constraints_dirichlet.m`.

*Remark 13.* While we chose to present and discuss periodic and Dirichlet micro constraints, other constraints through Neumann or Robin boundary conditions could easily be implemented, following the procedure described for periodic or Dirichlet constraints.

### 3.1.3 Parabolic problems

As mentioned in the introduction, our code can handle parabolic problems without difficulties. Consider for example

$$\frac{\partial u^\varepsilon}{\partial t} = \nabla \cdot (a^\varepsilon \nabla u^\varepsilon) + f \text{ in } \Omega \times [t_0, T], \quad (3.1.15)$$

with initial conditions given by  $u(x, t_0) = u_0(x)$  and mixed Dirichlet and Neumann boundary conditions as for the elliptic problem (2.1.1). Using the simple backward Euler method in time leads to the following weak form for the HMM

$$\begin{aligned} \int_{\Omega} u_n^H v^H dx + \Delta t B(u_n^H, v^H) &= \int_{\Omega} u_{n-1}^H v^H dx \\ &+ \Delta t \left( \int_{\Omega} f v^H dx + \int_{\partial\Omega_N} g_N v^H dx - B(g_D, v^H) \right), \end{aligned} \quad (3.1.16)$$

$\forall v^H \in V_D^p(\Omega, \mathcal{T}_H)$  and  $t \in (t_0, T)$ , where  $u_n^H$  is an approximation of  $u(x, t_n)$ ,  $t_n = t_0 + n\Delta t$  and  $B(\cdot, \cdot)$  is defined in (2.4.6). This leads to the linear system

$$(M + \Delta t A) u_n^H = b + M u_{n-1}^H, \quad (3.1.17)$$

where  $A$  and  $b$  are as in (3.1.4) and obtained as described in Section 2 and the mass matrix  $M$  has entries given by  $M_{ij} = \int_{\Omega} \varphi_i^H \varphi_j^H dx$ . We described here the simplest time integration procedure. Of course more sophisticated solvers could be implemented. Notice that by using discontinuous Galerkin methods for the HMM [5], the mass matrix  $M$  becomes block diagonal and an explicit stabilized scheme as proposed in [1] could be applied for the time integration, avoiding the use of linear solver for the whole approximation procedure of parabolic multiscale problems. This may be an appealing strategy when the spatially discretized problem is of high dimension and contains complicated nonlinear terms. In this case, the implementation of iterative solvers requires some care (suitable preconditioners, etc.) while the above strategy is very easy to implement while being efficient.

### 3.1.4 Three dimensional problems

With only a few modifications we can extend the 2D code presented above to three dimensional problems. For the macro problem, it is convenient to use tetrahedral elements. In three dimension, the sampling domains are chosen to be cubes of size  $\delta$  and parallelepipeds meshes are suitable for the micro problems. At the macro level, the core structure using 3D shape functions is built upon standard procedure [22]. Macro assembly using micro solvers on sampling domains follows the procedure explained in Section 2 with obvious changes. Let us notice that some care is needed to implement the micro boundary conditions. For example, for periodic boundary conditions, a moment of reflection is needed to find all couples of boundary nodes on opposite sides while avoiding redundant couples. Figure 3.1.2(b) gives an example of such a node selection.

## 3.2 Numerical examples

In what follows, we present several numerical examples to illustrate the use, the versatility, the capabilities and the performance of our code.

In a first set of examples, we present an elliptic problem with non-uniformly periodic coefficients (Section 3.2.1), an elliptic problem with a random tensor (Section 3.2.2) and a time dependent multiscale (parabolic) problem (Section 3.2.3). We notice that problems with random tensors, similar to those chosen here, are widely used in the modeling of porous media flows. In a second set of examples, we investigate three dimensional simulations based on simplified real-world engineering problems: the steady state heat distribution in a heat sink (Section 3.2.4) and in a microprocessor (Section 3.2.5). Both the processor and the heat sink are assumed to be made of composite materials.

### 3.2.1 Problems with non uniformly periodic coefficients

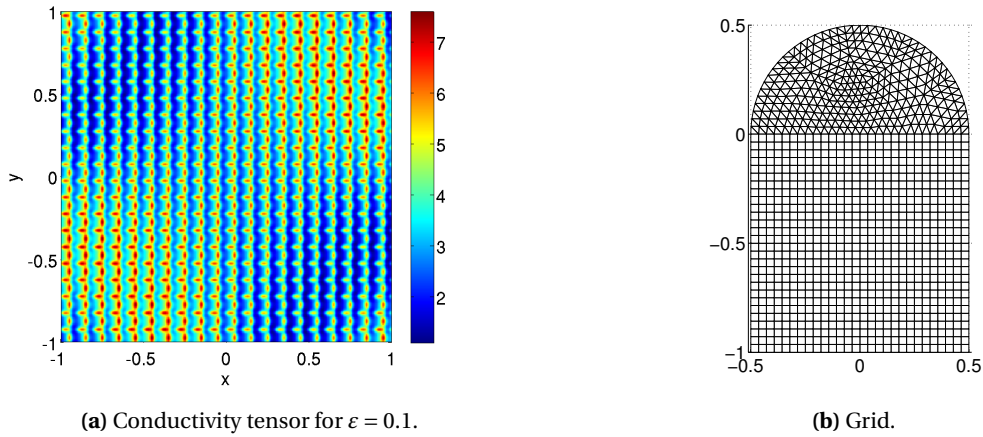
For the first example, we perform numerous numerical experiments to illustrate the convergence behavior of the FE-HMM in various norms, to show the impact of the size of the sampling domains on the numerical solution and to indicate how macro and micro meshes have to be refined in order to minimize the computational complexity with optimal convergence rates. For the other examples, we do not perform a rigorous numerical convergence study. Such studies have been reported elsewhere for the FE-HMM [14, 2, 4, 9, 6, 79]. We nevertheless occasionally complement our plots by giving energy norms of the exact, homogenized or FE-HMM solutions to illustrate some (convergence) aspects of the proposed algorithm.

We consider the two scale problem taken from [9]

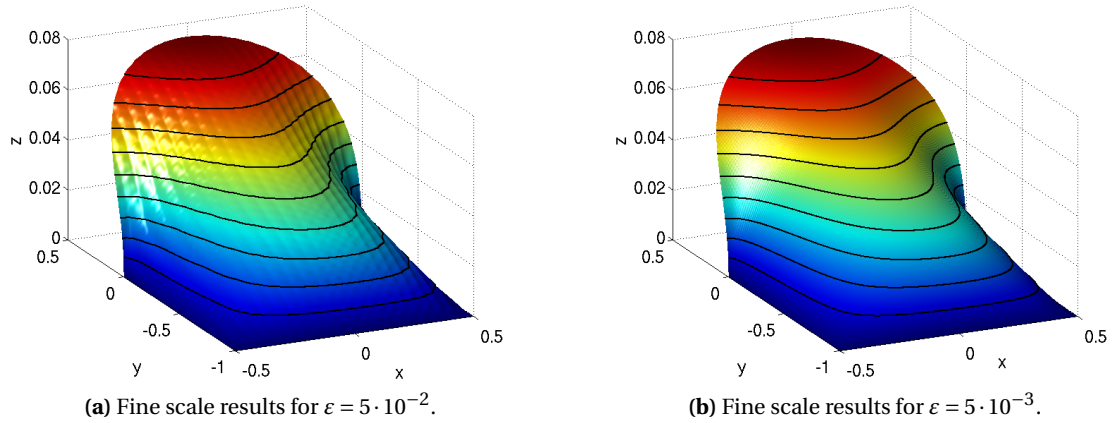
$$-\nabla \cdot \left( a \left( x, \frac{x}{\varepsilon} \right) \nabla u^\varepsilon \right) = f(x) \quad \text{in } \Omega, \quad (3.2.1)$$

$$u^\varepsilon = 0 \quad \text{on } \partial\Omega_D, \quad (3.2.2)$$

$$-n \cdot \left( a \left( x, \frac{x}{\varepsilon} \right) \nabla u^\varepsilon \right) = 0 \quad \text{on } \partial\Omega_N, \quad (3.2.3)$$



**Figure 3.2.1:** Snapshot of the conductivity tensor (left picture) and sketch of the computational mesh (right picture) for the elliptic problem of Section 3.2.1.



**Figure 3.2.2:** Fine scale solutions of the problem described in Section 3.2.1 computed on a mesh with  $10^6$  degrees of freedom for two different  $\varepsilon$ .

where the conductivity tensor  $a^\varepsilon$  is given by

$$a\left(x, \frac{x}{\varepsilon}\right) = \frac{1.5 \sin(2\pi x_1/\varepsilon)}{1.5 \sin(2\pi x_2/\varepsilon)} + \frac{1.5 \sin(2\pi x_2/\varepsilon)}{1.5 \cos(2\pi x_1/\varepsilon)} + \sin(4x_1 x_2) + 1, \quad (3.2.4)$$

with  $x = (x_1, x_2)$ . We choose  $f = 1$ . A snapshot of the conductivity tensor is shown in Figure 3.2.1a.

The domain  $\Omega$  consists of a semi circle and a rectangle, meshed with 1137 nodes using 576 triangles and 784 quadrilaterals, respectively (see Figure 3.2.1b). This problem does not

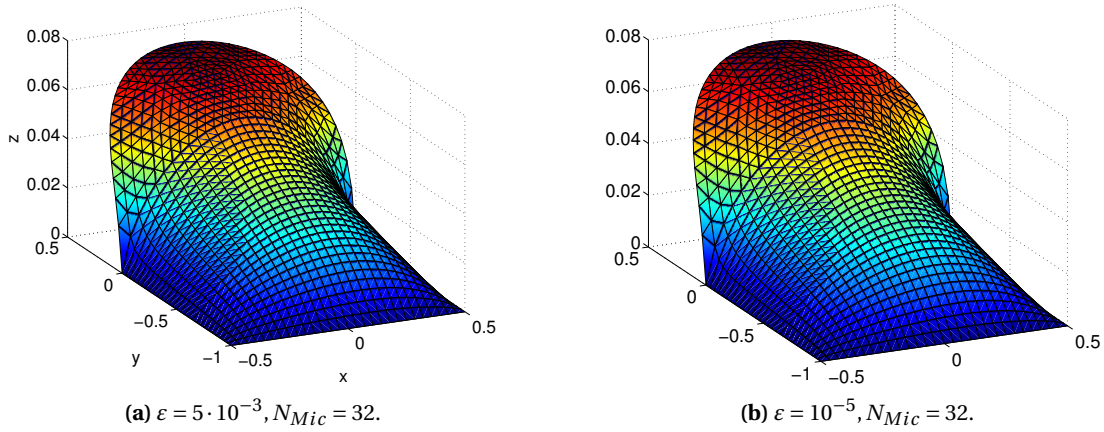


Figure 3.2.3: FE-HMM solutions of the elliptic problem described in Section 3.2.1 for two different  $\varepsilon$ . The FE-HMM captures the effective solution and is thus independent of  $\varepsilon$ .

|                         |                | $N_{mic} = 4$ | $N_{mic} = 8$ | $N_{mic} = 16$ | $N_{mic} = 32$ | finescale |
|-------------------------|----------------|---------------|---------------|----------------|----------------|-----------|
| $\varepsilon = 0.005$   | $\ u\ _A$      | 0.2105        | 0.2124        | 0.2128         | 0.2129         | 0.2146    |
|                         | $\ u\ _\infty$ | 0.0713249     | 0.072289      | 0.0725326      | 0.0725888      | 0.0736806 |
| $\varepsilon = 10^{-5}$ | $\ u\ _A$      | 0.2101        | 0.2124        | 0.2128         | 0.2129         | -         |
|                         | $\ u\ _\infty$ | 0.0710427     | 0.0722795     | 0.0725326      | 0.0725888      | -         |

Table 3.2.1: Energy and maximum norm for the solution of the problem described in Section 3.2.1.

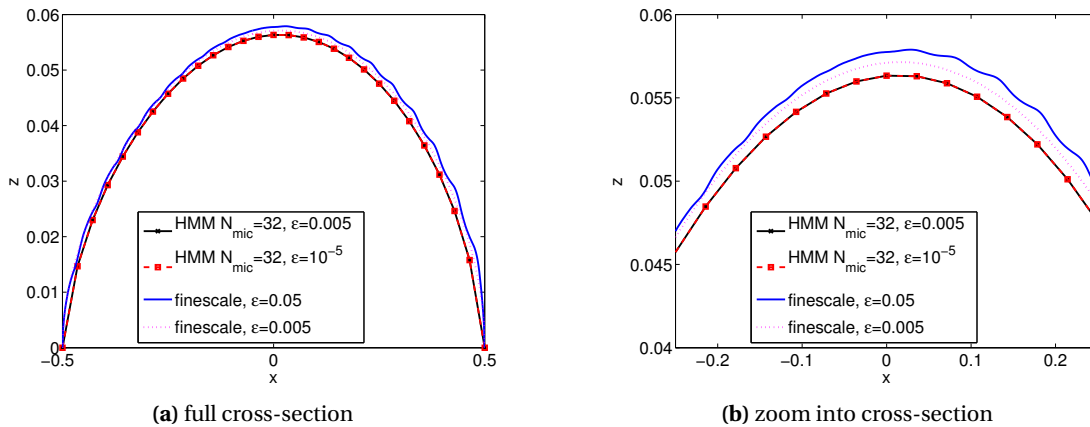
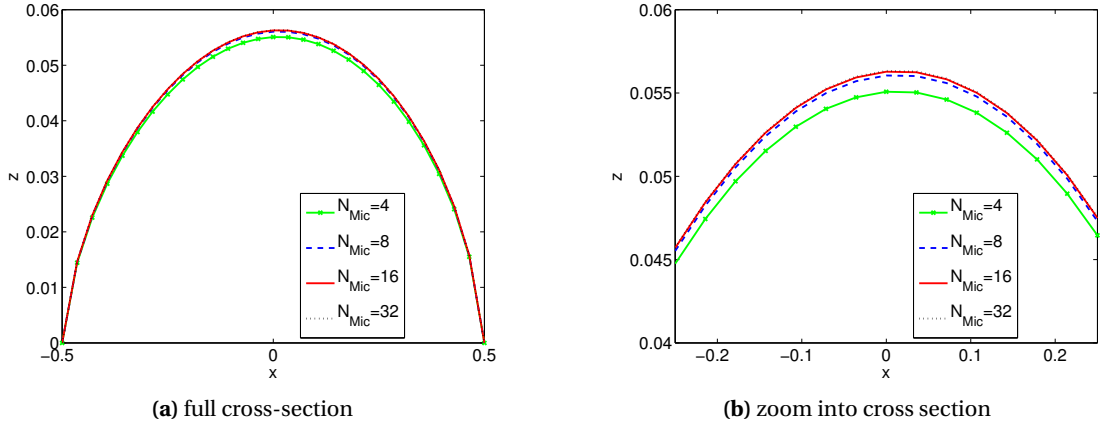


Figure 3.2.4: FE-HMM solution (with 1100 macro DOF) and fine scale solutions (with  $10^6$  DOF) of problem 3.2.1 for various  $\varepsilon$  (cross section of the solution of the problem described in Section 3.2.1 through  $y = 0$ ).



**Figure 3.2.5:** FE-HMM solutions of the problem described in Section 3.2.1 for  $\varepsilon = 0.005$  and  $N_{mic} = 4, 8, 16, 32$  (cross section of the solution of the problem described in Section 3.2.1 through  $y = 0$ ).

have an explicit analytical solution and we thus compute a finescale solution using a standard FEM with a fine mesh with about  $10^6$  degrees of freedom (DOF) in order to resolve the micro scale. Neumann boundary conditions (3.2.3) are given on the boundary of the semi circle  $\Omega_N$ , whereas the remaining part has Dirichlet boundary conditions (3.2.2). While the problem itself is not periodic, the fast scale is, and we thus choose  $\delta = \varepsilon$  for the size of the sampling domains and periodic micro boundary conditions. Figure 3.2.3 shows the solution generated by the FE-HMM.

In Figures 3.2.2 and 3.2.4 we compare the FE-HMM solution with the fine scale solutions obtained by choosing  $\varepsilon = 0.05$  and  $\varepsilon = 0.005$ . Fine scale solutions with smaller  $\varepsilon$  quickly get impossible to compute due to the increasing complexity. The plots of Figure 3.2.2 show the behavior of the fine scale solution as  $\varepsilon$  decreases. For  $\varepsilon = 0.005$  the oscillations are barely visible (remember that in the limit  $\varepsilon \rightarrow 0$  one gets the homogenized solution  $u^0$  of (2.4.2) as discussed in Section 2.1.2).

The FE-HMM captures the effective solution and is thus independent of  $\varepsilon$ . This fact is consistent with the numerical results provided in Table 3.2.1, where we compare the energy norm  $\|u\|_A := \sqrt{\int_{\Omega} a \nabla u \cdot \nabla u dx}$  and maximum norm  $\|u\|_{\infty} := \sup_{x \in \Omega} |u(x)|$  for different values of  $\varepsilon$  and  $N_{mic}$ . Here,  $N_{mic}$  refers to the number of discretization points in each direction of the  $d$ -dimensional sampling domains of the micro solver (2.4.7) of the FE-HMM, i.e. we have a total of  $N_{mic}^d$  DOF in each sampling domain. In Figure 3.2.5 we show the influence of the micro discretization in the macro solution of the FE-HMM by varying  $N_{mic}$ . The results presented in Section 2.4.1 allow to choose the appropriate micro mesh size according to the selected macro discretization. This is further discussed in Section 3.2.1.1.

**Sampling domain size.** We next report numerical results in the situation when  $\delta$ , the size of the sampling domain, is not an integer multiple of  $\varepsilon$ , i.e. when  $\delta$  does not comprise an integer number of  $\varepsilon$  micro oscillations. This situation arises even for periodic problems, if one does not know the exact size of the period. In Table 3.2.2, we study the influence of such “modeling

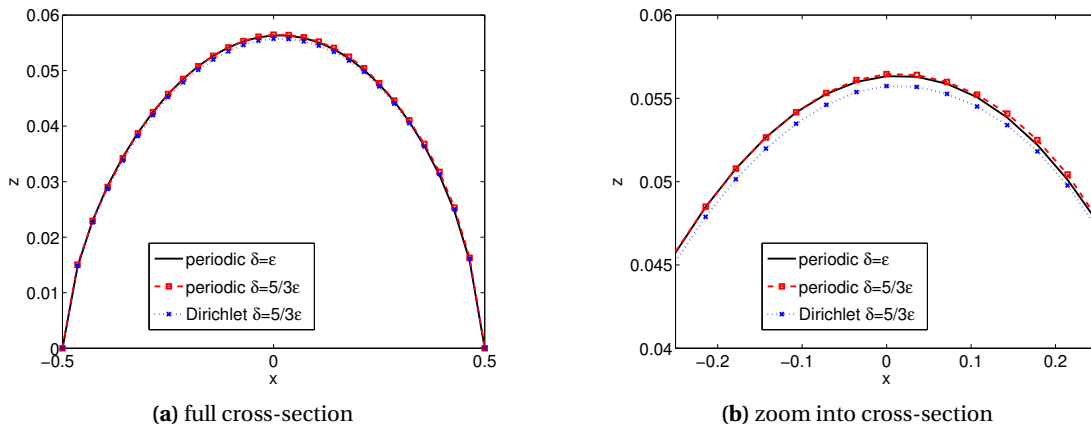
error” for the HMM. We choose  $\delta = \frac{5}{3}\varepsilon$ , and compare the HMM solution obtained with either periodic or Dirichlet boundary conditions for the micro problem. For each type of boundary conditions we also refine the discretization of the micro solver by increasing  $N_{mic}$ .

As can be seen, we get better solutions with periodic boundary conditions. The better behavior of periodic boundary conditions even for such non-matching sampling domains is well-known but yet not fully understood [118]. Cross-section plots for a fixed  $N_{mic}$  are presented in Figure 3.2.6. Further numerical results for the FE-HMM can be found in [2, 4, 79].

*Remark 14.* We emphasize that while it is convenient to compare the HMM solution  $u^H$  with a resolved fine scale solution, the former does usually not converges to the latter in the  $H^1$  or in the energy norm as  $\varepsilon \rightarrow 0$  (remember that  $u^H$  converges to  $u^0$  as macro and micro meshes tend to zero). In the  $L^2$  norm, the following convergence rate can be obtained  $\|u^\varepsilon - u^0\|_{L^2(\Omega)} \leq C\varepsilon$ , and by the triangle inequality and by using results of Section 2.4.1, convergence rate for  $u^H$  (to the fine scale solution in the  $L^2$  norm) can also be obtained (see [2] for details).

|           |                | $N_{mic} = 4$ | $N_{mic} = 8$ | $N_{mic} = 16$ | $N_{mic} = 32$ | finescale |
|-----------|----------------|---------------|---------------|----------------|----------------|-----------|
| periodic  | $\ u\ _A$      | 0.2085        | 0.2119        | 0.2130         | 0.2133         | 0.2146    |
|           | $\ u\ _\infty$ | 0.0702        | 0.0720        | 0.0726         | 0.0728         | 0.0737    |
| dirichlet | $\ u\ _A$      | 0.2080        | 0.2107        | 0.2117         | 0.2119         | 0.2146    |
|           | $\ u\ _\infty$ | 0.0699        | 0.0714        | 0.0719         | 0.0720         | 0.0737    |

**Table 3.2.2:** Energy and maximum norm for the FE-HMM and fine-scale solution of the elliptic problem described in Section 3.2.1 with a non-integer number of periods within each sampling domain ( $\delta = \frac{5}{3}\varepsilon$ ,  $\varepsilon = 5 \cdot 10^{-3}$ ).



**Figure 3.2.6:** FE-HMM solutions of the problem described in Section 3.2.1 with a non-integer number of periods within the sampling domain with  $\delta = \frac{5}{3}\varepsilon$ ,  $\varepsilon = 0.05$ ,  $N_{mic} = 32$  (cross section through  $y = 0$ ).



### 3.2.1.1 Performance of the method.

Finally, we want to investigate the dependence of  $L^2$  and  $H^1$  errors on the choice of the macro and micro meshes and give an indication of CPU-times for the FE-HMM.

In order to easily compare  $L^2$  and  $H^1$  errors for different mesh sizes, we consider the following quasi-1d-problem on a square domain for which the exact solution and the homogenized tensor can be computed analytically (see [2])

$$-\nabla \cdot \left( a \left( \frac{x}{\varepsilon} \right) \nabla u^\varepsilon \right) = f(x) \quad \text{in } \Omega = (0, 1)^2, \quad (3.2.5)$$

$$u^\varepsilon|_{\Gamma_D} = 0 \quad \text{on } \Gamma_D := \{x_1 = 0\} \cup \{x_1 = 1\}, \quad (3.2.6)$$

$$n \cdot \left( a \left( \frac{x}{\varepsilon} \right) \nabla u^\varepsilon \right) |_{\Gamma_N} = 0 \quad \text{on } \Gamma_N := \partial\Omega \setminus \Gamma_D. \quad (3.2.7)$$

We choose  $f(x) \equiv 1$  and set  $a \left( \frac{x}{\varepsilon} \right) = \cos \left( 2\pi \frac{x_1}{\varepsilon} \right) + 2$ . For the FE-HMM, we choose quadrilaterals for the macro- and micro-discretization.

**Optimal refinement strategy for the macro and micro meshes.** In what follows,  $N_{mac}$  refers to the number of discretization points in each direction of the domain  $\Omega = (0, 1)^2$ , while we recall that  $N_{mic}$  refers to the number of discretization points in each direction of the  $d$ -dimensional sampling domains (here  $d = 2$ ). Thus, the macro mesh  $H$  is given by  $H = 1/N_{mac}$  and the micro mesh by  $h = \delta/N_{mic}$  (in the following computations we choose  $\delta = \varepsilon$ ). As the sampling domain is of size  $\delta$ , we will also consider the scaled micro mesh given by  $h_s = h/\delta = 1/N_{mic}$ . The macro and micro DOF are then given by  $\mathcal{O}(M_{mac})$  and  $\mathcal{O}(M_{mic})$ , respectively, where  $M_{mac} = N_{mac}^d$  and  $M_{mic} = N_{mic}^d$ . Following the *a priori* estimates stated in Section 2.4.1 we see that

$$M_{mic} \propto M_{mac} \quad (L^2 \text{ norm}), \quad M_{mic} \propto \sqrt{M_{mac}} \quad (H^1 \text{ norm}),$$

i.e.,  $h_s \propto H$  or  $N_{mic} \propto N_{mac}$  (for the  $L^2$  norm) and  $h_s \propto \sqrt{H}$  or  $N_{mic} \propto \sqrt{N_{mac}}$  (for the  $H^1$  norm) are the best refinement strategies for optimal convergence rates with minimal computational cost. We thus obtain a complexity of  $\mathcal{O}(M_{mac} \cdot M_{mic}) = \mathcal{O}(M_{mac}^{3/2})$  floating point operations for a linear (macro) convergence rate in the  $H^1$  norm and  $\mathcal{O}(M_{mac} \cdot M_{mic}) = \mathcal{O}(M_{mac}^2)$  floating point operations for a quadratic convergence rate in the  $L^2$  norm. Here we assume that the cost (floating point operations) of the method is proportional to the total DOF (which is the case for example when using multigrid linear solver). For the above problem, we select a sequence of macro meshes  $H = 1/16, 1/32, 1/64, 1/128$  and adapt the micro mesh  $h_s$  according to the above criteria. This gives (selecting  $h_s = H$  for the  $L^2$  error and  $h_s = \sqrt{H}$  for the  $H^1$  error) the following macro and micro meshes written here in term of  $N_{mic}$  and  $N_{mac}$  ( $H = 1/N_{mac}$ ,  $h_s = 1/N_{mic}$ ):

|   |    |    |    |     |
|---|----|----|----|-----|
| $N_{mic}$                                     | 16 | 32 | 64 | 128 |
| $N_{mic}$ ( $L^2$ -micro refinement strategy) | 16 | 32 | 64 | 128 |
| $N_{mic}$ ( $H^1$ -micro refinement strategy) | 4  | 6  | 8  | 11  |

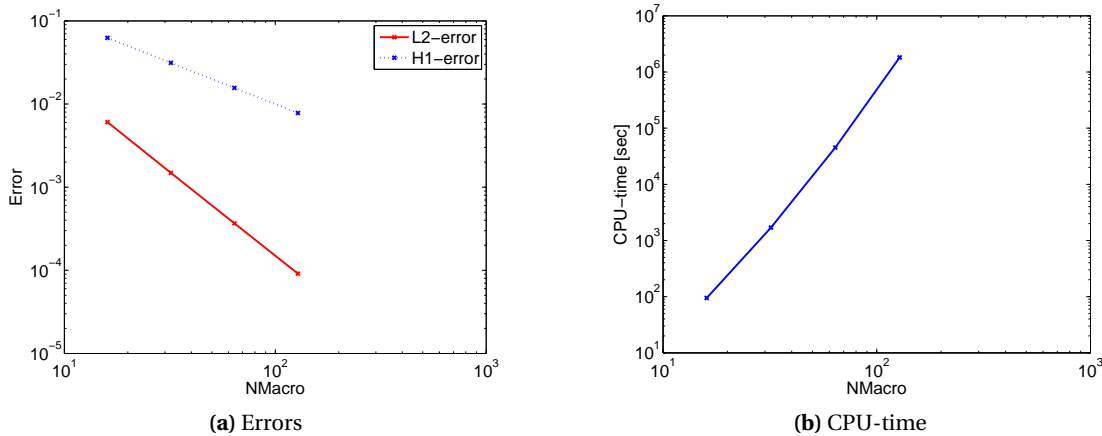
The results presented in Figures 3.2.7 and 3.2.8 show the errors and CPU-times corresponding to the aforementioned refinement-strategies (CPU-times measured on an AMD Opteron 2214 HE).

*Remark 15.* Because of the constant homogenized tensor we could solve the micro problem just once. However, to assess realistic computing times for two-scale problems, we provide the CPU-times that would be needed to solve the full problem.

As predicted by the *a-priori* estimates, the chosen refinement strategies achieve quadratic and linear convergence rates for the error in the  $L^2$  and  $H^1$  norms, respectively. We list in Table 3.2.3 the detailed errors.

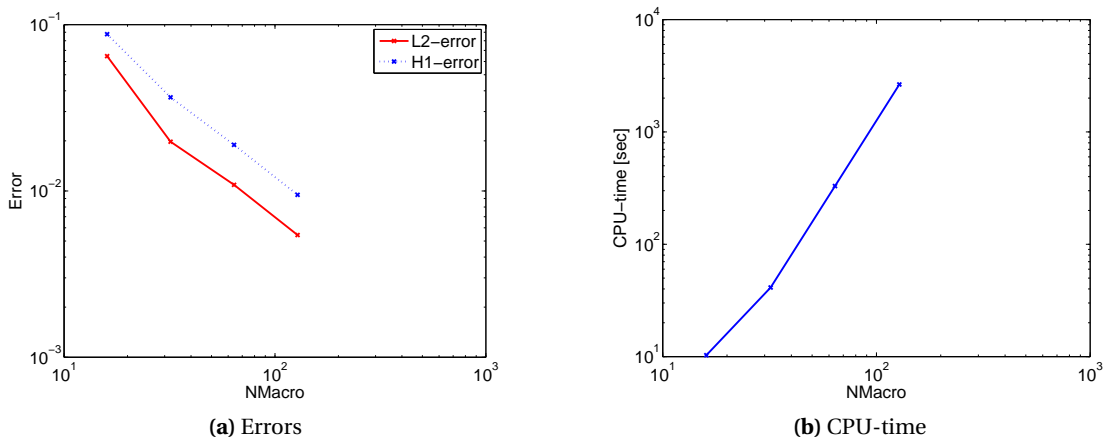
| $H$                  |                               | 1/16     | 1/32     | 1/64     | 1/128    |
|----------------------|-------------------------------|----------|----------|----------|----------|
| $L^2$ -approximation | $\ u^0 - u^H\ _{L^2(\Omega)}$ | 6.06e-03 | 1.48e-03 | 3.67e-04 | 9.13e-05 |
|                      | $\ u^0 - u^H\ _{H^1(\Omega)}$ | 6.25e-02 | 3.13e-02 | 1.56e-02 | 7.80e-03 |
| $H^1$ -approximation | $\ u^0 - u^H\ _{L^2(\Omega)}$ | 6.46e-02 | 1.98e-02 | 1.09e-02 | 5.43e-03 |
|                      | $\ u^0 - u^H\ _{H^1(\Omega)}$ | 8.76e-02 | 3.65e-02 | 1.89e-02 | 9.48e-03 |

**Table 3.2.3:** Comparison of the  $L^2$  and  $H^1$  errors for the test-case described in Section 3.2.1.1 using different refinement strategies.



**Figure 3.2.7:** Dependence of errors and CPU-time on the macro DOF for the test-case problem described in Section 3.2.1.1 ( $L^2$ -micro refinement strategy).

We see in Table 3.2.3 that an appropriate refinement strategy is crucial to achieve the optimal quadratic convergence rate in the  $L^2$  norm. We also see (in accordance with the



**Figure 3.2.8:** Dependence of errors and CPU-time on the macro DOF for the test-case problem described in Section 3.2.1.1 ( $H^1$ -micro refinement strategy).

theoretical estimates) that we can use a much larger micro mesh for  $H^1$  approximations than for  $L^2$  approximations. We emphasize that a correct refinement strategy is crucial for efficient computations. Finally, we observe as expected that the convergence rates are independent of  $\varepsilon$ .

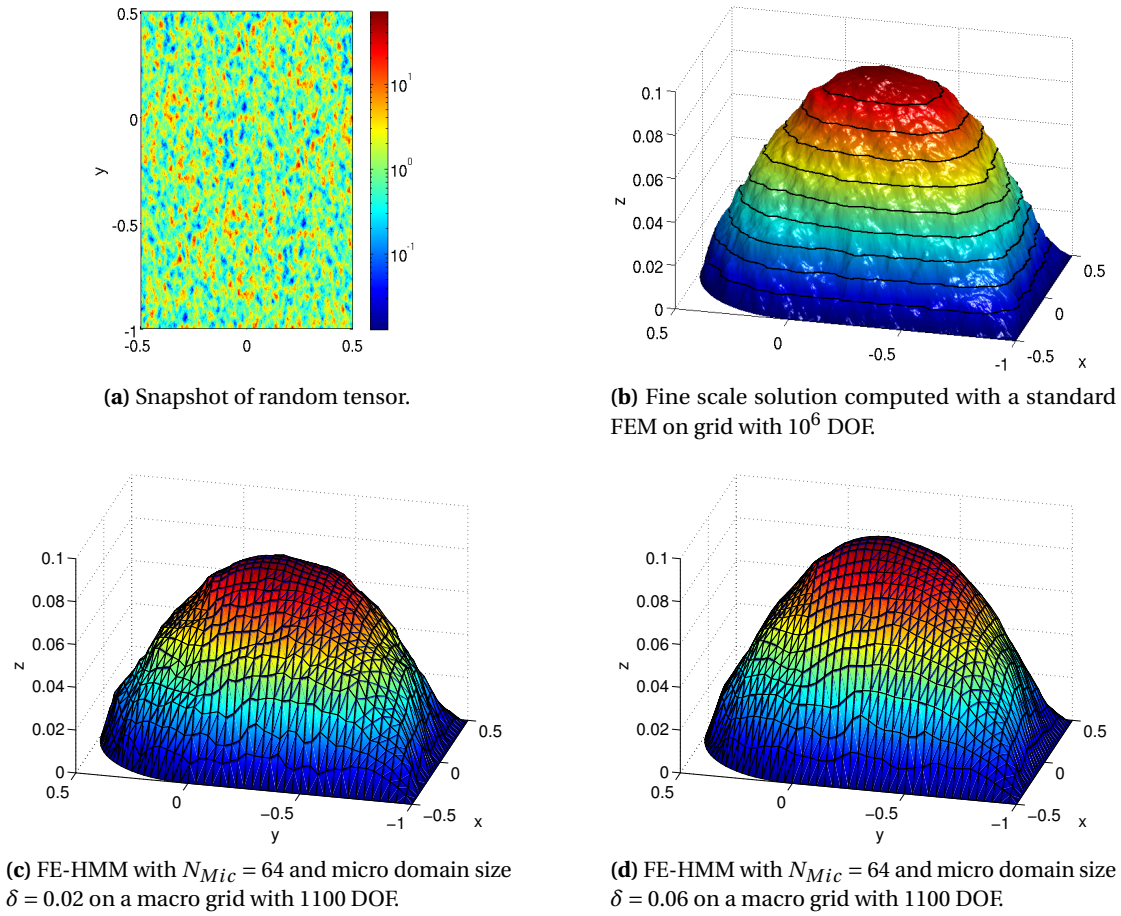
*Remark 16.* A comparison with a fine-scale solution is not useful in the present context of performance-evaluation. Indeed, as the performance of the FE-HMM is independent of the fine scale  $\varepsilon$ , this parameter can be tuned arbitrarily for the fine scale problem, leading to arbitrary (high) CPU-time for the corresponding fine scale solution.

**Parallel implementation of the FE-HMM.** We remark that an implementation on a parallel computer is easily possible and leads – due to the independence of the different micro problems – to a near-optimal speedup. A parallel version of the FE-HMM (on eight nodes) has been used for the 3D-problems in Sections 3.2.4 and 3.2.5.

### 3.2.2 Problems with random tensor

We again consider problem (3.2.1), but this time with a random tensor  $a^\varepsilon(x)$ . The domain  $\Omega$  is the same as in the previous example of Section 3.2.1 except that zero Dirichlet boundary conditions are chosen on the entire boundary. Again, triangles are used in the circular part of the domain while quadrilaterals are used in the remaining computational domain. We notice that random models for the fine scale are often used for problems involving the pressure equation in porous media flow [115].

This example shows how to address problems where the tensor is not given by an explicit function, but only as a random field. While the stochastic field for the considered problem is computer-generated, our code could be used with a real-life tensor obtained from imaging techniques (for example through scanning or microscopy). The realization of the log-normal stochastic field with mean zero and variance  $\sigma = 1$  is generated by the moving ellipse average method [115, Section 4.1]. We set the correlation lengths of the stochastic field to be  $\varepsilon_1 = 0.01$



**Figure 3.2.9:** FE-HMM and fine scale solution of the random problem described in Section 3.2.2. Snapshot of the tensor and comparison of the fine-scale solution with FE-HMM solutions with differently sized micro-domains.

and  $\varepsilon_2 = 0.02$ . A snapshot of this tensor is shown in Figure 3.2.9a.

We compute a reference solution on a fine grid of about  $10^6$  degrees of freedom (DOF), see Figure 3.2.9b, and compare the solution with the FE-HMM on the macro coarse grid with 1100 DOF. For the FE-HMM we present results for various sizes of the sampling domain. We first choose  $\delta = 0.02$  and then  $\delta = 0.06$ . We clearly see in Figures 3.2.9c and 3.2.9d that the profile of the solution is closer to the reference solution as the sampling domain contains more correlation length of the random field. This observation can also be seen in Table 3.2.4 when comparing the energy norm of the various solutions obtained with the FE-HMM to the energy norm of the reference solution.

|                 | $N_{mic} = 4$ | $N_{mic} = 8$ | $N_{mic} = 16$ | $N_{mic} = 32$ | $N_{mic} = 64$ | finescale |
|-----------------|---------------|---------------|----------------|----------------|----------------|-----------|
| $\delta = 0.02$ | 0.2352        | 0.2415        | 0.2439         | 0.2449         | 0.2454         | 0.2583    |
| $\delta = 0.06$ | 0.2313        | 0.2454        | 0.2520         | 0.2551         | 0.2567         | 0.2583    |

**Table 3.2.4:** Energy norms of the FE-HMM and fine-scale solutions of the problem described in Section 3.2.2 with a random tensor. FE-HMM results are given for various micro sampling-domain sizes  $\delta \times \delta$  and DOF  $N_{mic}^2$ .

|                         |                | $N_{mic} = 4$ | $N_{mic} = 8$ | $N_{mic} = 16$ | $N_{mic} = 32$ | finescale |
|-------------------------|----------------|---------------|---------------|----------------|----------------|-----------|
| $\varepsilon = 0.005$   | $\ u\ _A$      | 0.8354        | 0.8491        | 0.8525         | 0.8533         | 0.8697    |
|                         | $\ u\ _\infty$ | 1.0839        | 1.1060        | 1.1117         | 1.1130         | 1.1379    |
| $\varepsilon = 10^{-5}$ | $\ u\ _A$      | 0.8313        | 0.8490        | 0.8525         | 0.8533         | -         |
|                         | $\ u\ _\infty$ | 1.0773        | 1.1058        | 1.1117         | 1.1130         | -         |

**Table 3.2.5:** Energy and maximum errors for the FE-HMM and the fine-scale solutions of the parabolic problem described in Section 3.2.3 at  $t = 1$ . The computation of the fine-scale solution with  $\varepsilon = 10^{-5}$  would involve around  $10^{14}$  DOF and has therefore not been performed.

### 3.2.3 Parabolic problems

As mentioned in Section 3.4, our code can handle parabolic problems without difficulties. As an example, we consider

$$\frac{\partial u^\varepsilon}{\partial t} = \nabla \cdot (a^\varepsilon(x)u^\varepsilon) + f \text{ in } \Omega \times [0, T]. \quad (3.2.8)$$

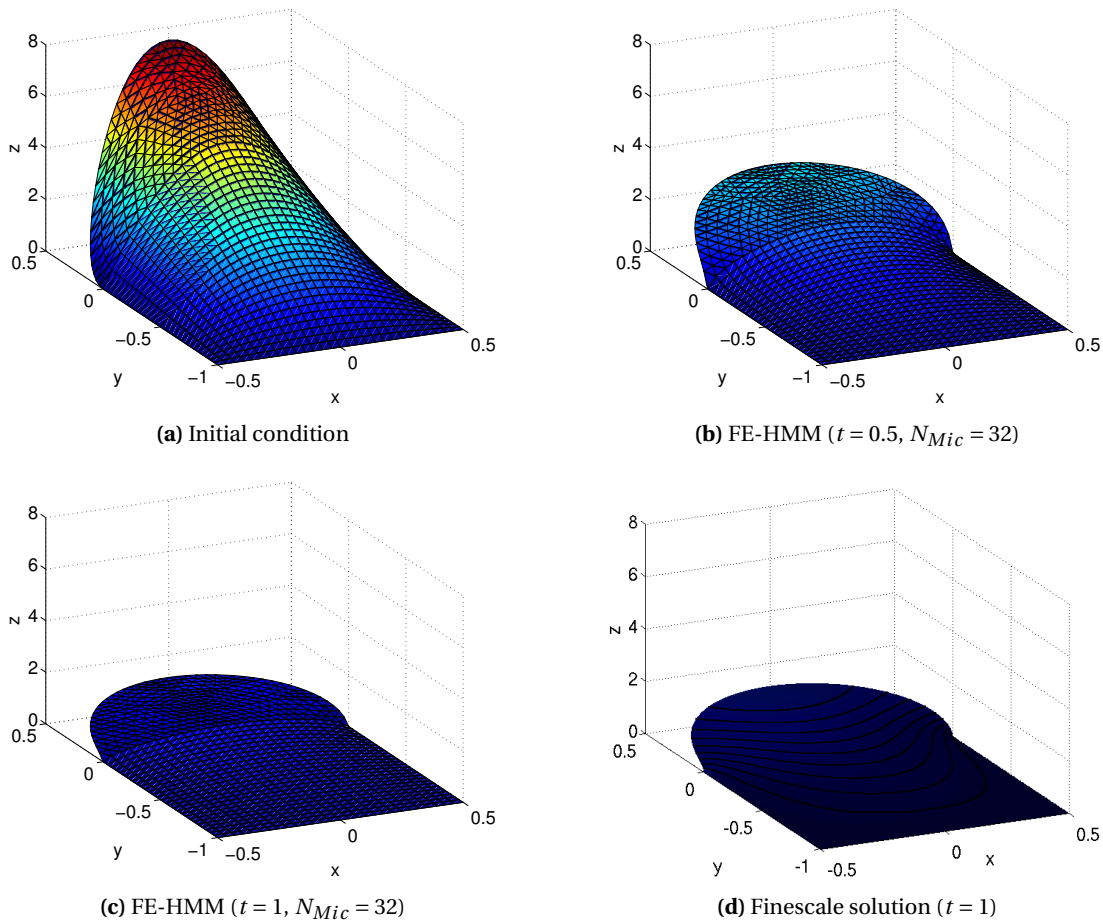
For simplicity, we consider an implicit Euler scheme for time integration. We choose  $T = 1$  and a time step of  $\Delta t = 0.1$ .

We keep the same domain and boundary conditions as in problem 3.2.1, set  $f = 0$  and consider the initial condition  $u_0 = -10(x - 0.5) \cdot (x + 0.5) \cdot (y + 1)$ . The tensor used is a scaled version of the tensor (3.2.4), given by

$$a\left(x, \frac{x}{\varepsilon}\right) = 0.1 \cdot \left( \frac{1.5 \sin(2\pi x_1/\varepsilon)}{1.5 \sin(2\pi x_2/\varepsilon)} + \frac{1.5 \sin(2\pi x_2/\varepsilon)}{1.5 \cos(2\pi x_1/\varepsilon)} + \sin(4x_1 x_2) + 1 \right).$$

Since we consider a time-independent tensor  $a^\varepsilon$  here, we only have to solve the micro problems once and can use the same stiffness matrix whenever we solve the linear equation arising from the implicit Euler scheme. The code corresponding to this problem is given in `fe_hmm2d_para.m`. More general problems with tensors of the form  $a(x, \frac{x}{\varepsilon}, t)$  can be computed in a similar manner, but in this case the stiffness matrix has to be updated at each time step.

Results for this parabolic problem solved with the FE-HMM code are sketched in Figure 3.2.10. The results at  $t = 1$  are extremely close to each other and are indistinguishable to the eye. A comparison of the energy and infinity norms of various numerical solutions is given in Table 3.2.5. Again, we clearly see that the FE-HMM is independent of  $\varepsilon$ .



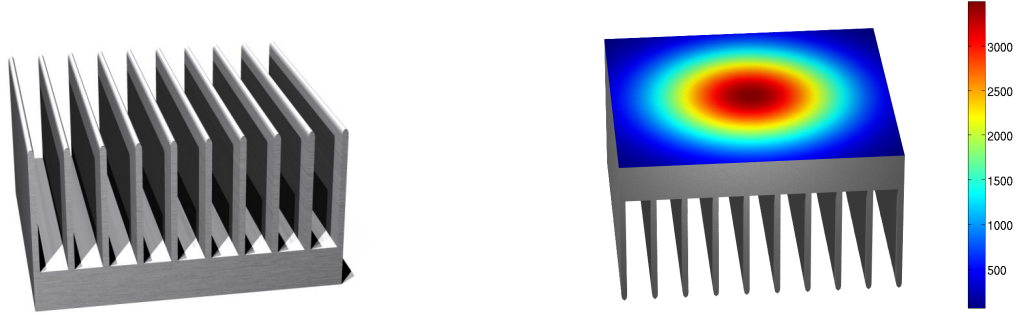
**Figure 3.2.10:** FE-HMM (1100 macro DOF) and fine scale ( $10^6$  DOF) solution of the parabolic problem described in Section 3.2.3, obtained with the described MATLAB code ( $\epsilon = 0.005$ ).

### 3.2.4 Three dimensional problem: steady state heat distribution in a heat sink

We now present a 3D multiscale problem, the steady state heat conduction of a heat sink made of a layered material. Crucial for the efficiency of today's microprocessors is a sophisticated cooling process which is usually achieved with the help of a heat sink, where advanced composite materials are used in order to obtain the most efficient cooling procedure. These heat sinks are typically mounted on top of a microprocessor to dissipate heat away from it, often with the help of a fan built on top of it. However, since we want to avoid discussing modeling issues, we neglect the effects originating from such fans and more refined models.

**Equations.** Heat is transferred away in three ways: conduction, convection and radiation, the latter of which we ignore. Heat conduction within the heat sink and its different components can be described in terms of Fourier's law

$$-\nabla \cdot (a \nabla u) = f, \tag{3.2.9}$$



(a) Heat sink, which can be mounted on a microprocessor to dissipate heat.

(b) Boundary conditions. The side shown in color is connected to the processor package and represent incoming heat flux. Neumann boundary conditions (3.2.12) are assigned to this face. The color bar refers to the incoming heat flux from the processor (measured in  $\frac{W}{m^2}$ ), reflecting the fact the major part of the incoming flux is located close to the processor core. The remaining (grey) part is modeled with Robin boundary conditions (3.2.11).

**Figure 3.2.11:** Model of the heat sink simulated in Section 3.2.4.

where  $a$  is the thermal conductivity tensor,  $u$  is the temperature of the body and  $f$  is a heat source. Convective heat transfer with the surrounding air is expressed by the Robin boundary condition

$$n \cdot (a \nabla u) = q_0 + \alpha (u_{amb} - u),$$

where  $\alpha$  is the heat transfer coefficient,  $u_{amb}$  is the ambient temperature and  $q_0$  represents the heat flux entering the domain.

Using the above heat transfer models leads to the following problem describing the steady-state temperature distribution of the heat sink

$$-\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) = f \quad \text{in } \Omega, \quad (3.2.10)$$

$$n \cdot (a^\varepsilon \nabla u^\varepsilon) + \alpha u^\varepsilon = g_R \quad \text{on } \partial\Omega_R, \quad (3.2.11)$$

$$n \cdot (a^\varepsilon \nabla u^\varepsilon) = g_N \quad \text{on } \partial\Omega_N, \quad (3.2.12)$$

where  $\Omega$  is the whole domain and  $\partial\Omega_R$  and  $\partial\Omega_N$  are the surfaces of the heat sink with Robin and Neumann boundary conditions, respectively. The right hand side of the Robin boundary conditions is given by  $g_R = q_0 + \alpha u_{amb}$ . Notice that we add (as usual) the superscript  $\varepsilon$  to emphasize on the multiscale (composite) nature of the material. The Robin boundary conditions (3.2.11) involve extra terms  $\int_{\partial\Omega_R} g_R v ds$  and  $-\int_{\partial\Omega_R} \alpha u v ds$  to be incorporated into the weak form (2.1.3). For the FE-HMM, this results in an extra contribution  $R$  to the stiffness matrix  $A$ ,

given by

$$R_{ij} = \int_{\partial\Omega_R} \alpha \varphi_i^H \varphi_j^H ds, \quad (3.2.13)$$

and an extra contribution to the load vector  $l$  given by

$$r_j = \int_{\partial\Omega_R} g_R \varphi_j^H ds. \quad (3.2.14)$$

While we consider here two-scale tensors with periodicity in the micro scale, i.e. non uniformly periodic tensors, other tensors for realistic materials obtained via imaging techniques and given as data points could be easily taken into account.

### 3.2.4.1 Setup

The side of the heat sink mounted on top of the processor has an incoming heat flux that we model by Neumann boundary conditions

$$n \cdot (a \nabla u) = g_N \text{ on } \partial\Omega_N,$$

see Figure 3.2.11b. The incoming heat flux entering  $\partial\Omega_N$  is given by  $g_N(x) = 3493 e^{-1000(x^2+y^2)} \frac{W}{m^2}$ , where the numerical values are chosen such that we have a total incoming power of  $P = \int_{\partial\Omega_N} g_N dx = 10W$ .

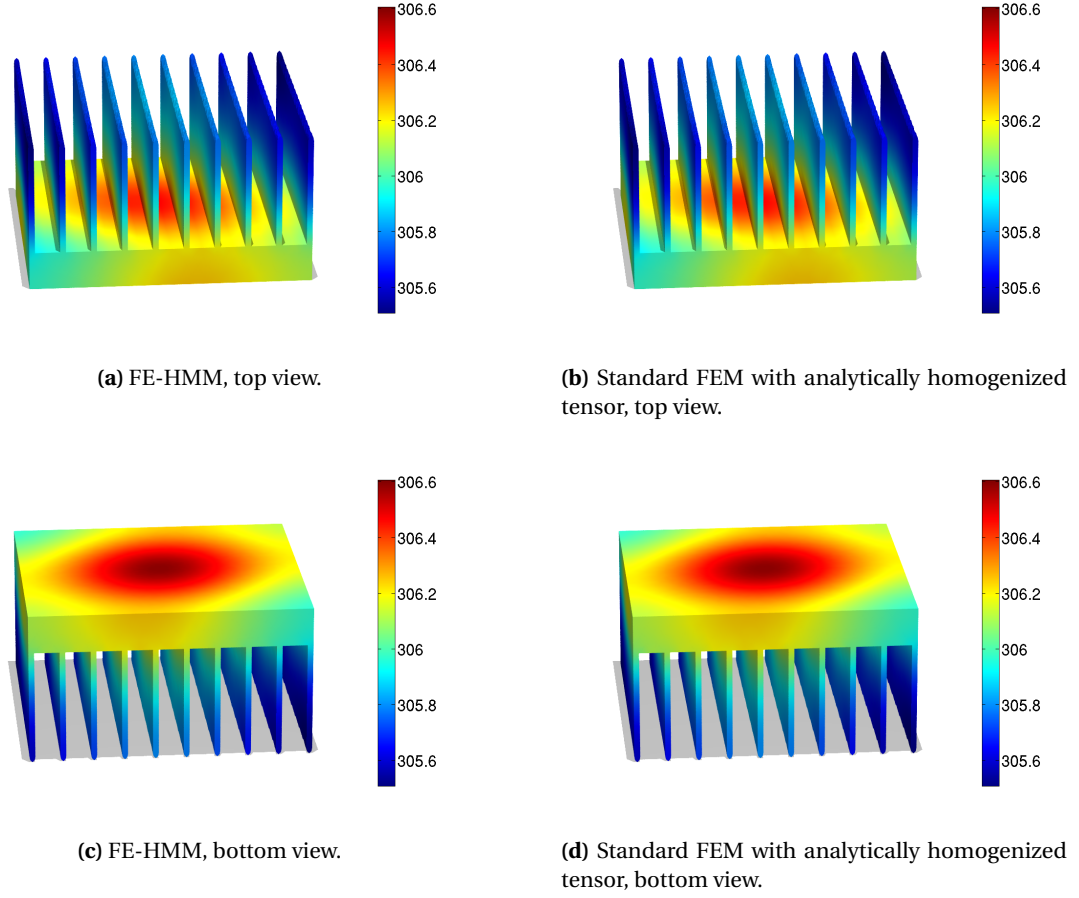
In our computations, we will assume that the heat sink is made of a layered material. The advantage of choosing such a material is that an analytic formula for the homogenized material is available. This allows to compare the solution of the FE-HMM with an exact solution. Needless to say that for most of materials used in cooling procedures (e.g. [39, 121]), analytical formulas are not available and the use of computational procedures as described in this thesis are unavoidable. We note that recent studies suggest the use of carbon nanotubes to act as a heat sink [35].

*Remark 17.* We emphasize that for the chosen tensor, there is no advantage in using the FE-HMM since the homogenized tensor is constant and can be analytically computed. Of course for such problems, we need only to call the micro solver once when using the FE-HMM, and no knowledge about the correct averaging procedure is needed (see [6, Sect. 3.3.2]). The purpose of using such a simple tensor as considered here is that we can make precise comparisons with various averaging procedures. A more general tensor (for which no explicit homogenized tensor is available) is used in Section 3.2.5, where we study the heat distribution in a micro-processor.

We set the multiscale tensor to be

$$a^\varepsilon(x) = \begin{pmatrix} a_{11}^\varepsilon(x) & 35 & 0 \\ 35 & a_{22}^\varepsilon(x) & 0 \\ 0 & 0 & 200 \end{pmatrix}, \quad (3.2.15)$$





**Figure 3.2.12:** Steady-state heat distribution in a heat sink (problem described in Section 3.2.4). Comparison of the results obtained with the FE-HMM and a standard FEM solved with the analytically homogenized tensor (3.2.16). The macro mesh used in both computations has 17,000 DOF. For the FE-HMM, we choose  $\varepsilon = 10^{-6}$  and  $N_{mic} = 16$  (as the FE-HMM is independent of  $\varepsilon$  (in the periodic case) any other value of  $\varepsilon$  give the same results). The color bar represents the temperature in K.

where

$$\begin{aligned} a_{11}^\varepsilon(x) &= [500 / (5 + 3.5 \cdot \sin(2\pi x_1 / \varepsilon))] \cdot e^{10 \cdot x_3} \\ a_{22}^\varepsilon(x) &= [500 / (5 + 3.5 \cdot \cos(2\pi x_1 / \varepsilon))] \cdot e^{10 \cdot x_3}. \end{aligned}$$

For such tensors it is well-known that explicit analytic formulas are available for the homogenized tensor (see for example [45, Chap. 5.4]). For the tensor (3.2.15) we obtain

$$a^0(x) \approx \begin{pmatrix} 100.0 \cdot e^{10 \cdot x_3} & 35.0 & 0.0 \\ 35.0 & 140.0 \cdot e^{10 \cdot x_3} & 0.0 \\ 0.0 & 0.0 & 200.0 \end{pmatrix}, \quad (3.2.16)$$

for the homogenized tensor. In our numerical experiments, the heat sink has dimensions  $87.5 \times 91.9 \times 50 \text{ mm}^3$ . The macro mesh was generated by CUBIT [102] and consists of 70,000 tetrahedra with 17,000 grid points and the average tetrahedra volume is  $2.3044 \text{ mm}^3$ .

**Arithmetic and harmonic mean.** We also compare the results of the FE-HMM with numerical results obtained from tensor averaged using more naive averaging procedures as the arithmetic or harmonic means. To compute these averaged tensors, we keep the macro variables fixed and average the tensor element-wise over the micro scale only.

$$a_{ij}^{\text{arithmetic}}(x_k) = \int_Y a_{ij}(x_k, y) dy, \quad (3.2.17)$$

$$a_{ij}^{\text{harmonic}}(x_k) = \left( \int_Y (a_{ij}(x_k, y))^{-1} dy \right)^{-1}, \quad (3.2.18)$$

where  $Y$  is the 3D unit cube. This leads to the following tensors

$$a^{\text{arithmetic}}(x) = \begin{pmatrix} 140.0 \cdot e^{10x_3} & 35.0 & 0 \\ 35.0 & 140.0 \cdot e^{10x_3} & 0 \\ 0.0 & 0.0 & 200.0 \end{pmatrix}, \quad (3.2.19)$$

$$a^{\text{harmonic}}(x) = \begin{pmatrix} 100.0 \cdot e^{10x_3} & 35.0 & 0.0 \\ 35.0 & 100.0 \cdot e^{10x_3} & 0.0 \\ 0.0 & 0.0 & 200.0 \end{pmatrix}. \quad (3.2.20)$$

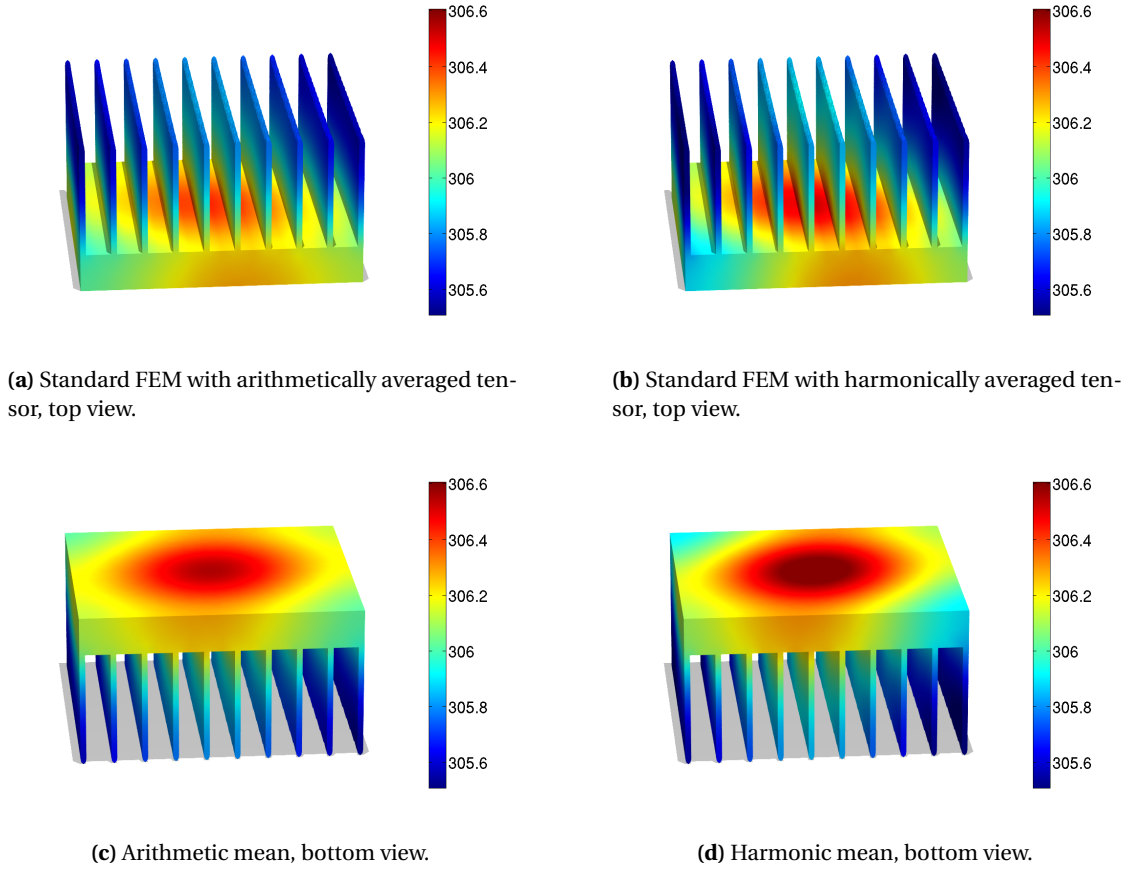
Compared to the homogenized tensor, we clearly see that the tensor obtained through arithmetic average overestimates the conductivity of the material while the tensor obtained through harmonic average underestimates the conductivity of the material.

### 3.2.4.2 Results

We compare in Figure 3.2.12 the results obtained with the FE-HMM to the solution of the FEM with the homogenized tensor. We see a very good qualitative agreement between the two solutions. Both solutions show an interesting effect, that is, the heat sink is more conductive in one diagonal direction. This is due to the anisotropy of the (homogenized) conductivity tensor.

Figure 3.2.13 shows the solutions obtained using the arithmetic and harmonic means. As expected by the over- and under-estimation of these conductivity tensors, a comparison to the homogenized solution presented in Figure 3.2.12 shows a significant discrepancy.

Table 3.2.6 provides more insights in the quality of the solution of the FE-HMM compared to the homogenized solution, by providing the values of  $\|u\|_\infty$ ,  $\inf_{x \in \Omega} |u(x)|$  and the energy norm  $\|u\|_a$  for various numerical experiments.



**Figure 3.2.13:** Steady-state heat distribution in a heat sink (problem described in Section 3.2.4). Comparison between the results obtained with an arithmetic and harmonic average of the multiscale tensor (3.2.15). We can see a significant difference with the results obtained using the homogenized tensor presented in Figure 3.2.12. The macro mesh used in both computations has 17,000 DOF. The color bar represents the temperature in K.

|        |                              | $N_{mic} = 4$ | $N_{mic} = 8$ | $N_{mic} = 16$ |
|--------|------------------------------|---------------|---------------|----------------|
| FE-HMM | $\ u\ _A$                    | 2.1668        | 2.1671        | 2.1671         |
|        | $\inf_{x \in \Omega}  u(x) $ | 305.527       | 305.527       | 305.527        |
|        | $\ u\ _\infty$               | 306.579       | 306.58        | 306.58         |

|              |                              | homogenized | arithmetic mean | harmonic mean |
|--------------|------------------------------|-------------|-----------------|---------------|
| standard FEM | $\ u\ _A$                    | 2.1874      | 2.1586          | 2.2516        |
|              | $\inf_{x \in \Omega}  u(x) $ | 305.509     | 305.54          | 305.43        |
|              | $\ u\ _\infty$               | 306.607     | 306.57          | 306.68        |

**Table 3.2.6:** Energy norm, minimal value and infinity norm for the solution of the problem described in Section 3.2.4. Comparison of the solutions obtained by the FE-HMM, a standard FEM with homogenized tensor (3.2.16), a standard FEM with tensor obtained by arithmetic or harmonic averages. The macro mesh used in all computations has 17,000 DOF ( $u$  represents the temperature in K).

### 3.2.5 Three dimensional problem: heat distribution in a microprocessor

To demonstrate the versatility of our code we finally present a real-world application, namely the steady-state temperature distribution in an embedded microprocessor as used e.g. in cellphones or automotive electronics. The increasing packaging density in microprocessors and the resulting growth in temperature requires new cooling methods, which are crucial for the efficiency and long term reliability of micro chips. The use of composite materials for lead frames and packaging material (see Figure 3.2.14) is crucial for the efficiency of the microprocessor. Understanding the effective properties of such devices is therefore of prime importance.

For example, various components of different materials (as metal and silicon) in a micro-

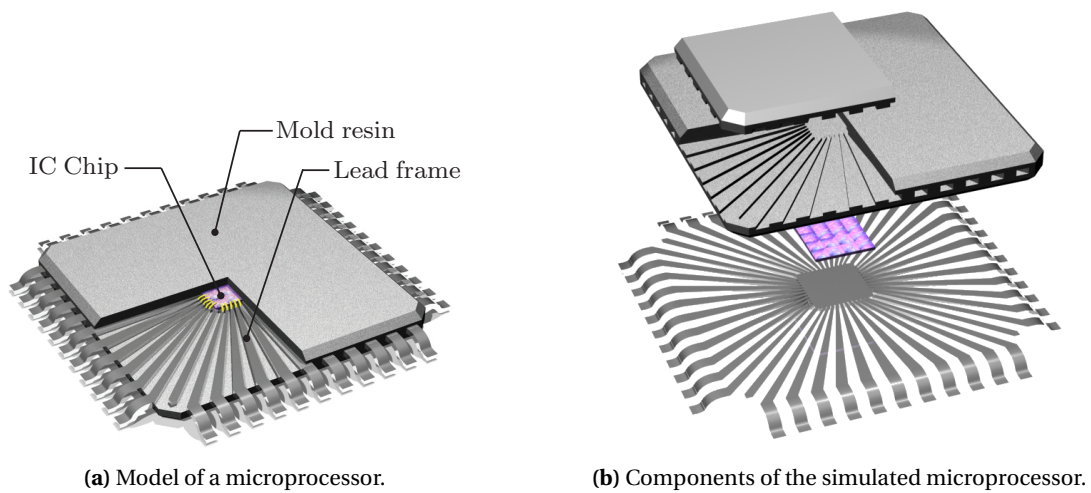


Figure 3.2.14: Model of the microprocessor used in Section 3.2.5.

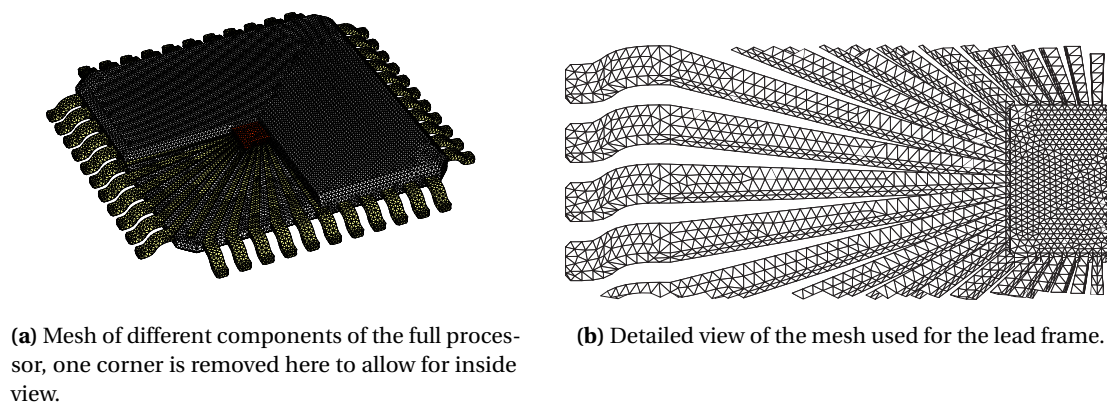


Figure 3.2.15: Coarse finite element mesh used in the 3D microprocessor problem described in Section 3.2.5 with a total of 81,000 grid points and 430,000 tetrahedra.

processor must have matching coefficients of thermal expansion (CTE), otherwise thermal stresses might cause component failure. Given that the IC chip is made of silicon, it is important to use a material for the lead frame which has a compatible CTE. Unfortunately, the obvious candidates for constructing the lead frame (such as copper or aluminum) have a CTE much higher than silicon. The way out of this problem is to use composite materials. For example nickel-iron alloys are widely used for lead frames. A second problem is to find composite materials (with matching CTEs) featuring a high thermal conductivity at the same time. We refer to [121] for details and an overview of modern composite materials with large thermal conductivities.

We will see on the example below, that our MATLAB code, although simple and short, is nevertheless capable of handling such challenging problems.

We consider a slightly simplified model of a modern embedded microprocessor which consists of the silicon IC chip, a lead frame, which in our example acts as a heat spreader, and a mold resin encapsulation (packaging), which covers the IC chip for protection (see Figure 3.2.14). While the simplified model does not reflect effects such as air flow over the package or heat conduction through the board on which the chip is mounted, these effects could be simulated without further difficulties with the proposed code. However, since we again want to avoid discussing modeling issues, we omit such effects originating from a more refined model.

### 3.2.5.1 Setup

As in Section 3.2.4, we use Fourier's law of cooling for the heat conduction within the microprocessor. Convection on the surface is again described by Robin boundary conditions. While conductivity tensors for realistic materials obtained via imaging techniques could be taken into account without further difficulties, we use simplified tensors here. To simulate the various composite materials, we use different conductivity tensors for each material. The diagonal entries of these tensors are given by

$$a_{ii,\text{leadframe}}^\varepsilon(x) = 400 \cos\left(2\pi \frac{x_i}{\varepsilon}\right) + 400 \exp\left(20\sqrt{x_1^2 + x_2^2}\right),$$

$$a_{ii,\text{resin}}^\varepsilon(x) = \begin{cases} 1.9 \cos\left(2\pi \frac{x_i}{\varepsilon}\right) + 2 & i = 1, 2, \\ 3.8 \cos\left(2\pi \frac{x_i}{\varepsilon}\right) + 4 & i = 3, \end{cases}$$

$$a_{ii,\text{chip}}^\varepsilon(x) = 140.$$

Notice that we center the chip at the origin  $x = (0, 0, 0)$  of the coordinate system. The tensor corresponding to the lead frame has a non-periodic slow variation, which models a change in the material structure from the center of the lead frame to the periphery. The tensor corresponding to the resin is chosen to be oscillating and anisotropic with a larger conductivity in the  $z$ -direction. Finally, the conductivity for the chip is assumed to be constant. There is no explicit analytical solution to this homogenization problem.

We assume a power of the chip of  $P_{\text{chip}} = 0.125\text{W}$  and consider a chip size of  $V = 2 \times 2 \times 0.2 \text{ mm}^3$ , which leads to an external heat flux  $f$  in (3.2.10) of  $f = \frac{P}{V} = 1.875 \cdot 10^8 \frac{\text{W}}{\text{m}^3}$ . We

set  $q_0 = 0$ . The ambient temperature is set to  $u_{amb} = 293.15\text{K}$ <sup>1</sup> and we fix the heat transfer coefficient  $\alpha = 20 \frac{\text{W}}{\text{m}^2\text{K}}$ , being a rough estimate for air cooling.

The macro mesh was again generated by CUBIT and consists of 81,000 grid points with a maximum tetrahedra volume of  $1.4 \cdot 10^{-3} \text{mm}^3$  (see Figure 3.2.15). The size of the full processor model is  $12.2 \times 12.2 \times 1 \text{mm}^3$ . With  $\varepsilon = 10^{-6} \text{m}$  as used below and about 10 grid points (in each direction) per oscillation length to resolve the microscopic composite heterogeneities, a full fine scale computation would lead to more than  $10^{12}$  grid points and would thus be computationally out of reach. For comparison purpose, we choose a relatively large  $\varepsilon = 5 \cdot 10^{-4} \text{m}$  to generate a reference fine scale solution on a mesh consisting of 3.9 million grid points and about 22 million tetrahedra. As we will see below, computations with realistic values of  $\varepsilon$  can be done without difficulties with the FE-HMM.

### 3.2.5.2 Results

We compare in Figure 3.2.16 a solution obtained by taking the arithmetic mean of the fast scale of the conductivity tensors, the solution of the FE-HMM and a fine-scale solution. We also compute a solution with the FE-HMM with  $\varepsilon = 10^{-6} \text{m}$ . We notice, as expected, that a naive average (such as the arithmetic average) gives a completely false result by overestimating the conductivity of the microprocessor.

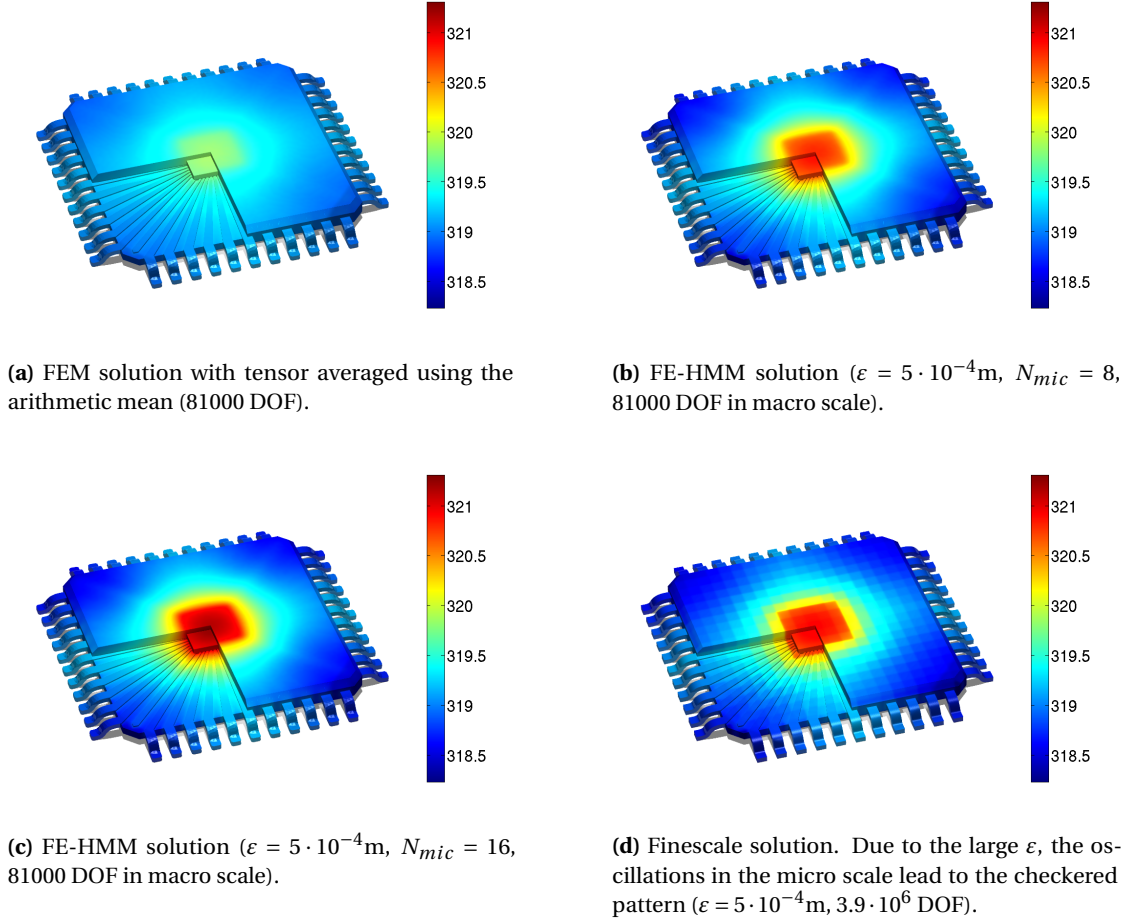
We observe in Figure 3.2.16 a very good qualitative agreement between the FE-HMM and the fine-scale solution. The results of the FE-HMM are independent of  $\varepsilon$ , as can be seen by comparing the results with  $\varepsilon = 5 \cdot 10^{-4} \text{m}$  and the results for small (realistic) heterogeneities ( $\varepsilon = 10^{-6} \text{m}$ ). This is again expected as the FE-HMM captures the effective (homogenized) solution of the multiscale problem (3.2.10).

Finally, we present in Figure 3.2.17 further results for the FE-HMM (with  $\varepsilon = 10^{-6} \text{m}$ ), where we keep the macro DOF fixed (81,000 grid points), while varying the micro degrees of freedom  $N_{mic}$ . We observe that the influence of the micro error at the macro scale is not negligible. If the small scales are not resolved on the sampling domain, a significant error occurs for the macro FE-HMM solution.

To get a rough estimate of the quality of the various experiments performed in this section, we provide in Table 3.2.7 the values of  $\min_{x \in \Omega} |u(x)|$ ,  $\|u\|_\infty$  and the energy norm  $\|u\|_a$  for the various numerical solutions.

---

<sup>1</sup>The heat in the immediate vicinity of the device can be much higher than room temperature. This should be reflected in a more realistic simulation.

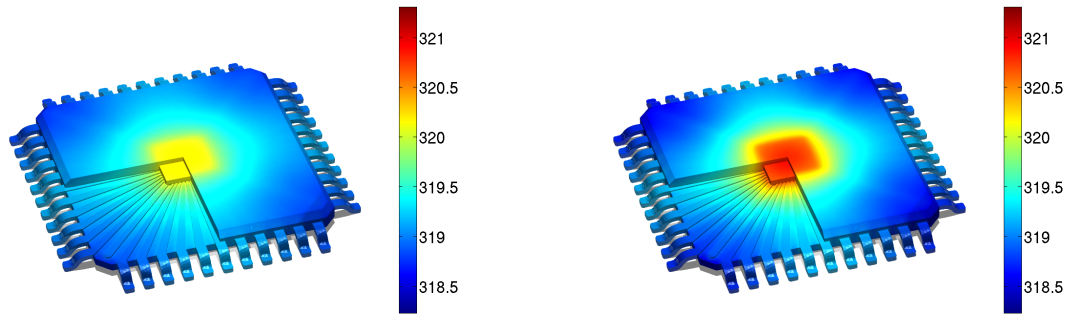


**Figure 3.2.16:** Steady state heat distribution in the microprocessor as sketched in Figure 3.2.14 for the problem described in Section 3.2.5. Comparison of the solutions obtained by a FEM with arithmetic averaging of the tensor, the FE-HMM and a finescale FEM. The color bar represents the temperature in K.

|                              |                              | $N_{mic} = 4$ | $N_{mic} = 8$ | $N_{mic} = 16$ | finescale | average |
|------------------------------|------------------------------|---------------|---------------|----------------|-----------|---------|
| $\epsilon = 5 \cdot 10^{-4}$ | $\ u\ _A$                    | 0.3906        | 0.4961        | 0.5514         | 0.5189    | 0.3122  |
|                              | $\inf_{x \in \Omega}  u(x) $ | 318.688       | 318.456       | 318.350        | 318.23    | 318.843 |
|                              | $\ u\ _\infty$               | 320.241       | 320.9         | 321.309        | 320.97    | 319.851 |
| $\epsilon = 10^{-6}$         | $\ u\ _A$                    | 0.3894        | 0.4963        | 0.5512         | -         | 0.3122  |
|                              | $\inf_{x \in \Omega}  u(x) $ | 318.688       | 318.454       | 318.351        | -         | 318.843 |
|                              | $\ u\ _\infty$               | 320.234       | 320.901       | 321.308        | -         | 319.851 |

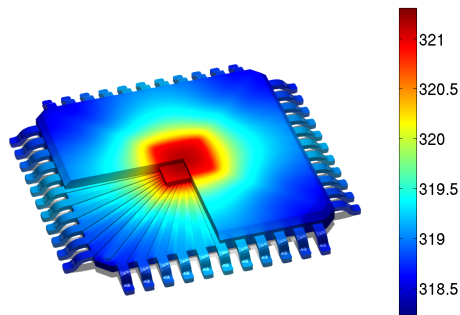
**Table 3.2.7:** Energy norm, minimal and maximum value for the solution of the 3D problem described in Section 3.2.5. Comparison of the solutions obtained by a FEM with arithmetic averaging of the tensor, the FE-HMM and a finescale FEM. Values are temperatures in K.

### Chapter 3. Implementation of the FE-HMM and Simulation of Heat Transfer Problems



(a)  $N_{mic} = 4$ ,  $\varepsilon = 10^{-6}$ m, 81000 DOF in macro scale.

(b)  $N_{mic} = 8$ ,  $\varepsilon = 10^{-6}$ m, 81000 DOF in macro scale.



(c)  $N_{mic} = 16$ ,  $\varepsilon = 10^{-6}$ m, 81000 DOF in macro scale.

**Figure 3.2.17:** Steady state heat distribution in the microprocessor as sketched in Figure 3.2.14 for the problem described in Section 3.2.5. Comparison of FE-HMM solutions with different resolutions of the micro problems. The color bar represents the temperature in K.



## 4 Conclusion of Part I

This Chapter is taken in part from [11, Section 5].

In this Part I of this thesis, we gave a review of the FE-HMM. In Chapter 2 we summarized homogenization, the single-scale FEM and the FE-HMM. In Chapter 3 we presented a short and versatile multiscale FE-HMM solver for PDEs (elliptic, parabolic) with highly oscillating coefficients. While many multiscale strategies have been developed in the past few years for multiscale PDEs, very few detailed numerical implementations were available so far. The algorithm is based on the heterogeneous multiscale method, which gives a methodology to couple efficiently macro and micro solvers. We discussed in detail the implementation of the multiscale FE solver and showed that the multiscale strategy can be built on the core structure of a standard FE code. The code also allows for different triangulations (simplicial, quadrilateral) and can trivially be parallelized.

To illustrate the performance of our code, we presented a number of numerical examples for both two and three dimensional problems with periodic, non-uniformly periodic and random tensors. Various boundary conditions and coupling conditions between macro and micro FE solvers were discussed. Furthermore, time-dependent problems were presented. We also showed an application of our code to the steady-state heat distribution in a heat sink and a microprocessor featuring a multiscale heat conductivity tensor. This simulation resembles the motivating engineering problem we stated in the introduction of this thesis in Chapter 1. Our numerical results demonstrated that, although simple, our code can nevertheless handle challenging problems.

The implementation builds the foundation for all numerical experiments of Part II of this thesis. There, we used a modified version of the implementation that is parallelized and optimized for efficiency rather than for comprehensibility to a broad audience. The implementation furthermore has been used as a basis for the numerical experiments in [13, 12, 10, 15, 16, 8].

We believe that the simplicity and the versatility of the proposed code could be useful for further developments in computational methods for multiscale PDEs, and that our code or algorithm could easily be integrated as a subroutine for more general multiscale computational

## **Chapter 4. Conclusion of Part I**

---

problems (e.g. involving optimization procedures, inverse or stochastic problems).

# **Adaptive Finite Element Heterogeneous Multiscale Methods**

**Part II**



## Introduction to Part II.

In many physical and engineering problems one is confronted with the challenge that local singularities, such as singularities caused by reentrant corners (see Section 5.7), deteriorate the overall accuracy of the numerical solution. This difficulty can be overcome by refining the mesh adaptively, and therefore putting more discretization points around the critical regions, where the solution is less regular. The goal is to equi-distribute the error among all elements in order to obtain an optimal mesh, where we have the maximal accuracy at the minimal computational cost. It is therefore crucial to know how to identify these critical areas, so that the surrounding mesh can be refined appropriately.

Besides knowing where to refine, it is very important to be able to obtain *reliable* estimates of the accuracy of the numerical solution, as we saw in the motivating example in Chapter 1. In Section 2.4.1 we recapitulated the *a priori* results for the FE-HMM. While these *a priori* error bounds hold before the calculation is carried out and hence do not depend on the numerical solution, they do depend on the *exact* solution, which is in general unknown. In addition, the *a priori* estimates only describe asymptotic error behavior and require a certain regularity of the solution, which might not be satisfied, like in the aforementioned example of reentrant corners. *A priori* estimates are therefore not suitable for providing information on the quality of an actual numerical computation for a given mesh. Hence, *a priori* results are not appropriate for driving a mesh refinement strategy in order to achieve a prescribed accuracy.

*A posteriori* error estimators, on the other hand, provide this information by using the numerically computed solution and given data of the problem. To be useful in practical applications, we would like to have the following properties from an *a posteriori* error estimator (see also [59] and [27, Chapter 6] for a longer discussion):

- the predicted error estimate should be close to the actual error, where we measure the error in a user-defined norm;
- the error estimation procedure should provide a useful stopping criterion, indicating that a given accuracy is reached;
- the error estimation procedure should be capable of steering an adaptive refinement process, where the *locally* estimated error is used to refine the mesh such that it equally distributes the error, and thus the computational effort, among all elements;
- the error estimate should be inexpensive to compute, compared to the total computation time of the problem<sup>1</sup>;
- the error bounds should be independent of unspecific constants;

---

<sup>1</sup>At the same time, the monetary cost of e.g. a sinking oil platform (see Chapter 1) has to be related to the computational cost of an estimator.

- sharp upper and lower bounds should be guaranteed. Upper bounds are sufficient to obtain a numerical solution which is below a user-specified tolerance; lower bounds ensure that the method is efficient in the sense that a (close to) minimal number of grid points is used to reach the prescribed tolerance.

*A posteriori* error estimates aim at providing error estimators which satisfy (to a certain degree) the above properties. As the exact solution is unknown, the estimates are solely based upon the numerical solution and the data. In reality, even for linear problems as considered in this work, it is in general not possible to guarantee practically useful error bounds that are *also* inexpensively computable<sup>2</sup>. However, the error estimator might still be suitable for driving a very efficient adaptive scheme and give a reasonable indication of the accuracy of the solution. For nonlinear problems, we refer the reader to [28, 59] and references therein.

Adaptive finite element methods take advantage of one or both of the following observations

1. Typically, in order to reach a certain accuracy of the solution, more degrees of freedom are needed in some parts of the domain than in others; the error on a uniform grid is typically not uniformly distributed among all elements. The non-uniform distribution of the error can e.g. be caused by singularities due to reentrant corners, which need to be specially taken care of.
2. Often scientists or engineers are not interested in the FEM solution per se, but in certain (often local) physical *quantities of interest* (such as point values or boundary fluxes) that are necessary to make a certain design decision. Adaptivity can reduce the complexity of the global solution because the solution does not have to be obtained very accurately everywhere, but only in those areas that are important to obtain an accurate quantity of interest.

While the adaptivity using *global norms* benefits from consequences of the first observation, *goal-oriented* adaptivity exploits the effects of observation one and two, but at the expense of having to solve a so-called dual problem in addition to the main (primal) FEM problem to obtain the estimate.

**Adaptive cycle.** Adaptive finite element methods follow an iteration cycle to adapt the mesh from iteration to iteration. The problem is *solved* for the numerical solution and an *a posteriori estimate* is computed based on the numerical solution. Then the (global) *a posteriori* error estimator is divided into (local) *error indicators* or *refinement indicators* on every element. These error indicators are used to steer local mesh modifications by *marking* elements which contribute “the most” – in a sense to be defined – to the global error. Finally, the marked

---

<sup>2</sup>The cost of the estimator should be small compared to the cost of the solver. To obtain error estimates in quantities of interest a dual problem must be computed, which in general is more expensive than the original (primal) numerical solution.

elements are *refined*. This leads to the following cycle

Solve → Estimate → Mark → Refine,

which is repeated until we reach a prescribed accuracy or a maximum number of cycles.

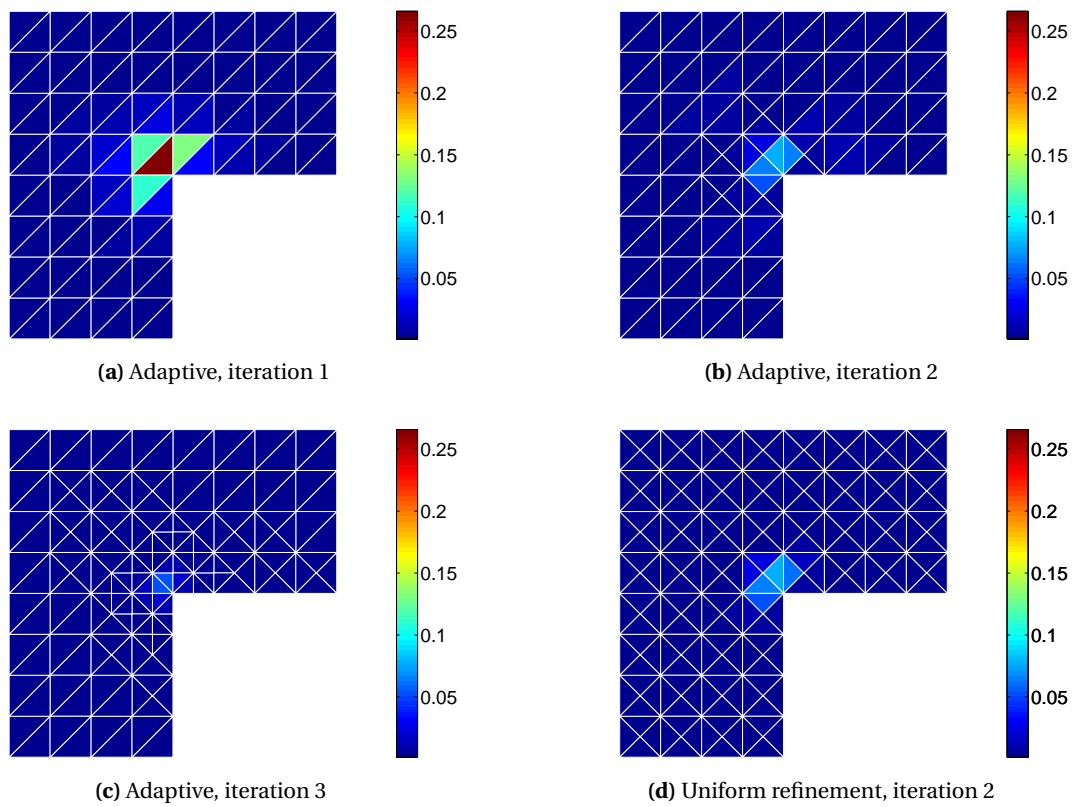
In order to increase the accuracy, we could instead of refining an element (*h-refinement*) also increase the polynomial degree (*p-refinement*<sup>3</sup>) or do both (*hp-refinement*, see [105, 77]), but throughout this work, we only consider *h-refinement*.

In Figure 4.0.1 we plot the local *error indicators* for an L-shape problem with a singularity caused by a reentrant corner. The color of the elements represents the estimated error. We can see that the algorithm refines those elements with the largest local error indicator and that the error decreases very rapidly with every iteration. We observe that at iteration 3 the error is more evenly distributed (the color is much more uniform among the different elements). While the overall estimated accuracy is comparable between Figure (b) and (d), many more elements are required for the uniform refinement used for Figure (d). In Figure 4.0.2 we show the mesh for a few iterations and finally plot a uniformly refined mesh. As we will see throughout this work, the computational cost for reaching a certain accuracy is significantly lower when using adaptivity compared to uniform refinement.

**Adaptivity in multiscale problems.** The results of this simple example indicate that adaptive methods can be highly efficient. Indeed, for single-scale FEM, adaptivity has been widely used in applications for about 30 years. We will see throughout this work that the gain in computational efficiency that is obtained when adding adaptivity to multiscale methods can be substantially higher than the gain that can be obtained when adding adaptivity to single-scale methods. The large gain in computational efficiency for multiscale methods is possible, because on the micro-level expensive micro computations can be avoided or re-used through the adaptive cycle. We will furthermore investigate how to couple adaptivity between the micro and the macro mesh, i.e., what to do with the micro mesh when we refine the corresponding macro element.

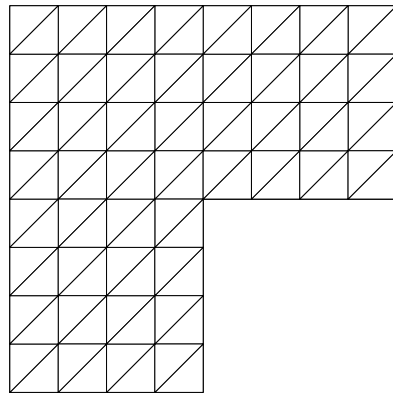
---

<sup>3</sup>This however requires local regularity of the solution in *classical* Sobolev spaces or, for the general case, the introduction of techniques such as the measurement of regularity in terms of *weighted* Sobolev Spaces, see [105, Section 4.2].

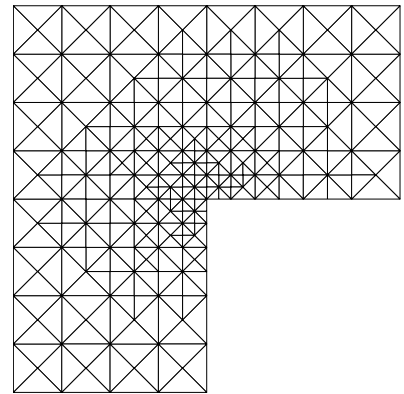


**Figure 4.0.1:** *A posteriori* error indicators on a sequence of adaptive meshes. The color of the element represents the estimated error. We can see that the estimated error decreases from iteration to iteration as the mesh is refined.

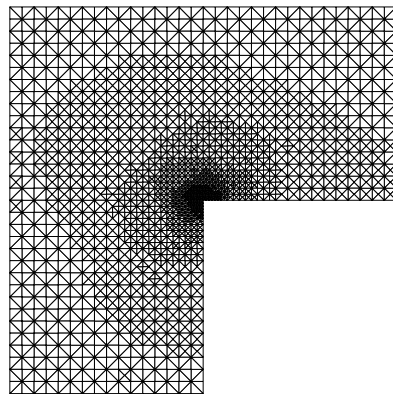




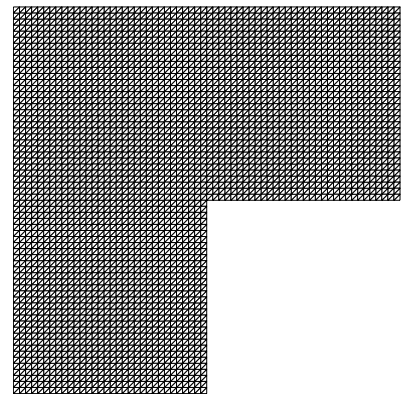
(a) Adaptive, iteration 1



(b) Adaptive, iteration 5



(c) Adaptive, iteration 10



(d) Uniform refinement

**Figure 4.0.2:** The adaptive meshes plotted at several iteration steps and a uniform mesh.

### Terminology.

We will use the following terminology. The word *error estimate* denotes a quantity that is used to approximate the unknown error; *upper* and *lower bound* denote quantities that are *always* larger and smaller, respectively, than the actual error. Therefore, error estimates are close to the actual error, but cannot be guaranteed, whereas the upper and lower bound might be less accurate, but are always guaranteed. We furthermore use the word *(local) error indicator* or *refinement indicator* (in context of the dual-weighted residual method) to denote local quantities on an element, which are used to drive an adaptive mesh refinement.

**Outline of Part II.** In **Chapter 5** we review the basics of adaptive single-scale FEM. In **Chapter 6** we derive *a posteriori* error estimates for the piecewise linear FE-HMM in the energy norm, give details of the algorithm and show numerical examples; we further compare the *a posteriori* estimates by Ohlberger obtained in a two-scale norm [92] to our estimates. In **Chapter 7** we extend these results to higher order FE. Finally, we use the techniques obtained in Chapter 7 in order to obtain the goal-oriented adaptive FE-HMM in **Chapter 8**. We end Part II with a conclusion in **Chapter 9**.

# 5 A short Review of Adaptive Finite Element Methods

In this Chapter we present a brief review of (single scale) *a posteriori* error estimators and adaptive FEM techniques and components. The single scale analysis and techniques form the foundation for our adaptive multiscale FE-HMM.

In Section 5.1 we present *a posteriori* estimates in global norms, such as the energy norm, in Section 5.2 we present the common framework of error estimation in quantities of interest and illustrate in Section 5.3 one specific method, the dual-weighted residual (DWR) method. We discuss marking schemes in Section 5.4 and show how a mesh is refined in Section 5.5. In Section 5.6 we mention adaptive techniques that go beyond the refinement of the macro (and micro) mesh. We end this Chapter in Section 5.7 with a look at how the regularity of the solution is affected in non-smooth domains and what benefits the use of higher order FE has for such problems.

## Model problem and notation.

In this Chapter we follow the notation we introduced for the single-scale FEM problem in Section 2.2. Given  $f \in L^2(\Omega)$  and a domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , we consider the single-scale, second-order elliptic equation

$$\begin{aligned} -\nabla \cdot (a \nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{5.0.1}$$

where  $a$  is symmetric, satisfies  $a(x) \in (L^\infty(\Omega))^{d \times d}$  and is uniformly elliptic and bounded. For simplicity, we only consider homogeneous Dirichlet boundary conditions. The discrete, weak-form of problem (5.0.1) is: find  $u^H \in V^H$  s.t.

$$B_{\text{classic}}(u^H, v^H) := \sum_{K \in \mathcal{T}_H} \int_K a \nabla u^H \cdot \nabla v^H \, dx = \int_\Omega f v^H \, dx \quad \forall v^H \in V^H, \tag{5.0.2}$$

where we used the spaces  $V$  and  $V^H$  defined as

$$V := H_0^1(\Omega),$$

and

$$V^H := V^p(\Omega, \mathcal{T}_H) = \{v^H \in H_0^1(\Omega); v^H|_K \in \mathcal{P}^p(K), \forall K \in \mathcal{T}_H\}.$$

Here, we use only simplicial elements and  $\mathcal{P}^p$  is defined as for (2.2.3). We emphasize that in this Chapter we use  $B_{\text{classic}}(\cdot, \cdot)$  to refer to the single-scale bilinear form as defined in (5.0.2), whereas we use  $B(\cdot, \cdot)$  to refer to the multiscale FE-HMM bilinear form (2.4.6).

## 5.1 Global error estimators

We will briefly illustrate various single-scale adaptive FE methods and start with the widely used recovery-based estimator by Zienkiewicz and Zhu. Then, we will describe residual-based estimators. Both belong to the group of so-called *global error estimators*, i.e. they estimate the error in terms of a global norm such as the energy norm. We will later see error estimators that can take local quantities of interest into account. For an extended overview of different adaptive methods, see [20, 59, 113, 114] and references therein.

### 5.1.1 Zienkiewicz-Zhu estimator

In general, the gradient of the FE solution is discontinuous across element interfaces; the basic idea of the Zienkiewicz-Zhu estimator [120, 59] is therefore to smooth the (discontinuous) gradients of the solution and quantify the error by comparing the smoothed and the non-smoothed gradient.

Let  $\varphi_i \in V^H$ ,  $i = 1, \dots, N$  be the FE basis functions and  $u^H$  be the FE-solution of (5.0.2). We determine the smoothed gradient  $\nabla u_\star^H = (r_1, \dots, r_d)^T$ , where  $r_i \in V^H$  by a standard  $L^2$  projection

$$\int_{\Omega} \varphi_i (\nabla u_\star^H - \nabla u^H) dx = 0, \quad i = 1, \dots, N.$$

The smoothed gradient often is only computed approximately. Then the error estimator in the energy norm is obtained as follows

$$\| \| e^H \| \|^2 \approx B_{\text{classic}}(e_\star^H, e_\star^H) = \int_{\Omega} a \nabla e_\star^H \cdot \nabla e_\star^H dx,$$

where  $e^H = u - u^H$  and  $\nabla e_\star^H = \nabla u_\star^H - \nabla u^H$ .

Although it is widely used and researched, the Zienkiewicz-Zhu estimator has several disadvantages including that it is assumed that smooth gradients imply accurate gradients. For a further discussion, see [59] and references therein.

### 5.1.2 Residual-based estimators

We will base our adaptive FE-HMM presented in Chapter 6 on the residual-based error estimators. This method seeks for an error estimate in the (global) energy norm and involves a direct computation of so-called *interior element residuals* and *jumps*. We will briefly show here the ansatz of the method for the classical single-scale case, and postpone a detailed proof and description for the FE-HMM until Chapter 6. We refer for details on the single-scale case to the work of Verfürth [113, 112] and Ainsworth and Oden [20, 19]; for the convergence of the adaptive method see [82, 48]. We will follow the presentation in [86].

We consider problem (5.0.1). We first define the jump and subsequently the element and jump residual and then find *a posteriori* error estimates.

**Definition 18.** (See [20, 113]). We define the *jump* as

$$\llbracket a \nabla u^H \rrbracket_e := \begin{cases} -a \nabla u^{H^+} \cdot n^+ - a \nabla u^{H^-} \cdot n^- & \text{for } e \notin \partial \Omega, \\ 0 & \text{for } e \subset \partial \Omega, \end{cases} \quad (5.1.1)$$

where where  $e$  is the interface between elements  $K^+$  and  $K^-$  with unit outward normals  $n^+$  and  $n^-$ , respectively.

**Definition 19.** (See [20, 113]). We define the *element residual*  $R_K(u^H)$  and the *jump residual*  $J_e(u^H)$  as

$$\begin{aligned} R_K(u^H) &:= f + \nabla \cdot (a \nabla u^H), \\ J_e(u^H) &:= \llbracket a \nabla u^H \rrbracket_e. \end{aligned}$$

Using Definition 18 and 19, we can construct the crucial component of the residual-based estimators, the *error representation formula* given by

$$B_{\text{classic}}(e^H, v) = \sum_{K \in \mathcal{T}_H} \int_K R_K(u^H) v \, dx + \sum_{e \in \mathcal{E}_H} \int_e J_e(u^H) v \, ds, \quad \forall v \in V,$$

where  $u^H \in V^H$  is the FE solution of (5.0.2). A calculation involving element-wise Cauchy-Schwarz inequality, bubble functions, using an appropriate test function  $v$  and applying various interpolation estimates leads to the upper bound for the error  $\| \| e^H \| \|$  in the energy norm given by

$$\| \| e^H \| \|_{\Omega}^2 \leq C \sum_{K \in \mathcal{T}_H} \eta_H^2(K) + \xi_H^2(K)$$

and to the following lower bound

$$\eta_H^2(K) \leq C \left( \| \| e^H \| \|_{\omega_K}^2 + \xi_H^2(\omega_K) \right).$$

Here, we used the so-called local residual

$$\eta_H^2(K) = H_K^2 \|R_K(u^H)\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \partial K} H_K \|J_e(u^H)\|_{L^2(e)}^2,$$

where  $\omega_K$  is a neighborhood of  $K$  and  $\xi_H(K)$  is the so-called data approximation error. We will describe the various terms involved in the above inequality and explain the details of the proof in Chapter 6. Having a lower and upper bound means that the method is both effective (i.e., the error is smaller than a prescribed tolerance) and efficient (i.e., the mesh is near optimal). The algorithm follows the cycle

Solve → Estimate → Mark → Refine

with an appropriate marking (see Section 5.4) and refinement (see Section 5.5) scheme.

## 5.2 Goal-oriented error estimators

Often, physicists and engineers are not interested in the error in the energy norm, but rather in the error regarding certain (often local) physical quantities of interest (QoI) that are necessary for making a design decision. This is achieved using so-called goal-oriented error estimators. We will follow the presentation in [28, 59] and introduce the general duality-based goal-oriented adaptivity framework. Two specific methods will be illustrated. A third method, the so-called dual-weighted residual (DWR) method will be described in the separate Section 5.3 in detail. We refer to [59, 31, 20, 58, 96, 88] for an overview of the general framework.

### General framework: duality and goal oriented estimators

We again consider the model problem (5.0.1). Suppose that we want to know the error  $e^H = u - u^H$  in terms of certain quantities of interests, which can be expressed in terms of a linear, bounded functional  $J(u)$ . Examples of quantities of interest include

- the *mean value* over a domain  $S \subset \Omega$

$$J(u) := \frac{1}{|S|} \int_S u \, dx.$$

- *point-wise errors* at a point  $x^* \in \bar{\Omega}$  such as

$$J(u) := u(x^*) \quad \text{or} \quad J(u) := \partial_x u(x)|_{x^*}.$$

Special care is required for point-wise quantities of interest, see Remark 22.

- *mean normal flux* over the boundary such as

$$J(u) := \int_{\partial\Omega} \partial_n u \, ds.$$

Here, the dual problem is not well posed and regularization should be used, see [28,

Example 3.5].

- *global norms*<sup>1</sup> such as

$$\|\nabla(u - u^H)\|_{L^2(\Omega)} \quad \text{or} \quad \|u - u^H\|_{L^2(\Omega)}.$$

The core of the method is an auxiliary (dual) problem, which “extracts” the necessary information from the primal solution to express the error in quantities of interest. We formulate the method for *linear*, bounded functionals  $J(u)$ , but remark that an extension to non-linear functionals is possible, see [28].

We emphasize that the quantity of interest must be well-defined; for example asking for the maximum error in the derivative or flux in a domain with reentrant corners is of no use, as the exact solution does not exist since the derivative of the solution around a corner point is usually unbounded (see [27] and Section 5.7).

Let  $J: V \rightarrow \mathbb{R}$  be a linear, bounded functional. We are interested in controlling the error in terms of the quantity  $J(e^H) = J(u) - J(u^H)$ . The **primal problem** is given by (5.0.1), which is in weak form: find  $u \in V$  s.t.

$$B_{\text{classic}}(u, v) = \int_{\Omega} a \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in V. \quad (5.2.1)$$

In discretized form, problem (5.2.1) reads: find  $u^H \in V^H$  s.t.

$$B_{\text{classic}}(u^H, v^H) = \int_{\Omega} f v^H \, dx \quad \forall v^H \in V^H(\Omega, \mathcal{T}_H).$$

Then let the *dual solution*  $z \in V$  be the solution to the **dual problem**: find  $z \in V$  s.t.

$$B_{\text{classic}}(\varphi, z) = J(\varphi) \quad \forall \varphi \in V.$$

We use the primal Galerkin-orthogonality

$$B_{\text{classic}}(u - u^H, \psi^H) = 0 \quad \forall \psi^H \in V^H(\Omega, \mathcal{T}_H)$$

to obtain

$$J(e^H) = B_{\text{classic}}(e^H, z) \quad (5.2.2)$$

$$= B_{\text{classic}}(e^H, z - \psi^H) \quad (5.2.3)$$

$$= B_{\text{classic}}(u, z - \psi^H) - B_{\text{classic}}(u^H, z - \psi^H) \quad (5.2.4)$$

$$= \int_{\Omega} f(z - \psi^H) \, dx - B_{\text{classic}}(u^H, z - \psi^H)$$

$$=: R(u^H)(z - \psi^H),$$

<sup>1</sup>The error in the presented norms may be estimated using a linear functional by using the technique presented in [28, Chapter 3.2, 3.3].

where  $R(u^H)(z - \psi^H)$  is called the *residual*. The function  $\psi^H \in V^H$  introduced in (5.2.3) can be chosen arbitrarily. We can therefore exploit the free choice to follow different goals, such as the reduction of the computational effort of the estimator or lowering its data approximation error.

*Remark 20.* We see from (5.2.2) that simply replacing  $z$  by a discrete solution  $z^H \in V^H$  will lead to  $J(e^H) = 0$  due to Galerkin-orthogonality. There are different strategies to overcome this problem (see [28, Chapter 4.1 and 5.2]):

- Approximate the dual solution  $z$  by a discrete solution  $z^{\mathcal{H}}$  in a “richer” space  $V^{\mathcal{H}}$ , such as a higher order polynomial space  $V^{\mathcal{H}} := V^{\tilde{p}}(\Omega, \mathcal{T}_H)$  where  $\tilde{p} > p$ .
- Approximate the dual solution  $z$  by higher-order interpolation. In [28] two-dimensional bilinear elements with hanging nodes are used; to approximate the dual solution  $z$ , a solution  $z^H \in V^H$  is used and a patch-wise biquadratic interpolation is performed to obtain  $z \approx I_{2H}^{(2)} z^H$ . We refer to [28, 87, 106] for a discussion about the approximation error.
- In the  $h$ -approach one divides each element  $K$  into finer subelements of the same polynomial order. This technique is mentioned in [96] and was used in [66].
- Approximation by local residual problems (which also need to be solved in a richer space than  $V^H$ ), which avoids the full solution of the dual problem, see [28, 96, 97, 88].

*Remark 21.* In our adaptive dual-weighted residual (DWR) FE-HMM, we will follow the first approach of seeking an approximation to  $z$  in a higher order polynomial space  $V^{\mathcal{H}}$ . This approach leads to a reliable method only if we can neglect the approximation error  $\|z - z^{\mathcal{H}}\|$ . While in principle this is the case for a sufficiently small global mesh size [87], rather coarse or strongly graded meshes are frequently used in context of adaptive FEM. It was shown by Nochetto et al. [87] that neglecting the approximation error can cause a severe underestimation of the error  $J(e^H)$ , thus leading to an unreliable method. The severe underestimation of the error can occur even in rather simple examples, see [87]. They proposed a safeguarded DWR FEM method, which leads to unconditionally reliable enhanced estimators which asymptotically coincide with the original DWR method. The main tool is a decomposition of the solution into a smooth part and into singular functions on the corners (see Section 5.7). However, only the case  $a(x) \equiv 1$  is considered in [87]. We further mention the very recent effort by Ainsworth and Rankin to develop guaranteed and fully computable bounds on the error in quantities of interest, see [21]. There, the discrepancy term  $B_{\text{classic}}(e^H, z - z^{\mathcal{H}})$  is estimated using *a posteriori* error estimates as  $|B_{\text{classic}}(e^H, z - z^{\mathcal{H}})| \leq \| \| e^H \| \| \| z - z^{\mathcal{H}} \| \|$ .

*Remark 22.* Special care has to be taken for some quantities of interest such as the point-wise and point-wise derivative errors, as the solution  $u \in H^1(\Omega)$  of the primal problem might not be continuous (in dimension  $d > 1$ ). Furthermore, point-wise derivatives of a FEM solution are in general not defined on element interfaces.



To estimate point-wise errors or point-wise derivative errors at a point  $x^*$ , Bangerth and Rannacher suggest in [28, Chapter 3.3] to use the regularized functional

$$J(u) := \frac{1}{|S^\epsilon|} \int_{S^\epsilon} u \, dx = u(x^*) + \mathcal{O}(\epsilon^2), \quad J(u) := \frac{1}{|S^\epsilon|} \int_{S^\epsilon} \partial_n u \, dx = \partial_n u(x^*) + \mathcal{O}(\epsilon^2),$$

where the domain  $S^\epsilon$  is the  $\epsilon$ -ball centered around  $x^*$ . Prudhomme and Oden suggest in [96] to use the so-called *mollification* (see [91, Chap. 2]), which introduces the regularized quantity of interest for the point-wise error

$$J_\epsilon(u) := \int_{\Omega} u(x) k_\epsilon(x - x^*) \, dx,$$

or for the point-wise directional derivative

$$J_\epsilon(u) := \int_{\Omega} \nabla u(x) \cdot n k_\epsilon(x - x^*) \, dx,$$

where the mollifier  $k_\epsilon$  is a parameter-dependent family of  $C^\infty(\mathbb{R})$  functions with compact support. The mollifier is commonly chosen in the form

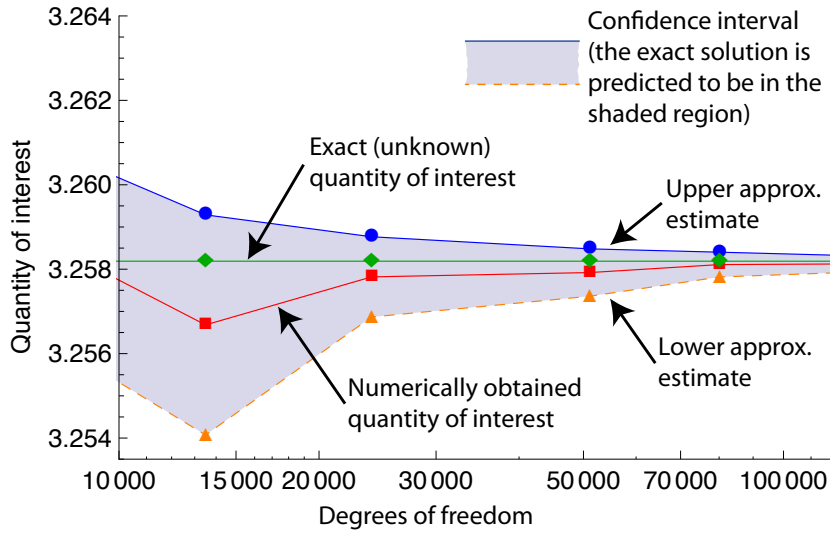
$$k_\epsilon(x) = \begin{cases} C \exp\left(-\frac{\epsilon^2}{\epsilon^2 - |x|^2}\right) & \text{if } |x| < \epsilon \\ 0 & \text{if } |x| \geq \epsilon, \end{cases}$$

where the normalization constant is given by  $\int_{\Omega} k_\epsilon(x - x^*) \, dx = 1$ .

**Confidence intervals.** The benefit of the duality-based goal-oriented method is not only that we can choose from a variety of different quantities of interest, but also that the method in general gives an estimate, which is – compared to the residual-based method – closer to the real error, as the duality based methods are based on an *exact* representation of the error in the quantity of interest. For global residual-based estimates, we can obtain an upper and a lower estimate for the error; these bounds however depend on an *unspecific constant* which in practice can be very difficult to estimate. In contrast to this, the duality-based estimates involve a *dual solution* and are therefore more expensive, but they are usually<sup>2</sup> independent of the aforementioned unspecific constant.

While there always is a data approximation error (thus the estimate based on the exact error representation usually does not coincide with the exact error), the duality-based estimates are independent of unspecific constants and can therefore be used for specifying a certain interval where we expect the true solution to be in. This is in general much more difficult to achieve with residual-based estimates. We can therefore provide the physicist or engineer not only with the *numerically* computed quantity of interest, but also with an estimate on the *exact* quantity of interest. This is illustrated in Figure 5.2.1. We see that by increasing the

<sup>2</sup>This is the case for most duality-based methods that are used in practice (such as the dual-weighted residual based method), but e.g. not for the energy-norm based method presented in Section 5.2.1.



**Figure 5.2.1:** The use of duality-based goal-oriented adaptive methods allows to specify – solely based on the *a posteriori* estimate – a confidence interval, where we expect the exact quantity of interest to be.

degrees of freedom, the numerical quantity of interest converges to the exact value and the width of the confidence interval converges to zero.

*Remark 23.* While in some goal-oriented methods such as [96], there are explicit upper and lower bounds, in the DWR method, we have an exact error representation that involves an unknown data approximation error. We *assume* the unknown data approximation error to be of equal size as the (absolute value of the) estimated error in the quantity of interest, which gives us an upper and lower approximation estimate. Therefore, the confidence interval is rather an indication than a reliable bound.

We have seen the basic framework building on (5.2.4), which is common to duality-based goal-oriented methods. We will now present three different specific ways of applying the framework to obtain error estimates.

### 5.2.1 Energy-norm based estimates

Starting from (5.2.3), one can simply apply the Cauchy-Schwarz inequality (and boundedness of  $a(x)$ ) to obtain the following upper bound

$$|J(e^H)| \leq |B_{\text{classic}}(u - u^H, z - z^H)| \leq C \|u - u^H\|_{H^1(\Omega)} \|z - z^H\|_{H^1(\Omega)},$$

or element-wise this can be written as

$$|J(e^H)| \leq \sum_{K \in \mathcal{T}_H} C \|u - u^H\|_{H^1(K)} \|z - z^H\|_{H^1(K)}.$$

We notice that the error in the quantity of interest converges faster than the error  $\|u - u^H\|_{H^1(\Omega)}$ . However, this type of error estimate will typically dramatically overestimate

the error, as it does not allow for cancelation of errors among different elements (see also [59]). The goal is thus to avoid the use of the Cauchy-Schwarz inequality.

#### 5.2.2 Goal-oriented error estimators by Oden and Prudhomme

Prudhomme and Oden have suggested in [96, 88] to use the parallelogram identity in order to obtain upper and lower bounds of the error. Let the dual error be  $\epsilon^H = z - z^H$ . Then the error representation is given by

$$J(e^H) = B_{\text{classic}}(e^H, \epsilon^H) = \frac{1}{4} \left\| \left\| s e^H + \frac{\epsilon^H}{s} \right\| \right\|^2 - \frac{1}{4} \left\| \left\| s e^H - \frac{\epsilon^H}{s} \right\| \right\|^2,$$

where the scaling factor  $s \in \mathbb{R}$  is  $s = \sqrt{\frac{\|\epsilon^H\|}{\|e^H\|}}$  and where the energy norm is denoted as  $\|v\| := \sqrt{B_{\text{classic}}(v, v)}$ . With the definition of the error indicators  $\eta_{\text{low}}^+, \eta_{\text{low}}^-, \eta_{\text{upp}}^+, \eta_{\text{upp}}^-$  which satisfy

$$\begin{aligned} \eta_{\text{low}}^+ &\leq \left\| \left\| s e^H + \frac{\epsilon^H}{s} \right\| \right\| \leq \eta_{\text{upp}}^+ \\ \eta_{\text{low}}^- &\leq \left\| \left\| s e^H - \frac{\epsilon^H}{s} \right\| \right\| \leq \eta_{\text{upp}}^-, \end{aligned}$$

we obtain the following lower and upper bounds

$$\frac{1}{4} (\eta_{\text{low}}^+)^2 - \frac{1}{4} (\eta_{\text{upp}}^-)^2 \leq J(e^H) \leq \frac{1}{4} (\eta_{\text{upp}}^+)^2 - \frac{1}{4} (\eta_{\text{low}}^-)^2.$$

Again we remark that an approximation to the dual solution has to be found in a different space than the primal solution; typically special spaces involving polynomials of higher orders are used (e.g. polynomials of degree  $p$  in the primal problem and polynomials of degree between  $p+1$  and  $p+q$  in the dual problem), we refer the reader to [89, 96, 97]. Prudhomme et al. have successfully used their goal-oriented method in complex engineering applications [97, 89] and showed error estimates that are in good agreement with the real error.

### 5.3 The Dual Weighted Residual (DWR) method

For our goal-oriented adaptive FE-HMM, we will use the so-called *dual weighted residual method* (DWR), see [31, 28] and references therein. The error in the quantity of interest is expressed by the following exact representation

$$\begin{aligned} J(e^H) &= B_{\text{classic}}(e^H, z - \psi^H) \\ &= \sum_{K \in \mathcal{T}_H} \left\{ \int_K R_{I,H}(z - \psi^H) dx + \int_{\partial K} R_{J,H}(z - \psi^H) ds \right\}, \end{aligned}$$

where  $\psi^H \in V^H$  is an arbitrary function, and the interior and jump residuals  $R_{I,H}$  and  $R_{J,H}$  are, respectively, given by

$$R_{I,H}|_K := f + \nabla \cdot (a \nabla u^H) \quad (5.3.1)$$

$$R_{J,H}|_e := -\frac{1}{2} \llbracket a \nabla u^H \rrbracket_e, \quad (5.3.2)$$

where we used the definition of the flux (5.1.1). Here we assume a conforming method, i.e., we ignore any type of quadrature. In the context of the FE-HMM, we will take the quadrature error into account, which leads to an additional term  $\xi_H(K)$ , the *data approximation error*. A rigorous analysis of the FE-HMM will follow in Section 8.5.

*Remark 24.* While in the literature typically integration by parts is applied to obtain (5.3.1) and (5.3.2), the integration by parts is not necessarily required (see also [59, Section 5.2.2]). Keeping the original form of the equation might facilitate the numerical evaluation of the expressions.

Unlike *a posteriori* estimators involving Cauchy-Schwarz inequalities (such as residual-based estimators), the DWR method allows for cancelation of errors among different elements over the domain. The DWR therefore does not come with the typical significant overestimation of the error and the estimates do not involve unspecific constants<sup>3</sup>. This however requires to have error estimators with local components that are signed. We therefore distinguish between the (always positive) *refinement indicators*  $\bar{\eta}_H(K)$ , used to drive the adaptive mesh refinement procedure and for constructing an upper error bound, and the *signed (local) error estimator*  $\eta_H(K)$  which is used to estimate the error  $J(e^H) \approx \sum_{K \in \mathcal{T}_H} \eta_H(K)$ . See also the terminology at the beginning of Part II.

**Proposition 25** (A posteriori DWR refinement indicators, error estimator and error bound [28]). *We have the following a posteriori element refinement indicators*

$$\bar{\eta}_H(K) = \left| \int_K R_{I,H}(z - \psi^H) dx + \int_{\partial K} R_{J,H}(z - \psi^H) ds \right|,$$

which yield the upper a posteriori error bound

$$|J(e^H)| \leq \sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K) + |\xi_H(K)|, \quad (5.3.3)$$

where  $\xi_H(K)$  is the data approximation error on  $K$ . Furthermore, we have the error estimator

$$J(e^H) = \sum_{K \in \mathcal{T}_H} \eta_H(K) + \xi_H(K),$$

where

$$\eta_H(K) = \int_K R_{I,H}(z - \psi^H) dx + \int_{\partial K} R_{J,H}(z - \psi^H) ds.$$

---

<sup>3</sup>We remark that for PDEs other than the linear elliptic problems considered throughout this thesis, constants may arise.

We will investigate the data approximation error in the case of the FE-HMM in detail in Section 8.5. We thus have an error estimate, which exploits the cancelation of errors among the elements and refinement indicators, which are always positive and can drive the mesh refinement.

**Proposition 26** (A posteriori error estimate [28]). *The inequality (5.3.3) can be further decomposed as*

$$|J(u^0 - u^H)| \leq \sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K) + |\xi_H(K)| \quad (5.3.4)$$

$$\leq \sum_{K \in \mathcal{T}_H} \rho_K \omega_K + \tilde{\xi}_H(K), \quad (5.3.5)$$

where the cell-residuals  $\rho_K$  and the weights  $\omega_K$  are given by

$$\rho_K := \left( \|R_{I,H}\|_K^2 + \frac{1}{H_K} \|R_{J,H}\|_{\partial K}^2 \right)^{1/2}$$

$$\omega_K := \left( \|z - \psi^H\|_K^2 + H_K \|z - \psi^H\|_{\partial K}^2 \right)^{1/2},$$

and where for the modified data approximation error  $\tilde{\xi}_H(K)$  we sum over the norms of all of its components.

*Remark 27.* Proposition 26 explains where the name dual-weighted residual comes from. The dual solution acts as a weight function for the cell-residuals.

The step from (5.3.4) to (5.3.5) does not add any further information either on the approximation of the dual solution  $z$  or on the approximation of the tensor  $a(x)$ . Therefore,  $\sum_{K \in \mathcal{T}_H} \rho_K \omega_K + \tilde{\xi}_H(K)$  is bigger than  $\sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K) + |\xi_H(K)|$  without increased reliability, see [87]. Thus we will use the DWR error estimate (5.3.3) in what follows.

Bangerth and Rannacher show in [28] several example problems where the dual solution needed in (5.3.5) can be estimated *a priori*. This however is not possible for general quantities of interest and therefore the dual problem has to be solved numerically. As we mentioned in Remark 20, we cannot simply choose  $z \approx z^H \in V^H$ . We will instead approximate  $z \approx z^{\mathcal{H}}$  using a higher order polynomial space  $V^{\mathcal{H}} := V^{\tilde{p}}(\Omega, \mathcal{T}_H)$  where  $\tilde{p} > p$ . This leads to the following Algorithm.

**Algorithm 28** (Adaptive DWR-FEM). *The algorithm for the adaptive, goal-oriented DWR FE method follows the Solve  $\rightarrow$  Estimate  $\rightarrow$  Mark  $\rightarrow$  Refine cycle with some modifications.*

**Solve.** *Solve the primal problem in  $V^H$  for  $u^H$  and the dual problem in  $V^{\mathcal{H}}$  for  $z^{\mathcal{H}}$ .*

**Estimate.** *Estimate the error in the quantity of interest  $J(u - u^H)$  by computing the indicators  $\eta_H(K)$ .*

**Mark.** *Mark the elements on a subset  $\tilde{\mathcal{T}}_H$  of  $\mathcal{T}_H$  based on the refinement indicators  $\bar{\eta}_H(K) = |\eta_H(K)|$ .*

**Refine.** Refine the marked elements (and some neighbors for mesh conformity) and update the mesh for the dual problem accordingly.

We use the same mesh for both the primal and dual solution (but with different shape functions), but other choices are possible, see [28].

### 5.3.1 Choice of $\psi^H$

Due to the Galerkin-orthogonality,  $\psi^H$  can be chosen freely. In the residual-based case, where we have an upper and a lower bound, the choice is motivated as follows. There, we choose  $\psi^H$  such that it brings the upper and lower estimate close to the exact error. As  $\nabla e^H$  is not necessarily continuous,  $\psi^H$  is chosen to be the Clément interpolant of  $e^H$ . In the DWR case, we work with an *exact* representation of the error. Therefore, the choice of  $\psi^H$  is not straightforward.

This freedom of choice is typically used with the aim of increasing the accuracy or reducing the computational cost of the error estimator. In [28, 99] it is suggested to choose  $\psi^H$  as an interpolation of  $z$  on a patch  $P$  of  $2 \times 2$  elements (there, rectangular elements and hanging nodes are used) by requiring the interpolant  $I_P z =: \psi^H$  to satisfy

$$\int_P (z - I_P z) dx = 0.$$

As  $z - I_P z$  will be orthogonal to all patch-wise constant functions, it is argued that the interior residual term is smaller by  $\mathcal{O}(H)$  than the jump residual term and thus can be ignored for  $H \rightarrow 0$ .

We consider the choice of  $\psi^H$  in the context of the DWR FE-HMM in Section 8.5.1.

### 5.3.2 Convergence rates

We only require the functional  $J(\cdot)$  to be linear and bounded. Therefore, no *a priori* estimates exist for general quantities of interest that could be used to extract useful general convergence rates, even for single-scale FEM. We know asymptotic convergence rates for a few quantities of interest, such as (presented for the case of piecewise linear elements)

- for the global  $H^1$ -norm we expect a convergence rate of  $J(e^H) \propto H$ ,
- for the global  $L^2$ -norm we expect a convergence rate of  $J(e^H) \propto H^2$ ,
- for the point-wise error we have a rough estimate of  $J(e^H) \propto H^3$ , see [28, Section 3.3].

As we will see in the following chapters, the (macroscopic) convergence rates are crucial to find an optimal multiscale algorithm. We will investigate this further in the case of the adaptive DWR FE-HMM in Section 8.3.1.

## 5.4 Marking schemes

We now seek for a suitable mesh refinement that reduce the error, while at the same time keeps the number of new elements small. Suppose that we have given local error indicators

$\bar{\eta}_H(K)$  from an *a posteriori* estimate

$$|J(e^H)| \leq \sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K).$$

If the stopping criterion  $|\sum_{K \in \mathcal{T}_H} \eta_H(K)| \leq tol$  is met, we stop the adaptive iteration cycle and accept  $u^H$  as the approximate solution of the problem, for which the quantity of interest  $J(u^H)$  (or energy norm  $\|u^H\|$ ) is within the tolerance  $tol$ . Otherwise, we continue with another iteration of the adaptive cycle. We sort the macro elements  $K_i \in \mathcal{T}_H$ ,  $i = 1, \dots, N$  of the macro FE mesh according to a decreasing local error indicator s.t.

$$\bar{\eta}_H(K_1) \geq \bar{\eta}_H(K_2) \geq \dots \geq \bar{\eta}_H(K_N).$$

For the mesh adaption itself, there are many different strategies. We want to present two of them in what follows and refer to [104, Chapter 1.5] and [28, Chapter 4.2] for an overview. We remark that some other methods target to directly construct an “optimal” mesh, which satisfies the error tolerance  $\bar{\eta}_H(K) \approx tol$  on its elements  $K$  without going through an iterative one-step process. See e.g. the *mesh-optimization strategy* in [28].

All strategies are based on the assumption that for an optimal mesh, the local error is equi-distributed among all the elements of the mesh. As the true error is unknown, Babuška and Rheinboldt suggest in [26] that if the local error indicators  $\bar{\eta}_H(K)$  are nearly equal on all elements, then the mesh is approximately optimal. We therefore mark a subset of the elements in  $\mathcal{T}_H$  (with a large local error indicator) for refinement and leave the rest of the elements unmarked. On the one hand, refining only few elements at every iteration will lead to a more optimal mesh. On the other hand, at every iteration the computationally expensive primal (and dual) FE solution has to be found; hence we would like to refine many elements, such that we only need a few iterations in order to reach a certain accuracy. The goal is therefore to find a balance between the two aforementioned opposing strategies.

### Marking strategy E

We first consider *marking strategy E* as proposed by Dörfler [48] and used in [82, 85].

**Algorithm 29** (Marking Strategy E). *For a given a user-defined parameter  $0 < \theta < 1$ , find a minimal subset  $\tilde{\mathcal{T}}_H$  of  $\mathcal{T}_H$  such that*

$$\sum_{K \in \tilde{\mathcal{T}}_H} \bar{\eta}_H(K)^2 \geq \theta^2 \bar{\eta}_H(\Omega)^2.$$

*Mark all elements in the subset  $\tilde{\mathcal{T}}_H$  for refinement.*

As shown in [82, 85], for the residual-based adaptive FEM and in the absence of data approximation errors, this marking strategy guarantees error reduction when used with an appropriate mesh refinement technique.

**Maximum strategy**

Another strategy, the so-called maximum strategy is used in [96] and given as follows (see also [104]).

**Algorithm 30** (Maximum strategy). *For a given user-defined parameter  $0 < \vartheta < 1$ , mark all elements  $K \in \mathcal{T}_H$  with*

$$\bar{\eta}_H(K) > \vartheta \max_{\tilde{K} \in \mathcal{T}_H} \bar{\eta}_H(\tilde{K}).$$

A typical value is  $\vartheta = 0.5$ .

**5.5 Mesh refinement techniques**

Once a set of triangles is marked for refinement, the actual element bisection is performed. It is crucial to ensure that after refining the mesh it is still conforming and shape regular. Among the various methods such as the *red* and *green refinement* [29] or the *longest edge bisection* [100], we will only consider the *newest vertex bisection* [81, 80]. In many adaptive FEM packages, the newest vertex bisection is used for 2D and longest edge bisection is used for 3D calculations [41, 42, 104]. We restrict ourselves to the description of bisection of triangles, as for quadrilaterals and hexahedra, if we bisect one element, mesh conformity requires the repetitive bisection of all their neighboring elements up to the boundary. For the bisection of tetrahedra, we refer the reader to [104, 41, 42].

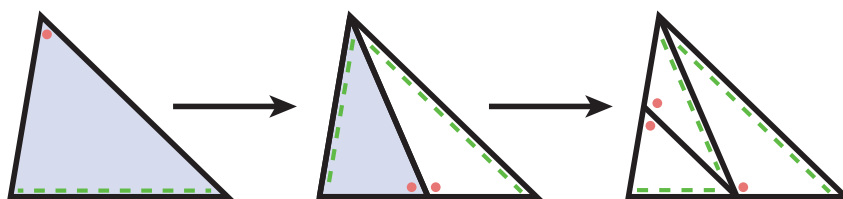
**Newest vertex bisection.** Let  $\mathcal{T}_H$  be a shape regular triangulation of  $\Omega$ . Then for every triangle  $K \in \mathcal{T}_H$  we label one of its vertices as the *newest vertex*. We call the opposing edge the *refinement edge*. Once an initial labeling is done, the refinement rule for newest vertex bisection is as follows (see also Figure 5.5.1 for an illustration)

1. We bisect a triangle  $K$  into two elements by connecting the newest vertex to the midpoint of the refinement edge.
2. We assign the new vertex at the midpoint of the refinement edge to be the *newest vertex* of the child elements.

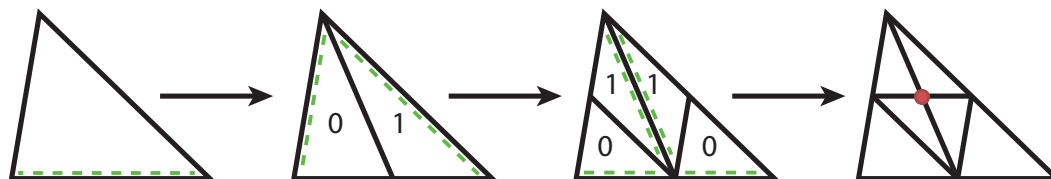
After bisecting the marked triangles, the mesh conformity is typically broken. Therefore additional steps are required to restore mesh conformity by refining the neighboring elements appropriately. This bisection process will not only lead to a series of nested grids, but also generate a mesh that is conforming and shape regular. In our numerical experiments we use the newest vertex bisection based on the implementation in [41, 42].

**Interior nodes.** Morin et al. show the convergence of adaptive residual-based FEM in [82], where they further demonstrate that a special condition on the refinement is required in order to ensure that an error reduction is obtained. The crucial point of this refinement method is the generation of an interior point at every marked element. We refer to [82, Example 3.6]





**Figure 5.5.1:** *Newest vertex bisection.* The marked element is shown in shaded blue, the newest vertex is indicated with a red dot and the refinement edges are shown with green dashed lines.



**Figure 5.5.2:** *Newest vertex bisection and additional steps to create an interior node (red).* The refinement edges are shown with green dashed lines.

for an illustrative example. In two dimensions, the newest vertex bisection can be extended to satisfy this property as follows. As shown in Figure 5.5.2, we first mark the appropriate elements for *two* bisections and refine them. This will already create a node at the midpoint of every interface of the original triangle. Second, we refine the two “grandchild” elements (marked with “1” in Figure 5.5.2) another time. This generates an interior node at the center of the original triangle. As usual, the surrounding elements have to be refined appropriately in order to maintain mesh conformity. For the extension in three spatial dimensions, we refer to [82].

We remark that in our numerical experiments we use the newest vertex bisection but do *not* use the additional steps that generate the interior node.

## 5.6 Adaptivity beyond the macro and micro mesh size

In an adaptive multiscale method, ideally not only the macro and micro mesh size should be chosen adaptively, but also the underlying model should be adapted accordingly. Here, we briefly illustrate the problems of adaptive bridging of scales and of the adaptive choice of the sampling domain size. We further mention the recent effort by Oden et al. regarding the adaptive control of models for multiphysics problems and refer the reader to [90] and references therein.

### Adaptive bridging.

In the adaptive FE-HMM we will encounter an issue that is specific to multiscale modeling. As the macro mesh is refined further and further, the size  $H_K$  of some macro elements  $K$  can become smaller than the size of the corresponding sampling domains  $K_{\delta_\ell}$  “within” the macro element. We illustrate this behavior in Figure 5.6.1. We furthermore see two other problems: parts of some sampling domains  $K_{\delta_\ell}$  are outside of the corresponding element  $K$  (see Figure

(c)), and some sampling domains may even be situated outside of the domain  $\Omega$ , as we can see in Figure (d). In some cases (such as periodic tensors of the form  $a^\varepsilon(x) = a\left(\frac{x}{\varepsilon}\right)$  and a sampling domain size  $\delta = \varepsilon$ ) we can find a physical interpretation, which still justifies the use of the FE-HMM. However, in most cases it is difficult to justify – from a physical and computational point of view – a sampling domain exceeding the macro element. We should therefore not only adaptively refine the mesh, but also adaptively switch the underlying physical model and therefore the modeling equations. For example, we could use homogenization for large macro elements and the full finescale equation for sufficiently small macro elements. A crucial point in model adaptivity is to find matching boundary conditions between the homogenized and the fine scale problems.

We mention for example the results by Larsson and Runesson [75] on the adaptive switching between a solver based on homogenization and a finescale solver (in the context of the VMM, see Section 2.3), which they call “adaptive bridging” [75]. There, they aim at bridging the scales “seamlessly”. We will not consider the problem of model adaptivity in this thesis.

**Adaptive choice of sampling domain size.**

As in general the periodicity  $\varepsilon$  is not known, the choice of the size of the micro sampling domains  $K_{\delta_\varepsilon}$  is not straightforward and ideally would be chosen adaptively. In the context of the variational multiscale method (see Section 2.3), Larson and Målqvist [74, 73] proposed an adaptive method, which selects the size of so-called patches for fine-scale problems based on *a posteriori* estimates. The selection of the patch size in context of the VMM roughly corresponds to the selection of the sampling domain size (when using oversampling techniques) of the micro problems of the FE-HMM.

We do not treat with the problem of adaptively determining the sampling domain size throughout this thesis, but mention that such kind of adaptivity could be coupled with our adaptive choice of the multiscale mesh.

**5.7 The use higher-order polynomials in non-smooth domains**

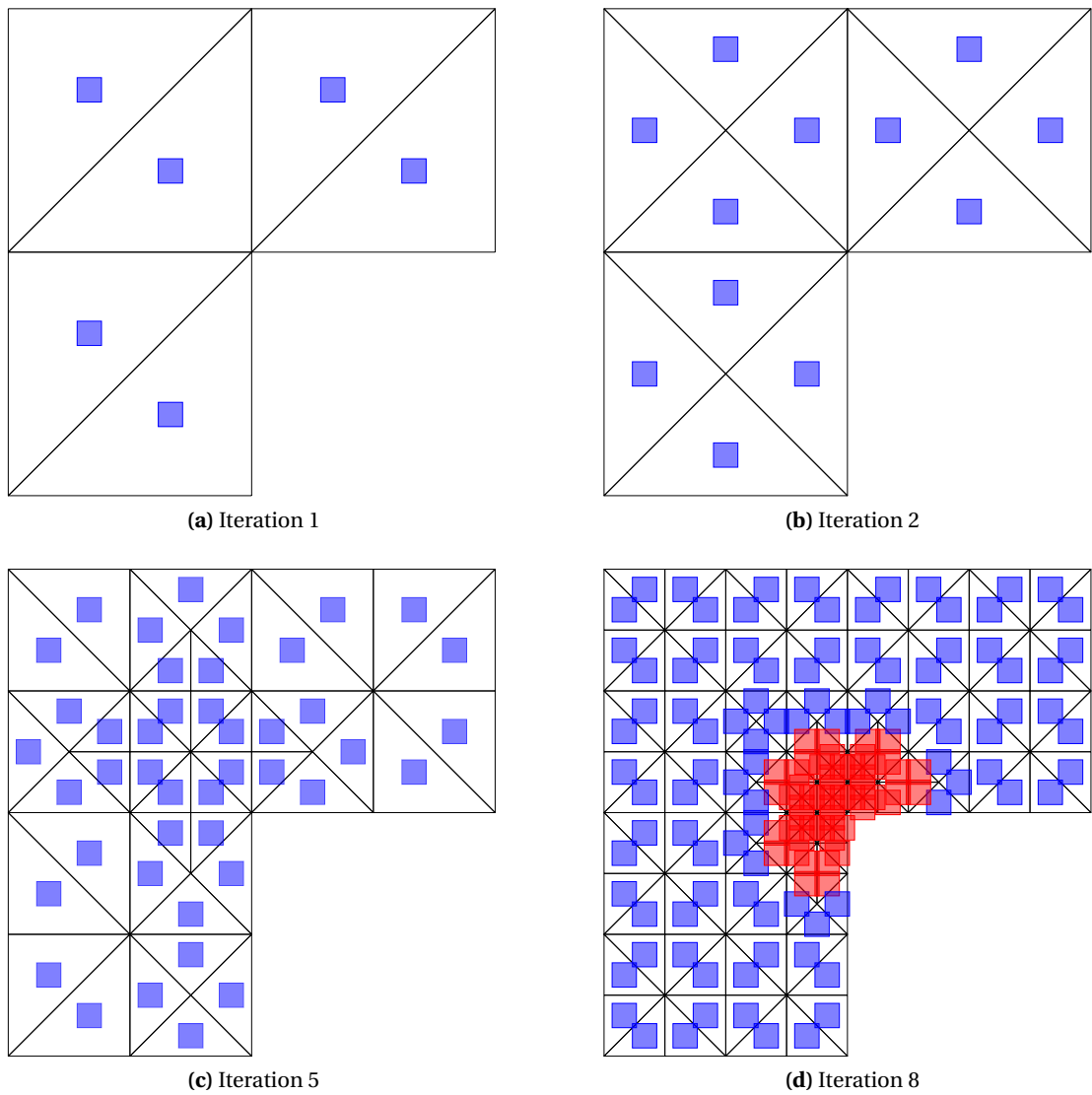
In many physical problems, we have non-smooth domains, such as domains with a reentrant corner. In what follows, we illustrate how non-smooth domains affect the regularity of the solution and we discuss the benefit of using high order FE for such problems.

For simplicity let us consider the following model problem (see [77, 105] for a more general case)

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \subset \mathbb{R}^2, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{5.7.1}$$

If  $\partial\Omega$  is smooth (i.e. there exists a parameter representation with  $C^\infty$  functions) then the classical elliptical *shift theorem* (see [18]) holds, which reads

$$f \in H^{k-1}(\Omega) \text{ implies } u \in H^{k+1}(\Omega), \quad k = 0, 1, 2, \dots$$



**Figure 5.6.1:** Macro mesh with corresponding micro sampling domains. Blue sampling domains indicate macro elements where  $|K| > |K_\delta|$ , red sampling domains indicate macro elements where  $|K| < |K_\delta|$ . From iteration 5 on we see that for some elements the sampling domain  $K_\delta$  is no longer a subset of the domain  $K$ . From iteration 7 on we have  $|K| < |K_\delta|$  for some elements; furthermore some sampling domains  $K_\delta$  lie outside of the domain  $\Omega$ . At this point, the underlying model should be switched adaptively.

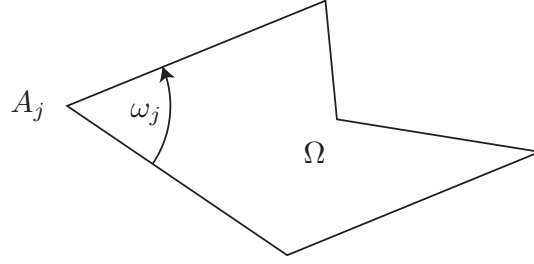


Figure 5.7.1: Polygonal domain as considered in Section 5.7.

In the presence of corners, the previous results are no longer valid, for example in the case of a polygonal domain  $\Omega$ , the above results only hold in general for  $k \in [0, k_0)$ , where  $k_0 = \pi / \max_j \omega_j$  and the interior angle  $\omega_j \in (0, 2\pi)$ .

Let us introduce polar coordinates  $(r_j, \varphi_j)$  for every vertex  $A_j$  of the polygon  $\Omega$ , see Figure 5.7.1. We define singular functions  $S_{lj}$  by

$$S_{lj}(r_j, \varphi_j) = \begin{cases} r_j^{l\pi/\omega_j} \sin\left(\frac{l\pi}{\omega_j} \varphi_j\right) & \text{if } l\pi/\omega_j \notin \mathbb{N} \\ r_j^{l\pi/\omega_j} \left( \ln r_j \sin\left(\frac{l\pi}{\omega_j} \varphi_j\right) + \varphi_j \cos\left(\frac{l\pi}{\omega_j} \varphi_j\right) \right) & \text{if } l\pi/\omega_j \in \mathbb{N}. \end{cases}$$

Then we can state a modified shift theorem as follows

**Proposition 31.** (See [77, Proposition 1.4.1]). *Let  $\Omega$  be a polygon with vertices  $A_j$ ,  $j = 1, \dots, J$  and corresponding interior angles  $\omega_j \in (0, 2\pi)$ . Let  $f \in H^k(\Omega)$ . Then for some  $a_{lj}(f) \in \mathbb{R}$  and  $u_0 \in H^{k+1}(\Omega)$  we can decompose the solution of (5.7.1) as*

$$u = \sum_{j=1}^J \sum_{l \in \mathbb{N}, l\pi/\omega_j < k} a_{lj}(f) S_{lj} + u_0,$$

where the coefficients  $a_{lj}(\cdot)$ , the so-called stress intensity factors, are linear functionals and where for each  $k$  there exists a constant  $C_k > 0$  independent of  $f$  such that

$$\sum_{j=1}^J \sum_{l \in \mathbb{N}, l\pi/\omega_j < k} |a_{lj}(f)| + \|u_0\|_{H^{k+1}(\Omega)} \leq C_k \|f\|_{H^{k-1}(\Omega)}.$$

Similar decompositions hold for general elliptic boundary value problems.

As the singular functions  $S_{lj}(r_j, \varphi_j)$  have a singularity only at  $r = 0$  while they are smooth for  $r > 0$ , the solution of problem (5.7.1) will be smooth in the interior of  $\Omega$  as long as the right hand side  $f$  is smooth. Therefore the singular behavior of the solution is restricted to the corners  $A_j$ . This justifies why when using higher order adaptive FEM on e.g. a domain with reentrant corners we may get a better overall convergence rate than we would expect considering only the (reduced) regularity of the solution. This will be confirmed in the numerical experiments in Section 7.6.1. A method very well suited for these kinds of problems is the

*hp*-FEM, which exploits the restriction of the singular behavior to corners by choosing a low polynomial degree and highly refined elements in the corner and a coarse elements with a high polynomial degree where the solution is smooth, see [77, 105].

## 5.8 Summary

In this Chapter, we have reviewed single-scale residual-based and duality-based *a posteriori* error estimates and the corresponding adaptive FEMs. We further pointed out challenges that come with the dual-weighted residual method, such as the approximation of the dual solution in a higher order FE space and its limited reliability, the choice of the function  $\psi^H$  or regularization of the functional. We briefly presented marking schemes and mesh refinement techniques. We illustrated adaptive techniques beyond mesh refinement, which aim at adaptivity of the model of multiscale problems. Further, the use of higher order FE in non-smooth domains was motivated.

In the Chapters 6, 7 and 8, we will introduce residual-based and goal-oriented adaptive multiscale FEMs, respectively. There, we will use techniques that allow us to relate the multi-scale components of the FE-HMM to classical, single-scale components. This relation will allow us to largely follow the single-scale analysis and techniques that we presented in this Chapter.



## 6 Adaptive FE-HMM

When combining multiscale finite element methods with adaptivity, new challenges arise. These challenges involve the quantification of the interplay between the micro and macro solutions. We seek for the design and analysis of an adaptive strategy that takes into account both micro and macro errors. The main questions are: How does the solution in the microscopic problem influence which macroscopic element we should refine? If we refine a macroscopic element, *how* do we need to refine the microscopic mesh of the corresponding sampling domain? How does the accuracy of the estimator depend on the error we make in the macro and micro solutions?

In this Chapter we present a detailed *a posteriori* error analysis for the FE-HMM and a suitable efficient and reliable adaptive FE-HMM algorithm. The analysis will give us answers to the above questions.

In order to analyze the multiscale FE-HMM, we relate its special multiscale components, such as the multiscale bilinear form, the multiscale fluxes and jumps to their equivalent classical, single-scale FEM counterparts. Using the equivalent counterparts enables us to follow classical residual-based adaptive FEM techniques in order to derive upper and lower error bounds. The relation between multiscale components and their classical counterparts has first been introduced in the context of the discontinuous Galerkin FE-HMM [5]. We will find that adaptive multiscale methods are much more efficient than non-adaptive ones. The huge gain in computational efficiency for multiscale methods is possible, because on the micro-level expensive micro computations can be avoided or re-used through the adaptive cycle. In this Chapter we restrict ourselves to piecewise linear FE for the macro and the micro problems.

In Section 6.1 we introduce the model problem. In Section 6.2 we state the main results consisting of upper and lower *a posteriori* bounds for the FE-HMM in the energy norm. The full analysis is presented in Section 6.3. Various numerical experiments that quantify both efficiency and reliability of our bounds are given in Section 6.4. We conclude in Section 6.5 with a detailed comparison to the *a posteriori* estimates obtained by Ohlberger in the two-scale norm [92].

The results presented in this Chapter have been announced in [10] (without proofs) and

published in [12]. Sections 6.2 to 6.4 of this Chapter are reprinted from [12, Sections 4-7] with permission from Elsevier. The references and citations were updated in order to fit into the framework of this thesis.

## 6.1 Model Problem

Given  $f \in L^2(\Omega)$  and a domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , we consider the second-order elliptic multiscale equation

$$\begin{aligned} -\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) &= f \quad \text{in } \Omega, \\ u^\varepsilon &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{6.1.1}$$

where  $a^\varepsilon$  is symmetric, satisfies  $a^\varepsilon(x) \in (L^\infty(\Omega))^{d \times d}$  and is uniformly elliptic and bounded. We consider in what follows only homogeneous Dirichlet boundary conditions, but emphasize that the *a posteriori* estimates can also be derived for more general boundary conditions (such as Neumann or Robin) following the lines of the results presented in this Chapter.

## 6.2 Main results

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

For simplicity, piecewise linear *macro* FE (simplicial elements) will be used. As a consequence, we will use a quadrature formula with  $\mathcal{L} = 1$  and integration node  $x_{K_\ell} = x_K$  localized at the barycenter of the macro element  $K$ , and a weight  $\omega_{K_\ell} = \omega_K = |K|$  (see Chapter 2). Moreover, we choose piecewise linear or bilinear *micro* FE, i.e.,  $q = 1$  in (2.4.9) and  $\mathcal{R}^1 = \mathcal{P}^1$  or  $\mathcal{Q}^1$ . The *a posteriori* error analysis will be generalized for higher order FEs in Chapter 7.

Let  $\mathcal{T}_H$  denote a conformal mesh and let  $\mathcal{E}_H$  be the set of interfaces. We label the two elements sharing an interface  $e \in \mathcal{E}_H$  as  $K^+$  and  $K^-$ . We consider the micro functions  $u_{K^+}^h$  and  $u_{K^-}^h$ , solutions of the micro problems (2.4.7), which correspond to the two sampling domains  $K_\delta^+$  and  $K_\delta^-$  of the elements  $K^+$  and  $K^-$ , respectively. These micro functions are constrained by the macro solution  $u^H \in V^1(\Omega, \mathcal{T}_H)$  of problem (2.4.12). We then introduce the following jump of *multiscale fluxes*

$$\overline{[a^\varepsilon(x) \nabla u^h]}_e := \begin{cases} \left( \frac{1}{|K_\delta^+|} \int_{K_\delta^+} a^\varepsilon(x) \nabla u_{K^+}^h dx - \frac{1}{|K_\delta^-|} \int_{K_\delta^-} a^\varepsilon(x) \nabla u_{K^-}^h dx \right) \cdot n_e & \text{for } e \notin \partial\Omega, \\ 0 & \text{for } e \subset \partial\Omega, \end{cases} \tag{6.2.1}$$

where the unit outward normal  $n_e$  is chosen to be  $n_e = n^+$ . We omit the index  $K_\delta$  for the micro



solutions  $u^h$  in  $\overline{[a^\varepsilon(x) \nabla u^h]}_e$  as the jump over  $e$  involves two sampling domains in adjacent elements. The multiscale fluxes are the building block in the derivation of our estimates and were first introduced in the context of multiscale discontinuous Galerkin methods [5, 7].

For each vector  $\mathbf{e}_i$ , where  $\{\mathbf{e}_i\}_{i=1}^d$  is the canonical basis in  $\mathbb{R}^d$ , we consider  $\psi_{K_\delta}^{i,h} \in S_h^1(K_\delta, \mathcal{T}_h)$ , the solution of the problem

$$\int_{K_\delta} a^\varepsilon(x) \nabla \psi_{K_\delta}^{i,h} \cdot \nabla z^h dx = - \int_{K_\delta} a^\varepsilon(x) \mathbf{e}_i \cdot \nabla z^h dx \quad \forall z^h \in S_h^1(K_\delta, \mathcal{T}_h). \quad (6.2.2)$$

It will be convenient for our analysis to introduce a numerically homogenized tensor (see [7]). We define this tensor  $a_K^0$ , constant on each macro element  $K$ , by

$$a_K^0 = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \left( I + J_{\psi_{K_\delta}^h(x)}^T \right) dx, \quad (6.2.3)$$

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

**Definition 32.** Let  $f^H$  be a piecewise constant approximation of  $f$ . Then the *local error indicator*  $\eta_H(K)$  on an element  $K$  is defined by

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

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

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

where  $a^0(x)$  is the unknown homogenized tensor of problem (2.4.2).

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

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

**Theorem 33** (*A posteriori* upper bound). *There exists a constant  $C > 0$  depending only on the shape regularity constant  $\gamma$ , the coercivity and continuity bound (2.1.2), the dimension  $d$  and the domain  $\Omega$  such that*

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

The next result gives an *a posteriori* lower bound.

**Theorem 34** (*A posteriori* lower bound). *There exists a constant  $C > 0$  depending only on the shape regularity constant  $\gamma$ , the coercivity and continuity bound (2.1.2) and the dimension  $d$  such that*

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

where the domain  $\omega_K$  consists of all elements sharing at least one side with  $K$ .

**Micro-macro refinement** The above two theorems for the *a posteriori* lower and upper bounds do not require any structure assumption on the oscillating tensor (such as periodicity) and only minimal assumptions on regularity. As we assume singularities in the macro scale, we do not consider explicit *a posteriori* estimates for the micro-problem (2.4.7). As the micro sampling domains have simple geometries (typically squares or cubes), singularities can only arise due to singularities in the microscale of the conductivity tensor. In that case, standard *a posteriori* methods could be used to refine the micro meshes in a non-uniform way.

We emphasize that the error indicator  $\eta_H$  does depend on the micro solutions and hence on the micro mesh. Thus, a criterion is needed to determine an appropriate size of the micro mesh as we refine the macro mesh through our adaptive procedure. Such a criterion can be deduced from the following theorem.

**Theorem 35** (Micro-macro refinement coupling). *Assume that (6.3.16) and (6.3.17) hold. Assume further that the cell problem (2.4.7) is solved with periodic boundary conditions if  $a^\varepsilon$  is periodic and  $\delta/\varepsilon \in \mathbb{N}$  and solved with Dirichlet boundary conditions otherwise. Then*

$$\sup_{x \in K} \|a_K^0 - a^0(x)\|_F \leq C \left( H_K + \left( \frac{h}{\varepsilon} \right)^2 \right) + err_{mod}, \quad (6.2.4)$$

where  $C$  is independent of  $H, h, \varepsilon$ , and  $err_{mod}$  is independent of  $H, h$ .

*Remark 36.* From estimate (6.2.4) we deduce that in order to minimize the error originating from the micro FEM, we have to refine the micro mesh in each sampling domain  $K_\delta$  as  $\frac{h}{\varepsilon} \propto \sqrt{H_K}$ . Here  $H_K$  is the size of the macro element  $K$  of the mesh  $\mathcal{T}_H$  obtained by the Algorithm 37 described below.

**Comparison with single scale results.** Our upper and lower bounds stated above are consistent with the usual (single-scale) residual based *a posteriori* estimates. Indeed, suppose  $a^\varepsilon = a\left(\frac{x}{\varepsilon}\right)$ , that exact micro functions are used in (2.4.7) with periodic boundary conditions and that  $\delta = \varepsilon$ . Then  $\|a_K^0 - a^0\|_F \equiv 0$  (see Section 6.3.5.2) and we recover the usual residual based indicator and data estimator [113, 33].

### 6.2.1 Algorithm

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

Solve  $\rightarrow$  Estimate  $\rightarrow$  Mark  $\rightarrow$  Refine

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

**Algorithm 37** (Adaptive FE-HMM).

**Solve.** For the macro and micro meshes obtained by REFINE, compute the micro solutions (only for the refined macro elements) and the macro solution  $u^H$  of (2.4.12) and compute and store the multiscale jumps  $\llbracket a^\varepsilon(x) \nabla \varphi_{i,K}^h \rrbracket_e$  (based on the FE basis functions) for the refined elements during the macro assembly process.

**Estimate.** Reconstruct the full multiscale jumps  $\overline{\llbracket a^\varepsilon(x) \nabla u^h \rrbracket_e}$  using the macro solution  $u^H$  of SOLVE and estimate the error by computing the indicators  $\eta_H(K)$  for all  $K \in \mathcal{T}_H$ .

**Mark.** Identify a subset  $\widetilde{\mathcal{T}}_H$  of  $\mathcal{T}_H$  based on the indicators  $\eta_H(K)$  following Dörfler's bulk-chasing strategy (Marking Strategy E, see [113, Chap. 4.1] or Section 5.4).

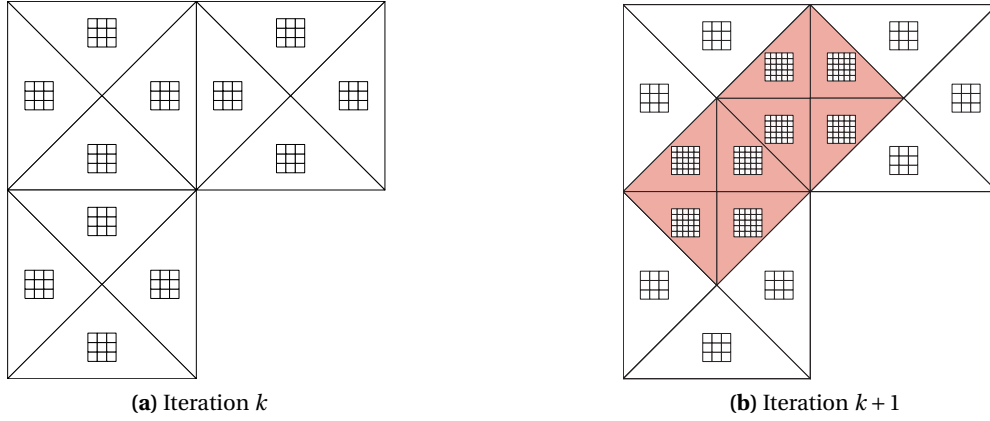
**Refine.** Refine the elements in the subset  $\widetilde{\mathcal{T}}_H$  and some neighboring elements in order to preserve mesh conformity. Update the micro mesh in the sampling domains of the refined macro elements according to  $h_{K_\delta} \propto \sqrt{H_K} \varepsilon$  (see Remark 36).

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

**Remark 38. Carry-over of information.** In contrast to standard (single-scale) adaptive FE methods, in the adaptive FE-HMM most of the computation time is used for solving the micro problems to obtain the entries of the macro stiffness matrix. A fundamental feature of an efficient implementation of an adaptive FE-HMM should thus be the carryover of reusable micro data from one iteration to the next.

In particular, this means that for every element we store the contributions  $A_{ij} = B(\varphi_{i,K}^H, \varphi_{j,K}^H)$  ( $\varphi_{i,K}^H$  are the macro basis functions) of the micro problems to the macro stiffness matrix and the components of the corresponding multiscale flux  $\overline{\llbracket a^\varepsilon(x) \nabla \varphi_{i,K}^h \rrbracket_e}$ . This is done with a small memory overhead in every iteration and we compute new data  $A_{ij}$  and  $\llbracket a^\varepsilon(x) \nabla \varphi_{i,K}^h \rrbracket_e$  only for those elements which are marked for refinement and carry-over the existing data  $A_{ij}$  and  $\overline{\llbracket a^\varepsilon(x) \nabla \varphi_{i,K}^h \rrbracket_e}$  for the remaining, unrefined macro triangles to the next iteration (see Figure 6.2.1 for an illustration). In the numerical experiments in Section 6.4 we illustrate the amount of work that can be saved following this strategy.

**Computation of the multiscale flux.** Let  $\{\varphi_{i,K}^H\}$  be the basis functions of  $V^1(\Omega, \mathcal{T}_H)$  and let  $\varphi_{i,K}^h$  be the micro solution of (2.4.7) constrained by  $\varphi_{i,K}^H$ , i.e.,  $(\varphi_{i,K}^h - \varphi_{i,K}^H) \in S^1(K_\delta, \mathcal{T}_H)$ , where



**Figure 6.2.1:** Adaptive mesh shown for two iterations. In Figure b), for all macro elements shown in white, the multiscale flux  $\llbracket a^\varepsilon(x) \nabla \varphi_{i,K}^h \rrbracket_e$  and the contributions  $A_{ij}$  to the macro stiffness matrix can be carried over to the next iteration and re-used, whereas for all macro elements shown in red, new solutions of the micro problems and corresponding multiscale fluxes must be computed.

$K_\delta \subset K$ . In SOLVE, we compute for every element  $K$

$$\varrho_{i,K}^h := \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla \varphi_{i,K}^h dx$$

right after computing the solutions  $\varphi_{i,K}^h$  of the micro problem (2.4.7) and store the three corresponding  $\varrho_{i,K}^h$  for later use. (In the case of two dimensions when using macro triangles this represents a  $2 \times 3$  matrix in each macro element  $K$  corresponding to the two components of  $\varrho_{i,K}^h$  for each macro basis function  $\varphi_{i,K}^H$ ,  $i = 1, \dots, 3$ ). The advantage of computing this quantity in the SOLVE instead of the REFINE step is that we do not need to store the full micro solution of each macro element (if we are not interested in a reconstructed full solution, this significantly reduces the memory requirement).

Denote by  $u_K^H(x) = \sum_{i=1}^3 \alpha_i \varphi_{i,K}^H(x)$  the representation of  $u^H$  in the element  $K$  with respect to the nodal basis. As the micro solution  $u_{i,K}^h(x)$  corresponding to  $u_K^H$  is linear with respect to  $u_K^H(x)$ , we can reconstruct  $u_K^h(x) = \sum_{i=1}^3 \alpha_i \varphi_{i,K}^h(x) = \sum_{i=1}^3 u^H(x_i) \varphi_{i,K}^h(x)$  where  $x_i$  are the nodes of the macro triangle  $K$ . In the ESTIMATE step, where the macro solution  $u^H$  is known from the previous SOLVE step, we reconstruct

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

and assemble  $\llbracket a^\varepsilon(x) \nabla u_{i,K}^h \rrbracket_e$  using the previously stored information. As can be seen from the aforementioned procedure, the computation of  $\llbracket a^\varepsilon(x) \nabla u_{i,K}^h \rrbracket_e$  can be done with minimal memory overhead and very small extra computational cost.

*Remark 39 (Coarsening).* We have not considered a coarsening strategy for two reasons. First, for linear elliptic problems, a quasi-optimal mesh can usually be obtained without the need

of coarsening strategies (see [104, Chapter 1.5.3]). Second, due to the macro-micro coupling, new extra micro problems must be solved when elements are coarsened, whereas information can be re-used if we omit to coarsen the macro elements.

## 6.3 Proof of the main results

### 6.3.1 Interpolation, trace and inverse estimates

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

**Clément interpolation operator** (see [46]).

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

$$\|v - I^H v\|_{L^2(K)} \leq CH_K \|\nabla v\|_{L^2(N(K))} \quad (6.3.1)$$

and

$$\|\nabla(v - I^H v)\|_{L^2(K)} \leq C \|\nabla v\|_{L^2(N(K))}, \quad (6.3.2)$$

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

**Trace inequality** (see [18, Thm. 3.10]).

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

$$\|v\|_{L^2(e)} \leq CH_e^{1/2} \|\nabla v\|_{L^2(K_e)} + CH_e^{-1/2} \|v\|_{L^2(K_e)}. \quad (6.3.3)$$

**Inverse inequality** (see for example [44]).

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

$$\|\nabla v^H\|_{L^2(K)} \leq CH_K^{-1} \|v^H\|_{L^2(K)}. \quad (6.3.4)$$

*Remark 40.* The combination of the Clément interpolation estimates and the trace inequality yields for an element  $K_e$  with side  $e \in \mathcal{E}_H$

$$\begin{aligned} \|v - I^H v\|_{L^2(e)} &\leq CH_e^{1/2} \|\nabla(v - I^H v)\|_{L^2(K_e)} + CH_e^{-1/2} \|v - I^H v\|_{L^2(K_e)} \\ &\leq CH_K^{1/2} \|\nabla v\|_{L^2(N(K_e))}. \end{aligned} \quad (6.3.5)$$

### 6.3.2 Error representation formula

The representation formula (6.3.7) in Lemma 42 is the central tool to derive our *a posteriori* bounds, as it allows to link the bilinear form for the homogenized solution with the FE-HMM. We first prove Lemma 41, needed to derive the representation formula.

**Lemma 41.** *Let  $v_K^h, w_K^h$  be the solutions of (2.4.7) constrained by  $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$  with boundary conditions given by (2.4.10) or (2.4.11). Then*

$$\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla v_K^h \cdot \nabla w_K^h dx = \frac{1}{|K|} \int_K a_K^0 \nabla v^H \cdot \nabla w^H dx.$$

*Proof.* The proof is similar to the proof of (A.1) in [14, App. A] or formula (63) in [6]. (For the convenience of the reader we will recall it). While in the aforementioned results, a specific structure of the tensor  $a^\varepsilon$  was used, we prove the result without any assumption on  $a^\varepsilon$  (except for the positivity and ellipticity).

First, we notice that the (unique) solution of (2.4.7) can be written as

$$v_K^h(x) = v^H(x) + \sum_{i=1}^d \psi_{K_\delta}^{i,h}(x) \frac{\partial v^H(x)}{\partial x_i}, \quad (6.3.6)$$

where  $\psi_{K_\delta}^{i,h} \in S^1(K_\delta, \mathcal{T}_h)$ ,  $i = 1, \dots, d$  are the solutions of (6.2.2).

From this we deduce

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

where we used that  $\frac{\partial w^H(x)}{\partial x_i}$  is constant,  $\psi_{K_\delta}^{i,h} \in S^1(K_\delta, \mathcal{T}_h)$  and equation (2.4.7). Recalling the definition (6.2.3) of  $a_K^0$  we obtain

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

where we used again that  $\nabla v^H$  and  $\nabla w^H$  are constant. □

We define the error as  $e^H := u^0 - u^H$  where  $u^0$  is the homogenized solution of (2.4.2) and  $u^H$  is the FE-HMM solution of problem (2.4.12). We shall now obtain an *error representation formula* which is crucial for the derivation of the *a posteriori* bounds.

**Lemma 42** (Error representation formula). *For all  $v \in H_0^1(\Omega)$ , we have*

$$B_0(e^H, v) = \int_{\Omega} f v dx - \sum_{e \in \mathcal{E}_H} \int_e \overline{[a^\varepsilon(x) \nabla u^h]}_e v ds + \sum_{K \in \mathcal{T}_H} \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla v dx \quad (6.3.7)$$

where  $u^H$  is the solution of (2.4.12) and  $u_K^h$  are the corresponding micro solutions (2.4.7) and where the multiscale jump  $\overline{[a^\varepsilon(x) \nabla u^h]}_e$  is defined in (6.2.1).

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

$$\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla u_K^h dx = a_K^0 \nabla u_K^H \quad (6.3.8)$$

which is obtained by using the expansion (6.3.6) and similar arguments as used for the proof of Lemma 41. Second, we prove that for all  $v \in H_0^1(\Omega)$  we have

$$\sum_{K \in \mathcal{T}_H} \int_K a_K^0 \nabla u^H \cdot \nabla v dx = \sum_{e \in \mathcal{E}_H} \int_e \overline{[a^\varepsilon(x) \nabla u^h]}_e v ds. \quad (6.3.9)$$

Integration by parts and the use of (6.3.8) gives

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

where we used the definition (6.2.1) of the multiscale flux. Finally, we obtain the error representation formula

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

□

## 6.3.3 Upper bound (Proof of Theorem 33)

To proceed with the proof of Theorem 33, we consider the error representation formula (6.3.7) and choose the test function  $v := e^H$ . We recall that  $I^H$  denotes the Clément interpolation operator (see (6.3.1),(6.3.2)) and  $f^H$  denotes a piecewise constant approximation of  $f$  over  $\mathcal{T}_H$ . By noting that

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

we obtain

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

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

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

With the help of the Clément interpolation estimates (6.3.1) and (6.3.5) we deduce that

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



The finite overlapping property of the neighborhoods  $N(K)$  allows us to estimate

$$\sum_{K \in \mathcal{T}_H} \|\nabla e^H\|_{L^2(N(K))}^2 \leq C \sum_{K \in \mathcal{T}_H} \|\nabla e^H\|_{L^2(K)}^2,$$

where  $C$  depends only on the shape regularity of the triangulation and the dimension  $d$ . Using the coercivity of  $B_0(\cdot, \cdot)$  and the triangle inequality yields

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

Using Poincaré inequality leads to

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

as stated in Theorem 33.

### 6.3.4 Lower bound (Proof of Theorem 34)

To derive the lower bound, we will use a construction involving bubble functions in a space  $\tilde{V}_H^1 \supset V_H^1$  which is defined over a refinement  $\tilde{\mathcal{T}}_H$  of  $\mathcal{T}_H$ . We assume that the refinement  $\tilde{V}_H^1$  is chosen such that every  $K \in \mathcal{T}_H$  has an interior node  $\tilde{x}_K \in K \setminus \partial K$  in  $\tilde{\mathcal{T}}_H$  and every edge  $e$  of  $\mathcal{T}_H$  not on the boundary  $\partial\Omega$  has an interior node in  $\tilde{\mathcal{T}}_H$ . We emphasize once again that the use of the representation formula (6.3.7) allows largely to follow the classical construction of [113]. In what follows we estimate successively interior and jump residuals.

#### 6.3.4.1 Interior residual

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

We choose  $v := \psi_K f^H \in H_0^1(\Omega)$  as a test function in the error representation formula (6.3.7) and obtain

$$\begin{aligned} \int_K f^H (\psi_K f^H) dx = & B_0(e^H, \psi_K f^H) - \int_K (f - f^H) \psi_K f^H dx \\ & - \int_K (a_K^0 - a^0(x)) \nabla u^H \cdot \nabla (\psi_K f^H) dx. \end{aligned}$$

Using the equivalence of norms on a finite-dimensional space we have (see [20, Theorem 2.2] for details)

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

Furthermore, the continuity of  $B_0(\cdot, \cdot)$ , the Cauchy-Schwarz inequality and the inverse in-

equality (6.3.4) give

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

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

$$\begin{aligned} H_K^2 \|f^H\|_{L^2(K)}^2 &\leq C \left( \|\nabla e^H\|_{L^2(K)}^2 + H_K^2 \|f - f^H\|_{L^2(K)}^2 \right. \\ &\quad \left. + \|(a_K^0 - a^0(x)) \nabla u^H\|_{L^2(K)}^2 \right). \end{aligned} \quad (6.3.10)$$

### 6.3.4.2 Jump residual

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

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

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

$$\begin{aligned} \int_e [[\overline{a^\varepsilon(x) \nabla u^h}]]_e \psi_e ds &= \overline{[[a^\varepsilon(x) \nabla u^h]]_e} \int_e \psi_e ds \\ &\geq C|e|^{-1/2} \left\| [[\overline{a^\varepsilon(x) \nabla u^h}]]_e \right\|_{L^2(e)} H_e^{d-1} \\ &\geq CH_e^{\frac{d-1}{2}} \left\| [[\overline{a^\varepsilon(x) \nabla u^h}]]_e \right\|_{L^2(e)}. \end{aligned} \quad (6.3.11)$$

Next, we set  $v = \psi_e$  in the representation formula (6.3.7) (recall that  $v \equiv 0$  on  $\Omega \setminus (K_1 \cup K_2)$ ) and obtain

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

where we used  $\|\nabla\psi_e\|_{L^2(K)} \leq C H_K^{-1} \|\psi_e\|_{L^2(K)} \leq C(H_K)^{\frac{d-2}{2}}$ , which follows from the inverse inequality (6.3.4). The inequality (6.3.11) and the above estimate for  $\|\psi_e\|_{L^2(K)}$  yield

$$H_e \left\| \overline{[a^\varepsilon(x) \nabla u^h]}_e \right\|_{L^2(e)}^2 \leq C \sum_{K_1, K_2} \left( H_{K_i}^2 \|f^H\|_{L^2(K_i)}^2 + \|\nabla e^H\|_{L^2(K_i)}^2 + H_{K_i}^2 \|f - f^H\|_{L^2(K_i)}^2 + \left\| (a_{K_i}^0 - a^0(x)) \nabla u^H \right\|_{L^2(K_i)}^2 \right). \quad (6.3.12)$$

### 6.3.4.3 Combining interior and jump residuals

We use the interior residual (6.3.10) to eliminate  $\|f^H\|_{L^2(K)}$  from the jump residual

$$H_e \left\| \overline{[a^\varepsilon(x) \nabla u^h]}_e \right\|_{L^2(e)}^2 \leq C \left( \|\nabla e^H\|_{L^2(\omega_e)}^2 + H_{\omega_e}^2 \|f - f^H\|_{L^2(\omega_e)}^2 + \left\| (a_K^0 - a^0(x)) \nabla u^H \right\|_{L^2(\omega_e)}^2 \right), \quad (6.3.13)$$

where  $H_{\omega_e} = \max_{i=1,2} H_i$  and  $\omega_e = K_1 \cup K_2$ . Adding the interior residual (6.3.10) to (6.3.13) leads to the desired lower bound

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

as stated in Theorem 34.

### 6.3.5 Data approximation

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

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

The first term of the right-hand side of this equality involves the usual data approximation term. The second term quantifies the accuracy of the macro-micro algorithm and it depends on the macro and micro meshes of the macro and micro FEMs, on the structure of the oscillating tensor  $a^\varepsilon$  and on the coupling condition between micro and macro solver. We first notice that

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

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

$$\bar{a}_K^0 = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \left( I + J_{\psi_K(x)}^T \right) dx, \quad (6.3.14)$$

where  $J_{\psi_K(x)}$  is a  $d \times d$  matrix with entries  $(J_{\psi_K(x)})_{ij} = (\partial \psi_K^i) / (\partial x_j)$ . This tensor is computed similarly to the tensor  $a_K^0$  in (6.2.3) but with functions  $\psi_K^i(x)$  solving (6.2.2) in the exact Sobolev space  $W(K_\delta)$  instead of its FE approximation  $S^1(K_\delta, \mathcal{T}_h)$ . We then consider the following decomposition

$$\|a_K^0 - a^0\|_F \leq \underbrace{\|a^0 - a^0(x_K)\|_F}_{err_{mac}} + \underbrace{\|a^0(x_K) - \bar{a}_K^0\|_F}_{err_{mod}} + \underbrace{\|\bar{a}_K^0 - a_K^0\|_F}_{err_{mic}}, \quad (6.3.15)$$

where  $x_K \in K$  is the quadrature node located at its barycenter. In the above equation,  $err_{mac}$  and  $err_{mic}$  stands for macroscopic and microscopic error, respectively. The analysis of the macroscopic and microscopic errors relies on regularity assumptions on the homogenized and fine scale tensors. This will be analyzed in Section 6.3.5.1. The term denoted by  $err_{mod}$  (the modeling error) needs additional assumptions on the structure of  $a^\varepsilon$  in order to be quantified (e.g., periodicity and random stationarity). This part does not depend on the discretization parameters. As it does not depend on the specific form of the numerical method, previously derived results can be used to analyze this contribution to the error. We will examine in Section 6.3.5.2 the case of a non-uniformly periodic tensor and comment on the case of a random tensor.

**Regularity assumptions.** For the oscillating tensor  $a^\varepsilon$  we assume

$$a^\varepsilon|_K \in W^{1,\infty}(K), \forall K \in \mathcal{T}_H \text{ and } |a_{ij}^\varepsilon|_{W^{1,\infty}(K)} \leq C_K \varepsilon^{-1} \text{ for } i, j = 1, \dots, d. \quad (6.3.16)$$

In the analysis we will often use a constant  $C = \max_K C_K$  independent of  $K$ . It is clear that if (6.3.16) is valid for an initial mesh, assumed to be aligned with the possible discontinuities of  $a^\varepsilon$ , it is still valid (with the same value of  $C$ ) for every mesh obtained by refining the initial one. In view of (6.3.16) we see that  $a^\varepsilon$  is allowed to be discontinuous in different macro elements but we assume that the macro mesh (i.e., the interface between two neighboring elements) is aligned with these discontinuities. For the homogenized tensor we assume

$$a_{ij}^0 \text{ are Lipschitz continuous in } K \text{ for any } K \in \bar{\mathcal{T}}_H. \quad (6.3.17)$$

*Remark 43.* Without further knowledge about the structure of the oscillating tensor  $a^\varepsilon$ , we will impose Dirichlet boundary conditions for (2.4.7) (or (6.2.3)). Assuming (6.3.16) one can show  $|\psi_K^i|_{H^2(K_\delta)} \leq C \varepsilon^{-1} \sqrt{|K_\delta|}$ , with  $C$  independent of  $\varepsilon$ , of the quadrature points  $x_K$  and the domain  $K_\delta$  (this follows from classical  $H^2$  regularity results, see for example [72, Chap. 2.6]). If  $a^\varepsilon = a(x, x/\varepsilon) = a(x, y)$  is  $Y$ -periodic in  $y$  and  $\delta/\varepsilon \in \mathbb{N}$ , then assuming (6.3.16),  $\delta/\varepsilon \in \mathbb{N}$  and periodic boundary condition for (2.4.7) (or (6.2.3)) one can show  $|\psi_K^i|_{H^2(K_\delta)} \leq C \varepsilon^{-1} \sqrt{|K_\delta|}$  (this follows from classical regularity results for solutions of periodic boundary value problems (see [34, Chap. 3])). Notice that in the periodic case, for more regular tensors  $a^\varepsilon(x)$ , one can obtain higher order estimates  $|\psi_K^i|_{H^{q+1}(K_\delta)} \leq C \varepsilon^{-q} \sqrt{|K_\delta|}$ ,  $q \in \mathbb{N}$  (see [9]).

### 6.3.5.1 Micro and macro data approximation (Proof of Theorem 35)

We start with the macro error. Assuming (6.3.17) directly gives the following estimate

$$\sup_{x \in K} \|a^0(x) - a^0(x_K)\|_F \leq CH_K, \quad (6.3.18)$$

where  $C$  only depends on the dimension  $d$  and the Lipschitz constant. For the micro error, we follow the ideas of [2]. In the precise form of (6.3.19), the micro error estimate has been proved in [7]. Assuming that (2.1.2) and (6.3.16) hold, we have

$$\|\bar{a}_K^0 - a_K^0\|_F \leq C \left(\frac{h}{\varepsilon}\right)^2, \quad (6.3.19)$$

where  $C$  depends only on the constant in (6.3.16) and the bound (2.1.2). Combining (2.1.2), (6.3.18) and (2.4.9) we obtain

$$\sup_{x \in K} \|a^0(x) - a_K^0\|_F \leq C \left( H_K + \left(\frac{h}{\varepsilon}\right)^2 + err_{mod} \right), \quad (6.3.20)$$

where  $C$  only depends on the coercivity and continuity bound (2.1.2), the dimension  $d$ , the Lipschitz constant (6.3.16) and where  $e_{mod}$  is independent of the discretization parameter of the micro and macro FE spaces.

*Remark 44* (Micro mesh refinement). From the estimates (6.3.18) and (6.3.19), we obtain the criterion for the micro mesh refinement used in Algorithm 37 in Section 6.2.

### 6.3.5.2 Modeling error

In this Section we discuss the error term  $\|a^0(x_K) - \bar{a}_K^0\|_F$  in (6.3.15). Additional assumption on the spatial structure of the oscillating tensor  $a^\varepsilon$  is required in order to give convergence rates for the modeling error. We will consider the case of non-uniformly periodic tensor and assume

$$a^\varepsilon = a(x, x/\varepsilon) = a(x, y) \quad Y\text{-periodic in } y, \quad (6.3.21)$$

where for simplicity we set  $Y = (0, 1)^d$ . We will sometimes refer to the variables  $x$  and  $y$  as slow and fast variables, respectively. Other tensors  $a^\varepsilon$  could be considered. For example, following the results in [53] we could also consider an appropriate random field  $a^\varepsilon = a(\omega, y)$  (with invariant statistics under integer shifts). The modeling error estimates from [36] or [53, Theorem 2.1] could be used directly to estimate  $\|a^0(x_K) - \bar{a}_K^0\|_F$ . We thus see that our *a posteriori* error analysis applies to a variety of tensors and is not restricted to the periodic case.

In the periodic case, the homogenized tensor  $a^0$  can be computed explicitly (see e.g., [32]). For  $x_K \in K$  it reads

$$a_{ij}^0(x_K) = \int_Y a_{ij}(x_K, y) + \sum_{k=1}^d a_{ik}(x_K, y) \frac{\partial \chi^j(x_K, y)}{\partial y_k} dy, \quad i, j = 1, \dots, d, \quad (6.3.22)$$

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

$$\int_Y a(x_K, y) \nabla \chi^j(x_K, y) \cdot \nabla z \, dy = - \int_Y (a(x_K, y) \mathbf{e}_j) \cdot \nabla z \, dy, \quad \forall z \in W_{per}^1(Y), \quad (6.3.23)$$

where  $\mathbf{e}_j$  is the  $j$ -th basis vector of  $\mathbb{R}^d$ . We will consider separately the coupling conditions (periodic and Dirichlet) for the micro FE space. To analyze the modeling error, we can use results obtained for the FE-HMM in [53, 2, 14, 6].

**Periodic micro boundary conditions.**

We assume that  $S^1(K_\delta, \mathcal{T}_h) \subset W_{per}^1(K_\delta)$ . In this case we assume that  $\delta/\varepsilon \in \mathbb{N}^*$ , i.e., the sampling domains cover an integer number of the exact period of the tensor  $a(\cdot, x/\varepsilon)$ . We then can derive the following error estimates (see [7, Lemma 5.2]). Assume that (6.3.21) and (6.3.16) hold. Then

$$\|a^0(x_K) - \bar{a}_K^0\|_F \leq C\delta \quad (6.3.24)$$

for the modeling error, where  $C$  is independent of  $H, h$ , and  $\varepsilon$ . If the decomposition in fast and slow variable of the tensor  $a(x, x/\varepsilon)$  is explicitly known, we can slightly modify the FE-HMM macro bilinear form (2.4.6) and micro problems (2.4.7) by replacing  $a^\varepsilon$  with  $a(x_K, x/\varepsilon)$ . In this case, performing a similar modification of the tensor  $a^\varepsilon$  of (6.2.3), (6.3.14), one can show that the modeling error vanishes (see [7, Lemma 5.2] for details)

$$\|a^0(x_K) - \bar{a}_K^0\|_F = 0. \quad (6.3.25)$$

**Dirichlet micro boundary conditions.**

In the case where the exact period is not known but an estimation of the size of the periodicity is available, the idea is to embed the periodic sampling domain  $K_\varepsilon$  in a larger cube  $K_\delta$  with  $\delta > \varepsilon$ . Here we do not assume that  $\delta/\varepsilon \in \mathbb{N}$ . Artificial boundary conditions are chosen for the micro solver. Various conditions are possible and we assume  $S^1(K_\delta, \mathcal{T}_h) \subset H_0^1(K_\delta)$ , see (2.4.9). Assume that (6.3.21) and (6.3.16) hold. Then

$$\|a^0(x_K) - \bar{a}_K^0\|_F \leq C \left( \delta + \frac{\varepsilon}{\delta} \right), \quad (6.3.26)$$

where  $C$  only depends on the domain  $\Omega$  and the bound (2.1.2). This estimate can be obtained following the line of the proof of [53, Thm. 3.2].

## 6.4 Numerical experiments

We present in this section a series of numerical experiments which verify the sharpness of the theoretical *a posteriori* estimates and confirm that our adaptive scheme is both efficient (i.e., the mesh is near optimal) and effective, i.e., the solution is reliable.

We will present various elliptic problems with a two-scale, quasi 1D tensor on a square domain (Section 6.4.1), a crack problem with a highly oscillating 2D tensor and a random tensor (Section 6.4.2), and finally an L-shaped problem with a highly oscillating 2D tensor

(Section 6.4.3).

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

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

- $\#el$ , number of macro elements for the specific mesh. It reflects the number of degrees of freedom.
- $EOC(e^H)$ , experimental order of convergence.

$$EOC := d \frac{\log(e_{k-1}^H / e_k^H)}{\log(\#el(k) / \#el(k-1))},$$

where  $e_k^H$  refers to the error  $\|u_0 - u_k^H\|_{H^1(\Omega)}$  in the  $k$ th refinement step and  $d$  is the physical dimension of the problem (the scaling with  $d$  allows to get a convergence rate independent of the dimension).

- $EOC(\eta)$ , experimental order of convergence of the indicator  $\eta_H(\Omega)$ .

$$EOC(\eta) := d \frac{\log(\eta_H^{k-1} / \eta_H^k)}{\log(\#el(k) / \#el(k-1))},$$

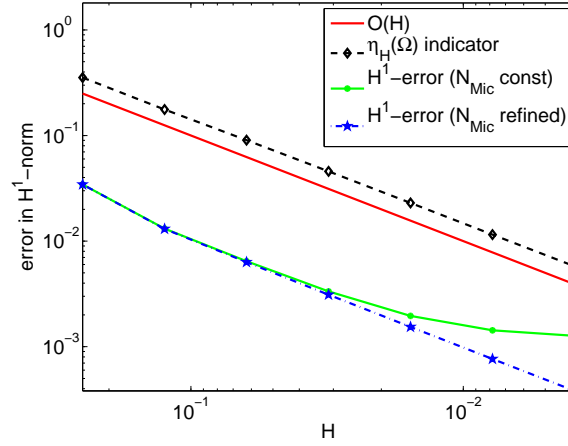
where  $\eta_H^k$  refers to the indicator in the  $k$ th refinement step and  $d$  is the physical dimension of the problem.

- $Z_e$ , the reduction factor  $Z_e := \frac{e_k^H}{e_{k-1}^H}$  of the error.
- Eff, the effectivity index  $\text{Eff} := \frac{\eta_H^k}{e_k^H}$  allows an estimate to the upper bound constant  $C$ .

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

The numerical experiments were performed using the FE-HMM code presented in [11] and the implementation of the mark and refine steps are based in part on the AFEM@Matlab code (see [43]).

*Remark 45.* In most of the following experiments, as we refine the mesh, we reach the point when  $H_K < \delta$  for some elements  $K \in \mathcal{T}_H$ , i.e., the macro element  $K$  is smaller than the sampling domain. Refining beyond this point is not computationally efficient and one should switch to



**Figure 6.4.1:** Errors and error estimate in the  $H^1$ -norm for the highly oscillatory problem described in section 6.4.1.

the fine scale solver for the whole (macro) triangle  $K$ . A precise study and analysis for such a modified algorithm will be presented elsewhere. We notice here that in the case of periodic coefficients taking a sampling domain larger than some macro elements still makes sense and allows us to check the efficiency and reliability of our estimates.

#### 6.4.1 Uniform refinement test

We consider the quasi-1D problem taken from [92],

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

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

As no singularity appears in the domain, we uniformly refine the macro triangles in every iteration step and compare the error to the indicator  $\eta_H(\Omega)$ . Parallel to the macro refinement, we refine the micro mesh according to Remark 36. The initial mesh for the micro FE spaces is chosen as  $\hat{h} = \frac{1}{8}$ .

We show in Figure 6.4.1 that the error in the  $H^1$  norm and the indicator  $\eta_H(\Omega)$  both converge to zero with rate  $\mathcal{O}(H)$  and thus match the prediction obtained by the *a priori* and *a posteriori* estimates, respectively.

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



In Table 6.4.1 we list various quantities (explained in detail at the beginning of this section) illustrating the quality of the refinement procedure.

The experimental order of convergence for both the error as well as the indicator confirm the theoretical linear convergence. The ratio Eff between them remains nearly constant, which confirms that our (here uniform) refinement is both effective and efficient; it furthermore gives an estimate for the *effectivity index*, the constant in the upper bound of Theorem 33. As the unspecific constant depends on e.g. the Clément interpolation estimates and the coercivity and continuity of the tensor, it is difficult to estimate, and may even be larger in more complex problems.

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

**Table 6.4.1:** Grid size  $H$ ,  $H^1$ -error, error indicator, experimental order of convergence for the error and the error indicator and the effectivity of the problem described in Section 6.4.1.

### 6.4.2 Crack problem

In our next experiment we consider a crack problem based on [82, Example 5.2] that exhibits a singularity in the macro domain. But unlike Example 5.2 of [82], we use here a two-dimensional, highly oscillating conductivity tensor. We consider the following problem

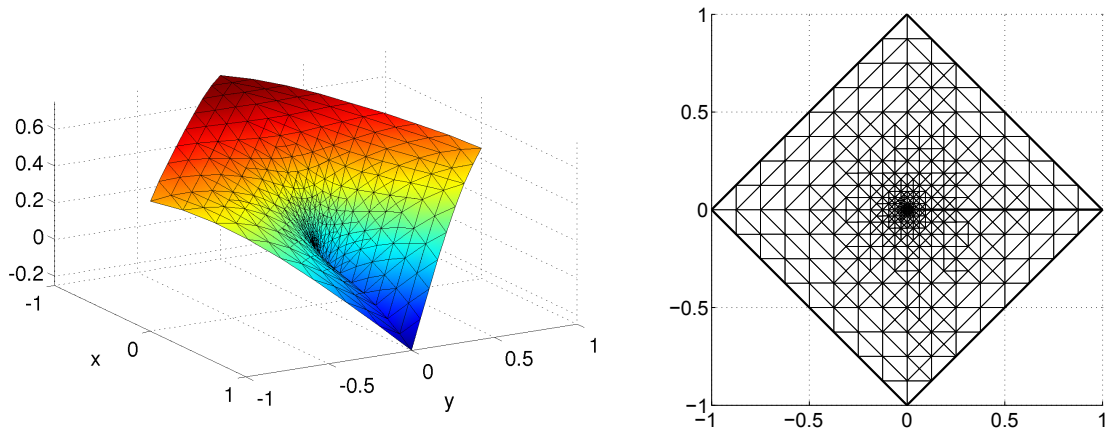
$$\begin{aligned} -\nabla \cdot (a^\varepsilon(x) \nabla u^\varepsilon) &= 1 \quad \text{in } \Omega \\ u^\varepsilon &= g_D \quad \text{on } \Gamma_D = \partial\Omega \end{aligned}$$

on a domain  $\Omega = \{|x_1| + |x_2| < 1\} \setminus \{0 \leq x_1 \leq 1, x_2 = 0\}$  with a crack along the positive  $x_1$ -axis (see Figure 6.4.2). We use the tensor

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

where we chose the coefficients of the tensor in such a way that the homogenized tensor coincides with the unit tensor  $I_2$  (see [70, Chap. 1.2]). The Dirichlet boundary conditions  $g_D = u^0$  match the exact homogenized solution  $u^0$  of the problem which is given (in polar coordinates) by

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



**Figure 6.4.2:** FE-HMM solution and mesh after 10 iterations for the crack-problem described in Section 6.4.2.

where  $x_1 = r \cos(\vartheta)$ ,  $x_2 = r \sin(\vartheta)$ . We emphasize that we use an analytically homogenizable tensor only to be able to compare our solution to the exact solution. Any other oscillating tensor could be used (see Section 6.4.2.2 for an experiment using a random tensor).

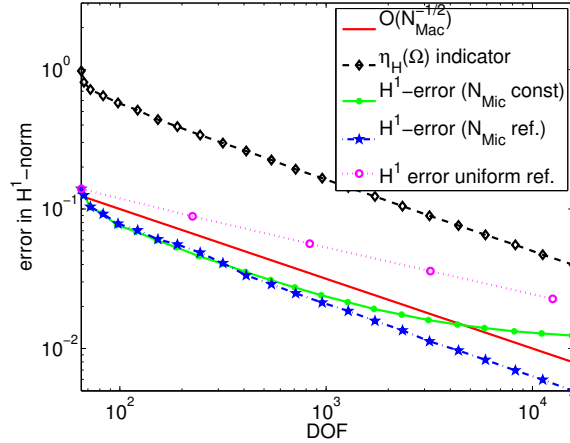
A solution of the problem is shown in Figure 6.4.2.

We choose  $\varepsilon = \delta = 10^{-3}$  and periodic micro boundary conditions. This time, we use our adaptive strategy to refine the mesh and select a total of 23 refinement steps (Dörfler’s bulk-chasing strategy is used for marking, with a parameter of  $\theta = 0.3$ , see [113, Chap. 4.1] or Section 5.4). We again use the relation  $\hat{h}_K = \sqrt{H_K}$  for our micro refinement strategy (see Algorithm 37) with an initial mesh of  $\hat{h}_K = \frac{1}{8}$ .

In Figure 6.4.3 we show the errors in the  $H^1$  norm. The rate of convergence of the error and the error indicator confirms the theoretical rate of  $\mathcal{O}(N_{mac}^{-1/d})$ , where  $N_{mac}$  denotes the macro degrees of freedom. We again plot the error obtained by using the same adaptive strategy without refining the micro mesh (fixed to  $\hat{h} = \frac{1}{8}$ ). As expected, the obtained asymptotic convergence rate is incorrect.

Finally, we provide a comparison with a uniformly refined FE-HMM starting from the same initial mesh. Obviously, the order of convergence is significantly lower than what can be obtained using adaptive methods. In order to get an accuracy of e.g.,  $\|u^0 - u^H\|_{H^1(\Omega)} \approx 0.07$  we need 16,384 macro elements in the fourth iteration of the uniformly refined FE-HMM, whereas only 796 macro elements are used in the 10th iteration of the adaptive FE-HMM. In order to reach an accuracy of  $\mathcal{O}(10^{-2})$ , approximatively  $\mathcal{O}(10^8)$  DOF would be needed for the uniform scheme.

For uniform refinement, micro solutions in the sampling domain of every triangle of the macro mesh have to be refined and recomputed at each step. In contrast, for the adaptive



**Figure 6.4.3:** Errors and error estimate in the  $H^1$ -norm for the crack problem described in Section 6.4.2.

| $\hat{h}$                              | 1/8 | 1/16 | 1/24 | 1/32  | 1/40 | 1/48 |
|--|-----|------|------|-------|------|------|
| <i>adaptive</i> FE-HMM, 10th iteration | 278 | 218  | 60   | 24    | 28   | 40   |
| <i>uniform</i> FE-HMM, 4th iteration   | -   | -    | -    | 16384 | -    | -    |

**Table 6.4.2:** Amount of micro problems with various  $\hat{h}$  (due to refinement) to be solved for the specific iteration in which we reach the accuracy  $\|e^H\|_{H^1(\Omega)} \leq 0.07$  in the Crack problem described in Section 6.4.2.

FE-HMM this needs only to be done for the macro triangles marked for refinement. In Table 6.4.2 we list the number of micro problems with a mesh size (complexity) of  $\hat{h}$  that need to be solved in the specific iteration when we reach an accuracy of  $\|e^H\|_{H^1(\Omega)} \leq 0.07$  for the adaptive and uniform refinement, respectively. Looking at the first line of Table 6.4.2, we see that with an adaptive scheme, the sampling domains of most of the macro elements need a micro mesh with a relative coarse resolution. This is in sharp contrast with a uniformly refined mesh, where all the sampling domains need to be solved with the same (fine) resolution.

As mentioned in Section 6.2.1, in an efficient implementation, one should store the contributions of the micro problems and the corresponding *multiscale flux* in every iteration and re-use these results for those elements which are not marked to be refined. In this way, the computational cost per iteration can be dramatically reduced, as only a fraction of the elements is refined at each iteration. For the Crack problem and the adaptive FE-HMM, we start with 256 micro problems with a micro mesh size chosen as  $\hat{h} = \frac{1}{8}$  in the first iteration. Four elements are refined into 8 elements, which require the solution of eight new micro problems with a micro mesh size  $\hat{h} = \frac{1}{8}$  (we round to get  $\hat{h} = \frac{1}{n \cdot 8}$ ,  $n \in \mathbb{N}^*$ ). On the other hand, the solutions of the 252 other micro problems will be re-used in the next iteration. After the second iteration, those 8 elements are yet again refined and we need to solve 16 new micro problems with micro mesh size  $\hat{h} = \frac{1}{16}$  and re-use all of the 252 solutions of the micro problems with micro mesh size  $\hat{h} = \frac{1}{8}$ . Summing over the 10 iterations, we only need to compute 436

| $\hat{h}$                                 | 1/8 | 1/16 | 1/24 | 1/32  | 1/40 | 1/48 |
|---|-----|------|------|-------|------|------|
| <i>adaptive</i> FE-HMM, <i>total cost</i> | 436 | 360  | 108  | 48    | 48   | 40   |
| <i>uniform</i> FE-HMM, <i>total cost</i>  | 256 | 1024 | 4096 | 16384 | -    | -    |

**Table 6.4.3:** Total amount of micro problems with various  $\hat{h}$  (due to refinement) to be solved to achieve an accuracy of  $\|e^H\|_{H^1(\Omega)} \leq 0.07$  in the Crack problem described in Section 6.4.2. (here we take into account all the iterations needed to reach the prescribed accuracy).

different micro problems of micro mesh size  $\hat{h} = \frac{1}{8}$ . In Table 6.4.3 we list the *total* amount of different micro problems (with different micro mesh) to be solved. For the uniform refinement, *every* triangle is divided into four new triangles, thus no information can be carried over from one iteration to the other.

In Table 6.4.4 we list various quantities illustrating the quality of the refinement procedure. It reveals that the error indicator converges with the same rate as the error itself. This is confirmed by the effectivity index which remains nearly constant over the iterations, showing that our adaptive refinement strategy is both effective and efficient.

In Table 6.4.5 we show the error and the experimental order of convergence when using a uniform refinement strategy. A comparison between Table 6.4.4 and 6.4.5 shows that a desired given accuracy of the numerical solution can be obtained with a much lower computational cost using an adaptive method (compare e.g., the number of elements needed to reach a certain accuracy).

### 6.4.2.1 Sampling domain size

We consider the crack problem in the situation when the size  $\delta$  of the sampling domain is not an integer multiple of  $\varepsilon$ . This situation might even arise for periodic problems when the exact size of the period is not known. We select Dirichlet boundary conditions in the micro problems and so estimates (6.3.26) apply.

We choose the same initial mesh size  $h$  for four different sampling domain sizes  $\delta_1 = \frac{4}{3}\varepsilon$ ,  $\delta_2 = \frac{5}{3}\varepsilon$ ,  $\delta_3 = \frac{11}{3}\varepsilon$  and  $\delta_4 = \frac{17}{3}\varepsilon$  (i.e.,  $\hat{h}$  differs), where  $\varepsilon = 10^{-5}$ . According to (6.3.26) one should take  $\delta \propto \sqrt{\varepsilon}$  (of course the value of  $\delta$  is unknown if  $\varepsilon$  is unknown. Furthermore, taking  $\delta = \sqrt{\varepsilon}$  can be computationally too expensive). We verify here that increasing the sampling domain size  $\delta$  does reduce the modeling error  $err_{mod}$  and hence improves the convergence rates. This is illustrated in Figure 6.4.4.

### 6.4.2.2 Random tensor

Many problems of interest are not periodic. As mentioned earlier, our adaptive algorithm does not rely on periodic problems (although the relation between micro and macro mesh does). To illustrate the versatility of the method, we test the behavior of the adaptive FE-HMM on the crack problem in Section 6.4.2 with a random tensor. This tensor is a log-normal

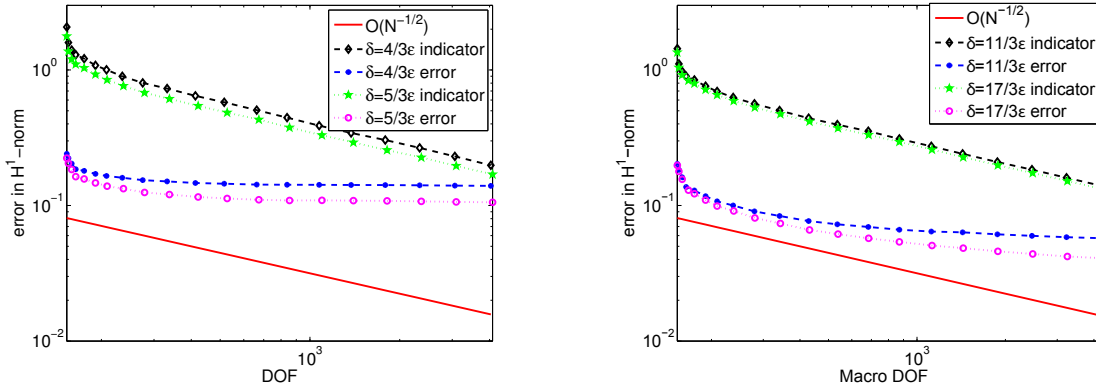
## 6.4 Numerical experiments

| iteration | #el   | $\ u^0 - u^H\ _{H^1(\Omega)}$ | $\eta_H(\Omega)$ | $EOC(e^H)$ | $EOC(\eta_H)$ | $Z_e$ | Eff  |
|-----------|-------|-------------------------------|------------------|------------|---------------|-------|------|
| 1         | 256   | 1.92e-01                      | 1.22e+00         |            |               |       | 6.33 |
| 2         | 260   | 1.73e-01                      | 9.41e-01         | 13.696     | 33.123        | 0.899 | 5.43 |
| 3         | 268   | 1.52e-01                      | 8.31e-01         | 8.715      | 8.255         | 0.876 | 5.49 |
| 4         | 284   | 1.24e-01                      | 7.58e-01         | 6.844      | 3.167         | 0.820 | 6.10 |
| 5         | 298   | 1.16e-01                      | 7.19e-01         | 2.960      | 2.170         | 0.931 | 6.21 |
| 6         | 329   | 1.02e-01                      | 6.43e-01         | 2.584      | 2.255         | 0.880 | 6.33 |
| 7         | 370   | 9.09e-02                      | 5.90e-01         | 1.934      | 1.485         | 0.893 | 6.49 |
| 8         | 432   | 8.12e-02                      | 5.23e-01         | 1.459      | 1.543         | 0.893 | 6.45 |
| 9         | 522   | 7.15e-02                      | 4.74e-01         | 1.340      | 1.036         | 0.881 | 6.62 |
| 10        | 648   | 6.41e-02                      | 4.23e-01         | 1.010      | 1.055         | 0.897 | 6.58 |
| 11        | 830   | 5.58e-02                      | 3.74e-01         | 1.125      | 1.005         | 0.870 | 6.71 |
| 12        | 1040  | 5.13e-02                      | 3.36e-01         | 0.746      | 0.940         | 0.919 | 6.54 |
| 13        | 1326  | 4.50e-02                      | 3.00e-01         | 1.082      | 0.948         | 0.877 | 6.67 |
| 14        | 1670  | 4.05e-02                      | 2.68e-01         | 0.896      | 0.964         | 0.902 | 6.62 |
| 15        | 2138  | 3.52e-02                      | 2.36e-01         | 1.136      | 1.019         | 0.869 | 6.71 |
| 16        | 2738  | 3.08e-02                      | 2.08e-01         | 1.085      | 1.022         | 0.874 | 6.76 |
| 17        | 3614  | 2.69e-02                      | 1.82e-01         | 0.988      | 0.965         | 0.872 | 6.76 |
| 18        | 4782  | 2.34e-02                      | 1.59e-01         | 0.989      | 0.970         | 0.871 | 6.80 |
| 19        | 6268  | 2.03e-02                      | 1.39e-01         | 1.057      | 1.010         | 0.867 | 6.85 |
| 20        | 8310  | 1.78e-02                      | 1.21e-01         | 0.940      | 0.961         | 0.876 | 6.80 |
| 21        | 10948 | 1.55e-02                      | 1.06e-01         | 1.009      | 0.998         | 0.870 | 6.85 |
| 22        | 14534 | 1.32e-02                      | 9.11e-02         | 1.114      | 1.036         | 0.854 | 6.90 |
| 23        | 19360 | 1.14e-02                      | 7.89e-02         | 1.031      | 1.010         | 0.863 | 6.94 |

**Table 6.4.4:** Iteration number, number of macro elements,  $H^1$ -error, error indicator, experimental order of convergence for the error and the error indicator, reduction factor and effectivity index for the crack problem described in Section 6.4.2.

| iteration                     | 1        | 2        | 3        | 4        |
|-------------------------------|----------|----------|----------|----------|
| #el                           | 256      | 1024     | 4096     | 16384    |
| $\ u^0 - u^H\ _{H^1(\Omega)}$ | 1.92e-01 | 1.36e-01 | 9.67e-02 | 6.85e-02 |
| $EOC(e^H)$                    |          | 0.496    | 0.495    | 0.497    |

**Table 6.4.5:** Iteration number, number of macro elements and experimental order of convergence of the  $H^1$  error for the crack problem described in Section 6.4.2 when using uniform refinement instead of an adaptive strategy.



**Figure 6.4.4:** Errors and error estimate in the  $H^1$ -norm for the crack problem with  $\delta \neq \varepsilon$  described in Section 6.4.2.1. We see that despite a further refinement of the macro and micro mesh, the error in the  $H^1$ -norm stagnates due to the dominating modeling error  $err_{mod}$  (see Section 2.4.1 and (2.4.15)). We further see that increasing the sampling domain size  $\delta$  decreases  $err_{mod}$  and therefore we obtain a lower overall error  $\|u^0 - u^H\|_{H^1(\Omega)}$ .

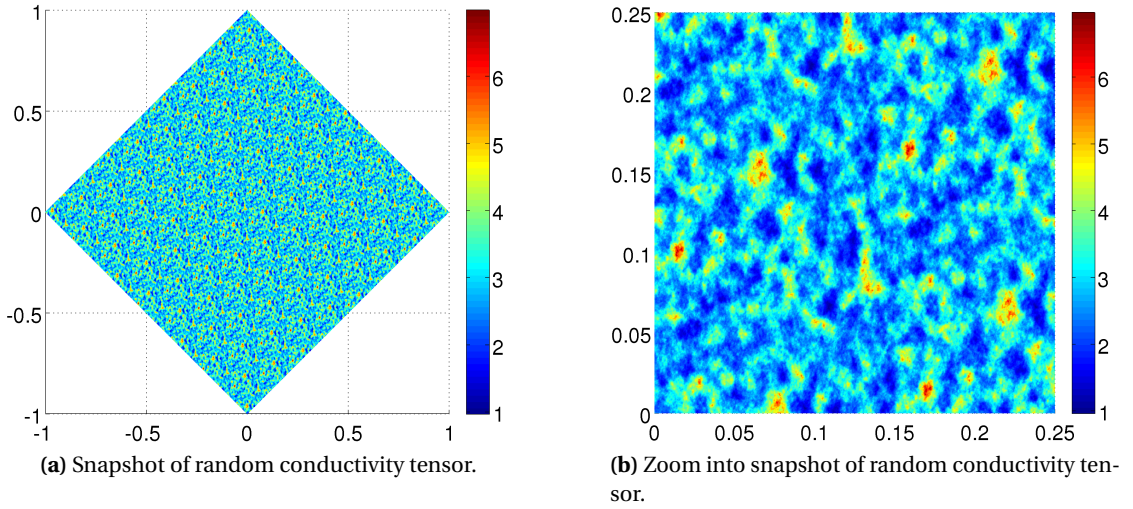
stochastic field generated by the moving ellipse average method (see [115, Section 4.1]). The numerically generated values of the tensor are given at  $9000^2$  discrete points and we use bilinear interpolation to obtain a smooth representation of  $a^\varepsilon(x)$ .

We set the correlation lengths of the stochastic field to  $\varepsilon_{x_1} = 0.0045$  and  $\varepsilon_{x_2} = 0.0065$  and choose the mean equal to be 1 and variance to be  $\sigma = 0.25$ . A snapshot of this tensor is shown in Figure 6.4.5.

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

We choose an initial micro mesh size  $\hat{h} = \frac{1}{8}$  and sampling domains of size  $\delta_1 = 0.005$ ,  $\delta_2 = 0.010$ ,  $\delta_3 = 0.020$  and  $\delta_4 = 0.040$  (we keep  $h$  fixed when increasing the sampling domain size). In Figure 6.4.6 we see that the indicator and the error follow the expected (optimal) convergence rate. Furthermore, increasing the sampling domain size reduces the modeling error and thus leads to a more accurate solution.

*Remark 46.* For  $H < \varepsilon$ , there is less than a period of the fine scale solution that is averaged (by the  $L^2$ -projection) on each macro element of the mesh  $\mathcal{T}_{H,FE-HMM}$  and as  $H \rightarrow 0$  the  $L^2$ -projection restores the behavior of the fine scale solution whose gradient has an  $\mathcal{O}(1)$  discrepancy with the solution obtained by the FE-HMM (see Section 2.4.1 and recall e.g., Remark 45).



**Figure 6.4.5:** Snapshot of the random conductivity tensor used in the crack problem described in Section 6.4.2.2.

### 6.4.3 L-shape problem

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

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

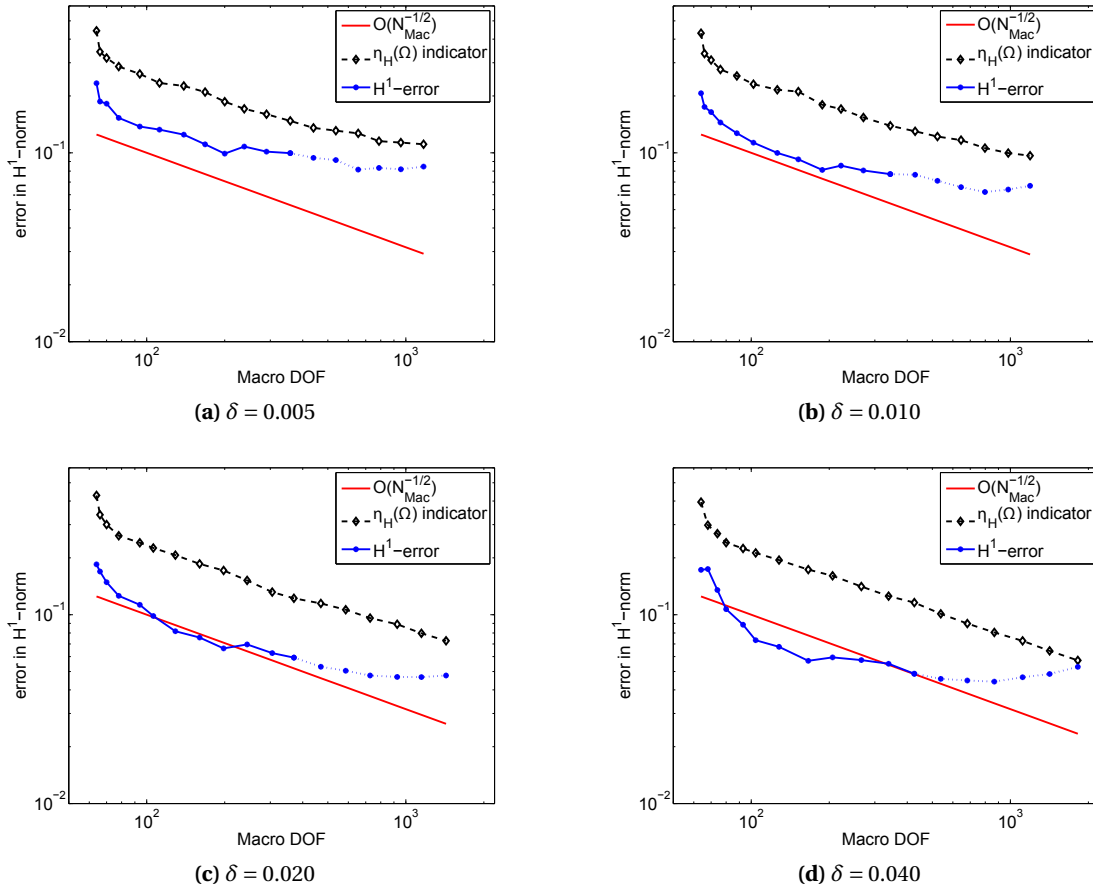
with  $\Omega = (-1, 1)^2 \setminus [0, 1] \times [-1, 0]$  (see Figure 6.4.7). As a first test, we consider the tensor

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

introduced in Section 6.4.2, where  $I_2$  is the  $2 \times 2$  unit matrix. The tensor has coefficients chosen such that the homogenized tensor matches the identity tensor. We will further investigate the same tensor with different coefficients

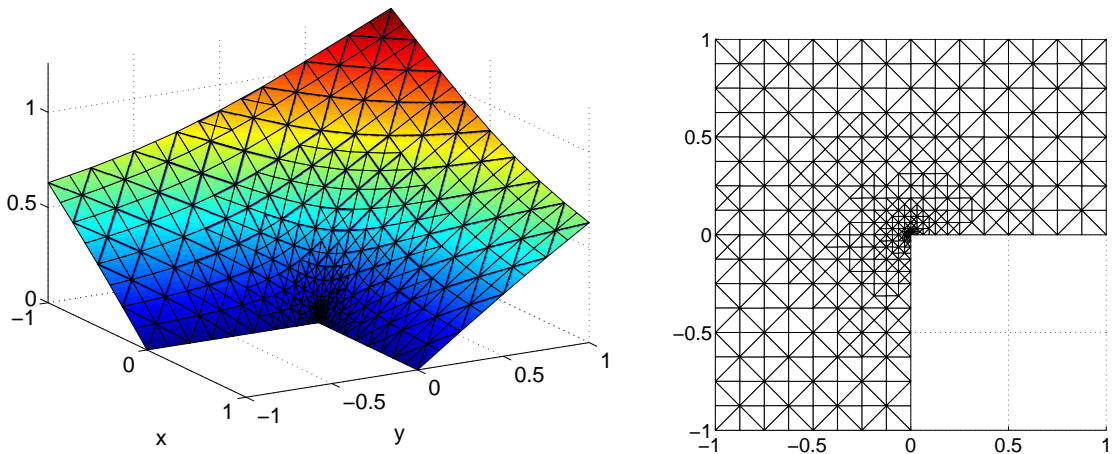
$$a_2^\varepsilon(x) = a_2\left(\frac{x}{\varepsilon}\right) = \frac{400}{21\sqrt{41}} \left( \sin\left(2\pi \frac{x_1}{\varepsilon}\right) + \frac{21}{20} \right) \left( \cos\left(2\pi \frac{x_2}{\varepsilon}\right) + \frac{21}{20} \right) \cdot I_2.$$

The homogenized problem corresponding to this latter tensor again leads to a problem with a homogenized tensor equal to  $I_2$ . As the tensor  $a_2^\varepsilon$  has a coercivity bound closer to zero, one expects a larger error than with the tensor  $a_1^\varepsilon$  (recall that our *a posteriori* error estimates depend on the bound (2.1.2)). For both  $a_1^\varepsilon$  and  $a_2^\varepsilon$ , an analytical homogenized solution exists and is given by  $u^0(r) = r^{\frac{2}{3}} \sin\left(\frac{2}{3}\vartheta\right)$  where  $x_1 = r \cos(\vartheta)$  and  $x_2 = r \sin(\vartheta)$ . We take the value of this exact solution for the Dirichlet boundary condition  $g_D = u^0$  (in the computation below

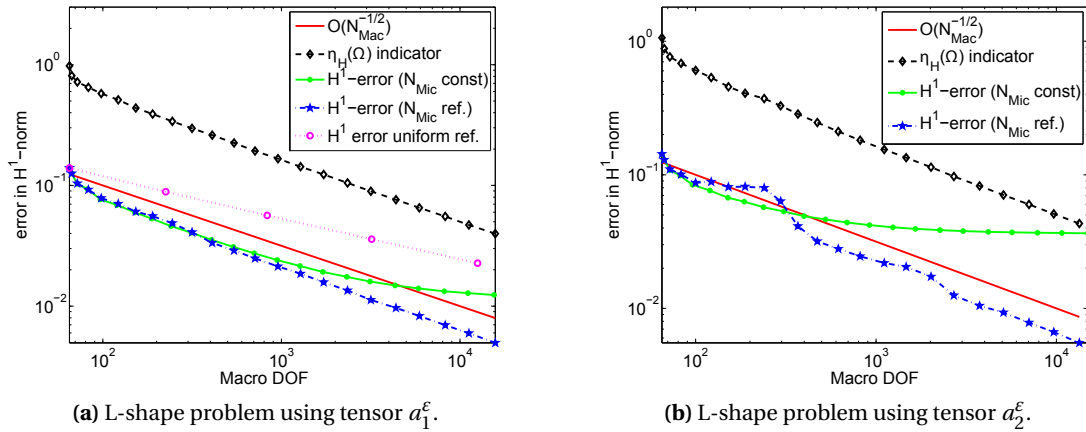


**Figure 6.4.6:** Error and error estimates in the  $H^1$ -norm for the crack problem with a random tensor described in Section 6.4.2.2 (for plotting reasons the indicator is scaled by a factor of 0.1). For the  $H^1$ -error, a continuous line is used in the case  $H > \varepsilon$ , where the projection (that is solely used for the computation of the error between the FE-HMM solution  $u^H$  and the fine-scale solution  $u_{FEM}^H$ ) is expected to resemble the error of the homogenized solution. For the case when for some  $K$  we have  $H < \varepsilon$ , where we expect the projection to resemble the error between  $u^H$  and the fine-scale solution, a dotted line is used. In the latter case of  $H < \varepsilon$  the error computed via projection therefore approaches an  $\mathcal{O}(1)$  discrepancy. See Remark 46.





**Figure 6.4.7:** FE-HMM solution and refined grid after 10 iterations using tensor  $a_1^\varepsilon$  of L-shape problem described in Section 6.4.3.



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

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

**Figure 6.4.8:** Errors and error estimate in the  $H^1$ -norm for the L-shape problem described in section. For the left plot we used tensor  $a_1^\varepsilon$ , for the right-hand side we used  $a_2^\varepsilon$ .

we use Algorithm 37 with Dörfler’s bulk-chasing marking strategy with parameter  $\theta = 0.3$ ).

In Figure 6.4.8 we compare the error and the indicator  $\eta_H(\Omega)$  in the  $H^1$  norm for the tensors  $a_1^\varepsilon$  and  $a_2^\varepsilon$ . We choose an initial mesh-size of  $\hat{h} = \frac{1}{8}$  and  $\delta = \varepsilon = 10^{-5}$  with periodic boundary conditions in the micro problems. For both problems we find that the corresponding indicator converges with the same (optimal) rate of  $\mathcal{O}(N_{mac}^{-1/d})$  as the errors between the FE-HMM solution and the exact homogenized solution, thus confirming numerically our theoretical estimates. As noted earlier we verify again that the asymptotic convergence rate is incorrect when using a constant micro mesh size  $\hat{h} = \frac{1}{8}$ . The asymptotic limit of the macroscopic error in Figure 6.4.8b (green continuous line with disks) reflects the error introduced through the micro problems and illustrates that the error stagnates even though we keep refining our macro grid. A comparison between Figures 6.4.8a and 6.4.8b shows that with constant micro mesh size the error is significantly larger when using  $a_2^\varepsilon$  as compared to  $a_1^\varepsilon$  (the tensor  $a_2^\varepsilon$  is closer to being singular and thus needs a higher resolution of the micro mesh in order to avoid a singular stiffness matrix).

Various quantities illustrating the quality of the refinement procedure and confirming the correct experimental order of convergence are reported in Table 6.4.6. In particular, we see that the effectivity index is approximatively constant indicating that our adaptive scheme is both effective and efficient. It is furthermore robust with respect to the change of tensors in the problem. In Tables 6.4.7 and 6.4.8 we compare again the number of sampling domains and the resolution of the mesh needed to solve the micro problems with adaptive and non-adaptive strategies. The results illustrate once more the importance of adaptive methods for multiscale problems.

## 6.4 Numerical experiments

| iteration | #el   | $\ u^0 - u^H\ _{H^1(\Omega)}$ | $\eta_H(\Omega)$ | $EOC(e^H)$ | $EOC(\eta_H)$ | $Z_e$ | Eff := $\frac{\eta_H}{e^H}$ |
|-----------|-------|-------------------------------|------------------|------------|---------------|-------|-----------------------------|
| 1         | 96    | 1.43e-01                      | 1.061            |            |               | 0.135 |                             |
| 2         | 100   | 1.30e-01                      | 0.880            | 4.886      | 9.145         | 0.905 | 6.80                        |
| 3         | 110   | 1.10e-01                      | 0.763            | 3.423      | 2.995         | 0.850 | 6.94                        |
| 4         | 130   | 1.01e-01                      | 0.682            | 1.064      | 1.343         | 0.915 | 6.76                        |
| 5         | 164   | 8.72e-02                      | 0.604            | 1.235      | 1.048         | 0.866 | 6.94                        |
| 6         | 206   | 8.85e-02                      | 0.534            | -0.133     | 1.082         | 1.015 | 6.02                        |
| 7         | 265   | 8.10e-02                      | 0.456            | 0.708      | 1.263         | 0.915 | 5.62                        |
| 8         | 335   | 8.15e-02                      | 0.408            | -0.052     | 0.943         | 1.006 | 5.00                        |
| 9         | 435   | 7.97e-02                      | 0.372            | 0.166      | 0.707         | 0.979 | 4.67                        |
| 10        | 542   | 6.34e-02                      | 0.329            | 2.090      | 1.128         | 0.795 | 5.18                        |
| 11        | 670   | 4.13e-02                      | 0.285            | 4.043      | 1.347         | 0.651 | 6.90                        |
| 12        | 866   | 3.18e-02                      | 0.246            | 2.036      | 1.143         | 0.770 | 7.75                        |
| 13        | 1156  | 2.79e-02                      | 0.211            | 0.907      | 1.064         | 0.877 | 7.58                        |
| 14        | 1545  | 2.45e-02                      | 0.181            | 0.882      | 1.039         | 0.880 | 7.41                        |
| 15        | 2123  | 2.19e-02                      | 0.154            | 0.729      | 1.014         | 0.891 | 7.04                        |
| 16        | 2787  | 2.05e-02                      | 0.134            | 0.483      | 1.022         | 0.936 | 6.58                        |
| 17        | 3876  | 1.73e-02                      | 0.114            | 1.030      | 0.998         | 0.844 | 6.58                        |
| 18        | 5238  | 1.25e-02                      | 0.097            | 2.148      | 1.060         | 0.724 | 7.75                        |
| 19        | 7312  | 1.05e-02                      | 0.083            | 1.062      | 0.977         | 0.838 | 7.87                        |
| 20        | 9934  | 9.30e-03                      | 0.071            | 0.765      | 1.020         | 0.889 | 7.58                        |
| 21        | 13772 | 7.80e-03                      | 0.060            | 1.074      | 0.998         | 0.839 | 7.69                        |
| 22        | 19010 | 6.70e-03                      | 0.051            | 0.996      | 1.008         | 0.852 | 7.63                        |
| 23        | 26430 | 5.50e-03                      | 0.043            | 1.129      | 1.011         | 0.830 | 7.81                        |

**Table 6.4.6:** Iteration number, number of macro elements,  $H^1$ -error, error indicator, experimental order of convergence for the error and the indicator, reduction factor and effectivity of the L-shape problem using tensor  $a_1^\varepsilon$  described in Section 6.4.3.

| $\hat{h}$                      | 1/8 | 1/16 | 1/24 | 1/32 | 1/40 | 1/48  | 1/56 | 1/64 | 1/80 |
|--------------------------------|-----|------|------|------|------|-------|------|------|------|
| <i>adaptive</i> FE-HMM, it. 13 | 0   | 862  | 246  | 64   | 72   | 30    | 18   | 24   | 24   |
| <i>uniform</i> FE-HMM, it. 5   | -   | -    | -    | -    | -    | 24576 | -    | -    | -    |

**Table 6.4.7:** Amount of micro problems with various  $\hat{h}$  (due to refinement) to be solved for the specific iteration in which we reach an accuracy  $\|e^H\|_{H^1(\Omega)} \leq 0.025$  in the L-Shape problem described in Section 6.4.3.

| $\hat{h}$                            | 1/8 | 1/16 | 1/24 | 1/32 | 1/40 | 1/48  | 1/56 | 1/64 | 1/80 |
|--------------------------------------|-----|------|------|------|------|-------|------|------|------|
| <i>adaptive</i> FE-HMM, <i>total</i> | 288 | 1556 | 388  | 112  | 96   | 48    | 36   | 36   | 24   |
| <i>uniform</i> FE-HMM, <i>total</i>  | 96  | 384  | 1536 | 6144 | 0    | 24576 | -    | -    | -    |

**Table 6.4.8:** Total amount of micro problems with various  $\hat{h}$  (due to refinement) to be solved to reach an accuracy of  $\|e^H\|_{H^1(\Omega)} \leq 0.025$  in the L-Shape problem described in Section 6.4.3 (here we take into account all the iterations needed to reach the prescribed accuracy).

## 6.5 A comparison to the *a posteriori* FE-HMM obtained in a two-scale framework

A first *a posteriori* error analysis for the FE-HMM was obtained by Ohlberger [92], where the problem (6.1.1) is considered. While we will use techniques close to the standard residual-based and goal-oriented FEM *a posteriori* analysis of [113, 20], the work in [92] relies on a reformulation of the FE-HMM in a two-scale framework. We briefly sketch the ansatz of the method in [92] with the goal of bringing the notation of [92] in accord to the notation of this thesis. We then put our results and the results of [92] in perspective.

The *two-scale convergence* was introduced by Nguetseng in [84] and generalized by Allaire in [23]. We recall its definition and state the two-scale homogenized equation.

**Definition 47.** (See [84]). Let  $\{v^\varepsilon\}$  be a sequence of functions in  $L^2(\Omega)$ . One says that  $\{v^\varepsilon\}$  two-scale converges to  $v_0 = v_0(x, y)$  with  $v_0 \in L^2(\Omega \times Y)$  if for any function  $\psi = \psi(x, y) \in L^2(\Omega; C_{per}^\infty(Y))$  we have

$$\lim_{\varepsilon \rightarrow 0} \int_{\Omega} v^\varepsilon(x) \psi\left(x, \frac{x}{\varepsilon}\right) dx = \frac{1}{|Y|} \int_{\Omega} \int_Y v_0(x, y) \psi(x, y) dy dx.$$

Based on the two-scale convergence theory, the so-called two-scale homogenized equation can be defined.

**Definition 48** (Two-scale homogenized equation). (See [23]). A pair

$$(u^0, u^1) \in H_0^1(\Omega) \times L^2(\Omega, W_{per}^1(Y))$$

is called the homogenized solution to problem (2.1.1) if it satisfies

$$\begin{aligned} & \int_{\Omega} \int_Y a(x, y) (\nabla_x u^0(x) + \nabla_y u^1(x, y)) (\nabla_x \Phi(x) + \nabla_y \varphi(x, y)) dx dy \\ &= \int_{\Omega} f(x) \Phi(x) dx \end{aligned}$$

for all

$$(\Phi, \varphi) \in H_0^1(\Omega) \times L^2(\Omega, W_{per}^1(Y)).$$

See also [45, Chapter 9] for a comprehensive introduction to two-scale convergence.

The crucial component of [92] is the reformulation of the FE-HMM in terms of the two-scale homogenized equation. We introduce the solution pair as

$$(u^H, \mathcal{K}_h[u^H]) \in V^1(\Omega, \mathcal{T}_H) \times V_H^0(\Omega, S^1(\mathcal{K}_{\delta_\varepsilon}, \mathcal{T}_h)),$$

where

$$V_H^0(\Omega, S^1(K_{\delta_\ell}, \mathcal{T}_h)) := \left\{ \varphi_h \in L^2\left(\Omega; W_{per}^1(Y)\right), |\varphi_h(\cdot, y)|_K \text{ is constant } \forall y \in Y \right. \\ \left. \text{and } \varphi_h(x, \cdot) \in S^1(K_{\delta_\ell}, \mathcal{T}_h) \forall x \in \Omega \right\},$$

where the micro FE space  $S^1(K_{\delta_\ell}, \mathcal{T}_h)$  is defined in (2.4.9).

Here  $\mathcal{K}_h[u^H]$  is the so-called discrete fine scale corrector  $\mathcal{K}_h[u^H]$  obtained from a reconstruction  $u^{H,\varepsilon}(x) - u^H(x)$  with  $u^{H,\varepsilon}(x)$  as in (2.4.13). (The reconstruction corresponds to the extension of  $(u^h - u^H)$ , which is only defined on  $K_\varepsilon$  on the whole macro element  $K$ ). Then we can state the following Lemma [92].

**Lemma 49** (Reformulation of the FE-HMM). *Under appropriate assumptions (see [92, Lemma 3.5]), the pair  $(u^H, \mathcal{K}_h[u^H])$  is the solution of*

$$\int_{\Omega} \int_Y a_h(x, y) (\nabla_x u^H(x) + \nabla_y \mathcal{K}_h[u^H](x, y)) (\nabla_x \Phi(x) + \nabla_y \varphi(x, y)) dx dy \\ = \int_{\Omega} f(x) \Phi(x) dx$$

for all

$$(\Phi, \varphi) \in V^1(\Omega, \mathcal{T}_H) \times L^2(\Omega, S^1(K_{\delta_\ell}, \mathcal{T}_h)),$$

where  $a_h(x, y)|_{K_j \times T_i} = a(x_j, y_i)$  is a discretized tensor.

Based on the reformulation, *a posteriori* upper and lower bounds are obtained in terms of the error

$$\|\nabla_x e^0 + \nabla_y e^1\|_{L^2(\Omega \times Y)}$$

in the two-scale norm, where  $e^0 := u^0 - u^H$  and  $e^1 := u^1 - \mathcal{K}_h[u^H]$ .

*A posteriori* error estimates for problems in perforated domains were considered in [61]; there error estimates in the macroscopic  $L^2(\Omega)$ -norm are obtained using the reformulation of the FE-HMM as above and subsequently employing duality arguments. Recent extensions include advection-diffusion problems [60] and problems with monotone operators [62].

**Comparison of the results of [92] to our results.**

The following discussion is taken in part from [12].

In the approach described in [92], the reformulation of the FE-HMM is based on a tensor product FEM with quadrature in the slow variable. Therefore, the *a posteriori* estimates are obtained in the two-scale norm over  $\Omega \times Y$  (with the physical domain  $\Omega$  and the domain of the microscopic variable  $Y$ ) and not in a norm related to the physical domain. It is not straightforward to derive *a posteriori* (or *a priori*) estimates in the energy norm – or in general

quantities of interest – for the physical domain from the results obtained in [92] and to our knowledge, such results have not yet been obtained. In the approach we follow, the *a posteriori* error estimates for the upper and lower bounds are derived in the energy norm of the physical domain. While the analysis in [92] is restricted to a class of homogenization problems where the tensor is given in an explicit two-scale form  $a^\varepsilon(x) = a(x, x/\varepsilon)$  and where the tensor features periodicity in the fast variables  $x/\varepsilon$  (see [92, Assumption 3.1]), we do not require an explicit decomposition of the tensor. Furthermore our analysis also applies to general (non-periodic) tensors (although it involves a data approximation error, which can only be explicitly estimated when we make additional spatial assumptions on the fast variable, such as periodicity or random homogeneity). The analysis in [92] requires an *a priori* knowledge of the exact periodicity  $\varepsilon$  of the problem in order to make the sampling domains  $K_\delta$  span exactly one period in each spatial direction, see [92, Remark 3.6] (in the more general case  $\delta > \varepsilon$  an extra error will be introduced). We derive estimates, which hold for general sampling domains of size  $\delta > \varepsilon$ . Further, the analysis in [92] assumes piecewise linear macro FE.

Our *a posteriori* error estimator  $\eta_H(K)$  only depends on  $a^\varepsilon$ ; the data approximation error however does involve the *unknown* homogenized tensor  $a^0(x)$ . As our techniques are designed to follow the classical framework to derive *a posteriori* error estimates, we obtain a data approximation error of the tensor, which reads in its most general form

$$\|(a_K^0 - a^0(x)) \nabla u^H\|_{L^2(K)}$$

(for piecewise linear macro FE), whereas the data approximation error in [92] is given by

$$\|(a_h(x, y) - a(x, y)) (\nabla_x u^H(x) + \nabla_y \mathcal{K}_h[u^H])(x, y)\|_{L^2(K \times T)}.$$

This means that our data approximation error involves the (unknown) exact homogenized tensor  $a^0(x)$ , whereas the data approximation error obtained in [92] involves estimates in terms of the (known)  $a(x, y)$ . Under the assumptions of [92], i.e.  $a^\varepsilon(x)$  is given in the form  $a(x, x/\varepsilon)$ ,  $\varepsilon$  is known exactly,  $a^\varepsilon$  is periodic and using periodic boundary conditions in the micro problems, we can choose  $\delta = \varepsilon$  and collocate in the slow variable, i.e. evaluate the tensor as  $a(x_{K_\ell}, x/\varepsilon)$  and the modeling error vanishes. The remaining macro and micro error can be estimated as  $CH_K$  and  $C\left(\frac{h}{\varepsilon}\right)^2$  and therefore our estimates can be made independent of the unknown homogenized tensor  $a^0(x)$ .

Our results also hold for higher order FE. The *a posteriori* error estimates in the energy norm obtained in this thesis are also instrumental for deriving *a posteriori* error estimates in *quantities of interest*, which are expressed by a linear bounded functional in the physical domain (see Chapter 8).

## 6.6 Summary

This summary is taken in part from [12].

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

The analysis and the method presented in this Chapter are limited to piecewise linear macro FE; in Chapter 7 we will present the extension of the analysis and method to higher order FE. This will allow us to find in Chapter 8 an estimate of the error in quantities of interests that are needed for a specific design purpose.





# 7 Higher order Adaptive FE-HMM

In Chapter 6 we introduced the *multiscale flux* and *multiscale jump*, which allowed us to derive *a posteriori* error estimates for the multiscale FE-HMM for piecewise linear FE. The derivation largely followed the error analysis of the classical single-scale residual-based adaptive FEM. We now want to extend the method for higher order FE. Can we find a *higher order multiscale flux* and *multiscale jump* similarly as before? While the multiscale jump for the piecewise linear case is a constant over the element interface, in the higher order case it is not. The challenge is now to construct a *higher order* multiscale jump that is continuous and that is solely based on the discrete solutions of micro problems within two adjacent macro elements. We will see how this can be done in this Chapter and derive *a posteriori* error estimates for the higher order FE-HMM. We mention that these constructions have first been used in the context of the discontinuous Galerkin FE-HMM [7].

In Section 7.1 we state the model problem. In Section 7.2 we introduce the *higher order* multiscale components and relate them to the classical, single-scale counterparts. In Section 7.3 we state the main results of our *a posteriori* error analysis for higher order FE. We further give algorithmic details. The main results are proved in Section 7.4 and the data approximation error is estimated in the separate Section 7.5. We present a numerical example, which confirms the theoretical estimates in Section 7.6.

## 7.1 Model problem

Let  $\mathcal{T}_H$  denote a conformal mesh and let  $\mathcal{E}_H$  be the set of interfaces. Given  $f \in L^2(\Omega)$  and a domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , we consider the second-order elliptic multiscale equation

$$\begin{aligned} -\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) &= f & \text{in } \Omega, \\ u^\varepsilon &= 0 & \text{on } \partial\Omega, \end{aligned} \tag{7.1.1}$$

where  $a^\varepsilon$  is symmetric, satisfies  $a^\varepsilon(x) \in (L^\infty(\Omega))^{d \times d}$  and is uniformly elliptic and bounded. We consider in what follows only homogeneous Dirichlet boundary conditions, but emphasize that the *a posteriori* estimates can also be derived for more general boundary conditions (such as Neumann or Robin) following the lines of the results presented in this Chapter. We discretize problem (7.1.1) using the FE-HMM, so that the problem in its weak, discrete form

reads: find  $u^H \in V^p(\Omega, \mathcal{T}_H)$  s.t.

$$B(u^H, v^H) = \int_{\Omega} f v^H dx \quad \forall v^H \in V^p(\Omega, \mathcal{T}_H), \quad (7.1.2)$$

where  $V^p(\Omega, \mathcal{T}_H)$  is given in (7.2.1) and where  $B(\cdot, \cdot)$  is given in (2.4.6).

## 7.2 Preliminaries

In this Section we present the necessary tools that are needed for extending our adaptive FE-HMM to higher order FEs. In what follows, we show how to relate the special *higher order* FE-HMM components such as the multiscale bilinear form, the multiscale fluxes and jumps to their equivalent classical, single-scale FEM counterparts. This will allow us to largely follow classical adaptive FEM techniques in order to derive error bounds, similarly to what we presented in Chapter 6. These techniques have first been introduced in the context of the discontinuous Galerkin FE-HMM [7]. The equivalence of the components builds the foundation for both the residual-based adaptive FE-HMM in Chapter 7 as well as the adaptive goal-oriented FE-HMM in Chapter 8.

### 7.2.1 Quadrature rules and interpolation

We consider the higher order FE space

$$V^p(\Omega, \mathcal{T}_H) = \{v^H \in H_0^1(\Omega); v^H|_K \in \mathcal{P}^p(K), \forall K \in \mathcal{T}_H\}, \quad (7.2.1)$$

where  $\mathcal{P}^p$  is the space of piecewise polynomials on the element  $K$  of degree  $p$ , see Chapter 2, and where for simplicity, we restrict ourselves to simplicial elements in what follows. We consider a given quadrature formula (QF)  $\{x_{K_\ell}, \omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$  as defined in Section 2.4 that satisfies condition **(Q2)** (see (2.4.5) on page 24). Condition **(Q2)** requires that the QF for simplicial elements, which is based on  $\mathcal{L}$  points is exact for polynomials of degree  $2p - 2$  if  $p > 1$ .

For two spatial dimensions, the bound  $\mathcal{L} \geq \frac{1}{2}p(p+1)$  holds (see [103, Section 1]), and similarly for three spatial dimensions we have that  $\mathcal{L} \geq \frac{1}{6}p(p+1)(p+2)$ . There exist some well-known quadrature formulae which minimize these inequalities in the following sense:

$$\begin{cases} \mathcal{L} = \frac{1}{2}p(p+1) & \text{for } d = 2, \\ \mathcal{L} = \frac{1}{6}p(p+1)(p+2) & \text{for } d = 3. \end{cases} \quad (7.2.2)$$

For a QF satisfying (7.2.2), we consider the following interpolation problem: for  $v \in C^0(K)$  find  $\Pi_v(x) \in \mathcal{P}^{p-1}(K)$  s.t.

$$\Pi_v(x_{K_\ell}) = v(x_{K_\ell}), \quad \ell = 1, \dots, \mathcal{L} \quad (7.2.3)$$

where the interpolation nodes  $x_{K_\ell}$  are given by the quadrature nodes of the QF  $\{x_{K_\ell}, \omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$ .

**Proposition 50.** *Assume that **(Q2)** (see (2.4.5)) holds for  $\sigma = 2p - 2$ . Then, the interpolation problem (7.2.3) has a unique solution.*

*Proof.* The space of polynomials in  $d$  variables of degree at most  $(p-1)$  on the element  $K$  is denoted by  $\mathcal{P}_d^{p-1}(K)$ . The dimension of the vector space is (see [47, Section 3])

$$\dim \mathcal{P}_d^{p-1} = \binom{(p-1)+d}{p-1} = \mathcal{L}.$$

We now show that the interpolating polynomial  $\Pi_v(x)$  is uniquely defined. For that, let us assume that there exists a polynomial  $\Pi_v(x)$  in  $\mathcal{P}_d^{p-1}(K)$  such that  $\Pi_v(x_{K_\ell}) = 0$ ,  $\ell = 1, \dots, \mathcal{L}$ , but  $\Pi_v(x) \not\equiv 0$ . The integration (2.4.5) of QF condition **(Q2)** is exact for polynomials up to a degree of  $2(p-1)$ , i.e., we have

$$\int_K (\Pi_v(x))^2 dx = \sum_{\ell=1}^{\mathcal{L}} \omega_{K_\ell} (\Pi_v(x_{K_\ell}))^2 = 0.$$

This is in contradiction with  $\Pi_v(x) \not\equiv 0$ , and therefore, it follows that  $\Pi_v(x) \equiv 0$  and that the interpolation problem has a unique solution.  $\square$

We remark that condition (7.2.2) is satisfied by various well-known QF (see e.g. [55, Chap. 8] or [119]). We refer the reader to [103] for a general discussion on the assumed condition and to [47] for an overview of the construction of cubature formulae. It is strongly advised to choose such a QF, as it minimizes the number of quadrature points and therefore the number of micro problems needed to be solved at each macro element. It is also possible to generalize the FE-HMM to the general case – for the analysis however, this will lead to additional technicalities.

### 7.2.2 Higher order multiscale fluxes and jumps

The goal is now to generalize the multiscale flux defined in (6.2.1) for higher order FE. We introduced in Section 6.2 an expression that relates the macro and the micro fluxes for piecewise linear FE-HMM. We found, for  $v^H \in V^1(\Omega, \mathcal{T}_H)$  the equation (6.3.8), which is

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

with a numerically homogenized tensor  $a_K^0$  (from eq. (6.2.3)), which is given by (see also [12, Lemma 9])

$$a_K^0 = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \left( I + J_{\psi_{K_\delta}^h(x)}^T \right) dx.$$

For higher order FE-HMM, we can write equivalent expressions for all  $\mathcal{L}$  micro solutions  $v_{K_\ell}^h$  on the  $\mathcal{L}$  sampling domains  $K_{\delta_\ell}$  located around the quadrature nodes  $x_{K_\ell}$  within the macro element  $K$ . For each vector  $\mathbf{e}_i \in \mathbb{R}^d$ ,  $i = 1, \dots, d$  we consider in every sampling domain  $K_{\delta_\ell}$  the solution  $\psi_{K_{\delta_\ell}}^{i,h} \in S^q(K_{\delta_\ell}, \mathcal{T}_h)$  of the problem

$$\int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla \psi_{K_{\delta_\ell}}^{i,h} \cdot \nabla z^h dx = - \int_{K_{\delta_\ell}} a^\varepsilon(x) \mathbf{e}_i \cdot \nabla z^h dx \quad \forall z^h \in S^q(K_{\delta_\ell}, \mathcal{T}_h). \quad (7.2.4)$$

We again introduce a numerically homogenized tensor denoted as  $a_K^0(x_{K_\ell})$ ,

$$a_K^0(x_{K_\ell}) = \frac{1}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \left( I + J_{\psi_{K_{\delta_\ell}}^h}(x) \right) dx, \quad (7.2.5)$$

where  $J_{\psi_{K_{\delta_\ell}}^h}(x)$  is a  $d \times d$  matrix whose entries are given by  $\left( J_{\psi_{K_{\delta_\ell}}^h}(x) \right)_{ij} = (\partial \psi_{K_{\delta_\ell}}^{i,h}) / (\partial x_j)$ . Notice that  $a_K^0(x_{K_\ell})$  does now depend on the quadrature point  $x_{K_\ell}$  that is corresponding to the sampling domain  $K_{\delta_\ell}$ . The tensors  $a_K^0(x_{K_\ell})$  are solely a tool for the analysis but need not be used for the computation of our error estimators. The following relation between the fluxes in the micro problem and the fluxes of the higher order macro problem holds (see Corollary of Lemma 51).

$$\frac{1}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_\ell}^h dx = a_K^0(x_{K_\ell}) \nabla v^H(x_{K_\ell}), \quad \ell = 1, \dots, \mathcal{L}. \quad (7.2.6)$$

This will allow us to extend the multiscale jump (6.2.1) to higher order FE. On every macro element  $K$  we define the interpolating polynomial  $\Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x)$  in  $(\mathcal{P}^{p-1}(K))^d$ , which satisfies

$$\Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x_{K_\ell}) = \frac{1}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_\ell}^h dx, \quad \ell = 1, \dots, \mathcal{L}. \quad (7.2.7)$$

We will refer to the interpolating polynomial  $\Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x)$  as the **higher order multiscale flux**. Using (7.2.6) we see that

$$\Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x_{K_\ell}) = a_K^0(x_{K_\ell}) \nabla v^H(x_{K_\ell}), \quad \ell = 1, \dots, \mathcal{L}. \quad (7.2.8)$$

Generalizing (6.2.1), we introduce for each interior interface  $e$  of two elements  $K^+$  and  $K^-$  the following **jump of higher order multiscale fluxes**

$$\llbracket \Pi_{a^\varepsilon \nabla v_{K_\ell}^h} \rrbracket_e(s) := \begin{cases} \left( \Pi_{a^\varepsilon \nabla v_{K_\ell^+}^h}(s) - \Pi_{a^\varepsilon \nabla v_{K_\ell^-}^h}(s) \right) \cdot n_e & \text{for } e \notin \partial\Omega, \\ 0 & \text{for } e \subset \partial\Omega, \end{cases} \quad (7.2.9)$$

where the unit outward normal  $n_e$  is chosen to be  $n_e = n^+$ . We will omit the index  $K_\ell^+$  and  $K_\ell^-$  for the micro solutions  $v_{K_\ell}^h$  in  $\llbracket \Pi_{a^\varepsilon \nabla v_{K_\ell}^h} \rrbracket_e(s)$  to emphasize that the jump over the edge  $e$  involves sampling domains  $K_{\delta_\ell}$  in two adjacent elements  $K^+$  and  $K^-$ .

### 7.2.3 FE-HMM bilinear form

We will now relate the FE-HMM bilinear form  $B(\cdot, \cdot)$ , which is given in (2.4.6) and reads

$$B(v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_\ell}^h \cdot \nabla w_{K_\ell}^h dx, \quad (7.2.10)$$

to a single-scale homogenized bilinear form (2.1.7) with quadrature,  $B_{0,H}(\cdot, \cdot)$ , which reads

$$B_{0,H}(v^H, w^H) := \sum_{K \in \mathcal{F}_H} \sum_{\ell=1}^{\mathcal{L}} \omega_{K_\ell} a^0(x_{K_\ell}) \nabla v^H(x_{K_\ell}) \cdot \nabla w^H(x_{K_\ell}). \quad (7.2.11)$$

**Lemma 51.** *For all micro solutions  $v_{K_\ell}^h, w_{K_\ell}^h$  of (2.4.7) in  $S^q(K_\delta, \mathcal{F}_h)$  that are constrained by the linearized macro solutions  $v_{\text{lin},\ell}^H, w_{\text{lin},\ell}^H \in V^1(\Omega, \mathcal{F}_H)$  as defined in (2.4.8), we have*

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

*Proof.* The proof can be found in Lemma 41 (or [12, Lemma 8]), we apply it for all numerically homogenized tensors  $a_K^0(x_{K_\ell})$  corresponding to the  $\ell = 1, \dots, \mathcal{L}$  sampling domains as before.  $\square$

**Corollary 52.** *Formula (7.2.6) holds.*

*Proof.* Immediately follows from Lemma 51.  $\square$

**Lemma 53.** *Consider the FE-HMM bilinear form (7.2.10) and a QF  $\{x_{K_\ell}, \omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$  that is exact for polynomials of degree  $2p-2$  (i.e. (Q2) with  $\sigma = 2p-2$  holds). Then, we have the following chain of equalities:*

$$\begin{aligned} B(v^H, w^H) &= \sum_{K \in \mathcal{F}_H} \sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K_{\delta_\ell}|} \int_{K_{\delta_\ell}} a^\varepsilon(x) \nabla v_{K_\ell}^h \cdot \nabla w_{K_\ell}^h dx \\ &= \sum_{K \in \mathcal{F}_H} \sum_{\ell=1}^{\mathcal{L}} \frac{\omega_{K_\ell}}{|K|} \int_K a_K^0(x_{K_\ell}) \nabla v_{\text{lin},\ell}^H \cdot \nabla w_{\text{lin},\ell}^H dx \\ &= \sum_{K \in \mathcal{F}_H} \sum_{\ell=1}^{\mathcal{L}} \omega_{K_\ell} a_{K_\ell}^0 \nabla v^H(x_{K_\ell}) \cdot \nabla w^H(x_{K_\ell}) \\ &= \sum_{K \in \mathcal{F}_H} \sum_{\ell=1}^{\mathcal{L}} \omega_{K_\ell} \left( \Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x) \right) \cdot \nabla w^H(x_{K_\ell}) \\ &= \sum_{K \in \mathcal{F}_H} \int_K \left( \Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x) \right) \cdot \nabla w^H dx. \end{aligned}$$

*Proof.* Follows immediately from Lemma 51, (7.2.6) and the definition of the higher order flux (7.2.7) and the exactness of the numerical quadrature (Q2), where  $\Pi_{a^\varepsilon \nabla v_{K_\ell}^h}(x) \in (\mathcal{P}^{p-1}(K))^d$ .  $\square$

This completes the required tools we need in order to extend the FE-HMM to higher order polynomial degrees.

### 7.3 Main results

We now present the main results of the adaptive FE-HMM, extended to higher orders. In what follows we only point out differences compared to the piecewise linear case. For more details

on the algorithm, we refer to Chapter 6.

Let  $f^H$  be an approximation of  $f$  in the space  $\{g^H \in L^2(\Omega); g^H|_K \in \mathcal{P}^m(K), \forall K \in \mathcal{T}_H\}$ . The choice of  $m$  will depend on  $p$  and one usually chooses  $m = p - 1$ .

**Definition 54.** The *local error indicator*  $\eta_H(K)$  on an element  $K$  is defined by

$$\eta_H(K)^2 := H_K^2 \left\| f^H + \nabla \cdot \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right\|_{L^2(K)}^2 + \frac{1}{2} \sum_{e \in \partial K} H_e \left\| \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right\|_{L^2(e)}^2. \quad (7.3.1)$$

The *data approximation error*  $\xi_H(K)$  on an element  $K$  is defined by

$$\xi_H(K)^2 := H_K^2 \left\| f^H - f \right\|_{L^2(K)}^2 + \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)}^2, \quad (7.3.2)$$

where  $a^0(x)$  is the unknown homogenized tensor of problem (2.4.2).

In what follows we will sometimes consider the indicators  $\eta_H(K)$  and data approximation terms  $\xi_H(K)$  on a subset  $\omega = K_{i_1} \cup K_{i_2} \cup \dots \cup K_{i_n}$ ,  $K_{i_j} \in \mathcal{T}_H$  of the domain  $\Omega$ . We will denote the summation of these quantities over all elements  $K \subset \omega$  as  $\eta_H(\omega)^2$  and  $\xi_H(\omega)^2$ .

The following theorem gives an *a posteriori upper bound* for the error between the macroscopic FE-HMM solution  $u^H$  and the homogenized solution  $u^0$ .

**Theorem 55** (*A posteriori upper bound*). *There exists a constant  $C > 0$  depending only on the shape regularity constant  $\gamma$ , the coercivity and continuity bound (2.1.2) and the dimension  $d$  such that*

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

We will present the proof in Section 7.4.1. The next result gives an *a posteriori lower bound*.

**Theorem 56** (*A posteriori lower bound*). *There exists a constant  $C > 0$  depending only on the shape regularity constant  $\gamma$ , the coercivity and continuity bound (2.1.2) and the dimension  $d$  such that*

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

where the domain  $\omega_K$  consists of all elements sharing at least one side with  $K$ .

The proof will be shown in Section 7.4.2.

**Comparison with single scale results.** We suppose that  $a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y)$  is  $Y$ -periodic in  $y$  and  $a(\cdot, y)|_K$  is constant for  $K \in \mathcal{T}_H$ , i.e. the homogenized tensor  $a^0(x)$  is piecewise constant in any  $K \in \mathcal{T}_H$ . We further assume that exact micro functions are used in (2.4.7) and that  $\delta = \varepsilon$  with periodic micro boundary conditions, then  $a_K^0 = a^0(x)|_K$  and we recover again the usual (classical) residual based indicator and data estimator as in [113, 33].

**Algorithm.** The algorithm for the higher order adaptive FE-HMM follows the algorithm for the piecewise linear case presented in Chapter 6.2.1 consisting in the cycle

Solve → Estimate → Mark → Refine.

The micro-macro refinement strategy accounts for the higher macro FE polynomial degree according to Theorem 57.

### 7.3.1 Macro-micro refinement

In the same manner as in the piecewise linear case, the above two theorems for the *a posteriori* lower and upper bounds do not require any structural assumption on the oscillating tensor (such as periodicity) and only minimal assumptions on regularity. We assume singularities to appear in the macro scale and thus do not consider explicit *a posteriori* estimates for the micro-problem (2.4.7). Again, as the error indicator  $\eta_H$  *does* depend on the micro solutions and hence on the micro mesh, we need a new criterion which determines how to choose the size of the micro mesh when we refine a macro element. We can adapt Theorem 35 to the higher order case as follows

**Theorem 57** (Micro-macro refinement coupling). *Let Assumption 61 hold. Assume further that the cell problem (2.4.7) is solved with periodic boundary conditions if  $a^\varepsilon$  is periodic and  $\delta/\varepsilon \in \mathbb{N}$ ; the cell problem (2.4.7) is solved with Dirichlet boundary conditions otherwise. We furthermore consider the case where  $a^\varepsilon = a(x, x/\varepsilon) = a(x, y)$  is  $Y$ -periodic in  $y$  and  $a(\cdot, y)|_K$  is constant. Then*

$$\sup_{x \in K} \left\| a^0(x) \nabla u^H(x) - \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right\|_{L^2(K)} \leq C \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \quad (7.3.3)$$

where  $C$  is independent of  $H, h, \varepsilon$ , and  $err_{mod}$  is independent of  $H, h$ .

The proof will be provided in Section 7.5.

*Remark 58.* From equation (7.3.3) we can deduce the ideal coupling (in the sense of minimizing the macro and micro error at the same rate). The limitation to the case where  $a^0(x)$  is constant on  $K$  leads to an estimate, which does not explicitly depend on  $H$ . As we want the data approximation error to be of the same size as the error

$$\|u^0 - u^{HMM}\|_{H^1(\Omega)} \leq C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right)$$

itself, it follows that we refine the micro mesh in each sampling domain  $K_{\delta_\ell}$  according to  $\frac{h_K}{\varepsilon} \propto (H_K)^{\frac{p}{2q}}$ , where  $H_K$  is the size of macro element  $K \in \mathcal{T}_H$ . The size  $h_K := h(K)$  may vary in different macro elements  $K$ , but is kept constant within  $K$  for all of the  $\mathcal{L}$  sampling domains  $K_{\delta_\ell}$ . We can rewrite the micro mesh size  $h_K$  as a scaled (i.e., independent of  $\varepsilon$ ) micro mesh size  $\hat{h}_K = \frac{h_K}{\varepsilon} = (N_{mic})^{-1/d}$ , where  $N_{mic}$  denotes the degrees of freedom of the micro problem on  $K_{\delta_\ell}$  and  $d$  is the spatial dimension. Assuming appropriate regularity on the micro problem,

we have that as  $q$  gets larger  $\frac{h_K}{\varepsilon} \propto 1$  or  $\hat{h}_K = (N_{mic})^{-1/d} \propto 1$ , hence almost no micro DOF are used for the micro problem. This approach is followed in [9], where a spectral method is used in the micro problems.

Standard *a posteriori* methods could be used on the micro problems to refine the micro meshes differently in the various sampling domains  $K_{\delta_\ell}$  (depending for example on different contrast of the small scale tensor). Accordingly, the micro mesh could also be refined in a non-uniform way.

**The pitfalls and remedy of following the asymptotic refinement.**

While the coupling in Theorem 57 is asymptotically correct, strictly keeping this strategy might (depending on the order of the macro and micro FE  $p$  and  $q$ , respectively) quickly lead to micro problems with an overwhelming number of degrees of freedom. We will see this happening in the numerical experiment in Section 7.6.1, where we start with a micro problem of  $N_{mic} = 8^2$  DOF and after only 20 iteration cycles we reach  $(4, 194, 304)^2$  DOF. Ideally, we would like the macro and micro error to be of the same size, not of the same asymptotic convergence rate, i.e. the proportionality factor might be obstructive.

To overcome this problem, we could apply *a posteriori* estimates on the micro problem. An alternative approach is the use of a reduced basis approach for the solution of the micro problems. There, a few representative basis functions with a large information content are computed once at an offline stage. Typically, only a few reduced basis functions are needed if there is some regularity w.r.t. the slow parameter. In the online stage, the reduced basis functions are used to approximate the solution of the micro problems with high accuracy. As the conductivity tensor of the micro problems will not change significantly from one macro element to its neighbors (provided regularity w.r.t. the slow variable), this approach can lead to a remarkable reduction in computation cost. See [95, 101, 38] for an overview of reduced basis methods and [37] for an application in context of homogenization. We mention the recent effort of using reduced basis techniques in context of the FE-HMM [8].

**7.4 Proof of the main results**

We define the error as  $e^H := u^0 - u^H$ , where  $u^0(x)$  is the homogenized solution of (2.1.5) and  $u^H(x)$  is the FE-HMM solution of (7.1.2).

We first obtain the following *error representation formula*, which is the crucial component to relate the homogenized bilinear form  $B_0(\cdot, \cdot)$  as defined in (2.1.7) to the FE-HMM solution  $u^H$ . We remind the reader that the FE-HMM bilinear form is denoted as  $B(\cdot, \cdot)$  and is defined in (2.4.6).

**Lemma 59** (Error representation formula). *For all  $v \in H_0^1(\Omega)$ , we have*



$$\begin{aligned}
 B_0(e^H, v) &= \int_{\Omega} f v dx - \sum_{e \in \mathcal{E}_H} \int_e \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \llbracket e \rrbracket(s) v ds \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) v dx \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla v dx
 \end{aligned} \tag{7.4.1}$$

where we used the multiscale flux  $\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x)$  as defined in (7.2.7), the multiscale jump as defined in (7.2.9), and where  $u^H$  is the solution of (7.1.2).

*Proof.* We integrate the homogenized bilinear form by parts to obtain

$$\begin{aligned}
 &B_0(e^H, v) \\
 &= B_0(u^0, v) - B_0(u^H, v) \\
 &= \int_{\Omega} f v dx - \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \cdot \nabla v dx \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla v dx \\
 &= \int_{\Omega} f v dx - \sum_{K \in \mathcal{T}_H} \int_{\partial K} \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \cdot n v ds \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) v dx \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla v dx \\
 &= \int_{\Omega} f v dx \\
 &\quad - \sum_{e \in \mathcal{E}_H} \int_e \underbrace{\left( \left( \Pi_{a^\varepsilon \nabla u_{K_\ell^+}^h}(s) - \Pi_{a^\varepsilon \nabla u_{K_\ell^-}^h}(s) \right) \cdot n_e \right)}_{(*)} v ds \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) v dx \\
 &\quad + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla v dx.
 \end{aligned}$$

By noticing that  $(*) = \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \llbracket e \rrbracket(s)$ , the proof is complete.  $\square$

#### 7.4.1 Upper bound (Proof of Theorem 55)

We now prove the *a posteriori* upper bound. In what follows,  $I^H$  stands for the Clément interpolation operator (see (6.3.1), (6.3.2)) and  $f^H$  denotes a  $\mathcal{P}^m$  approximation of  $f$  over  $\mathcal{T}_H$ . We consider the error representation formula (7.4.1) and choose the test function  $v := e^H$ . We

have for the FE-HMM bilinear form that

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

and therefore using (7.4.1)

$$\begin{aligned} B_0(e^H, e^H) &= \int_{\Omega} f e^H dx - \sum_{e \in \mathcal{E}_H} \int_e \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} e(s) e^H ds \\ &\quad + \sum_{K \in \mathcal{T}_H} \int_K \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) e^H dx \\ &\quad + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla e^H dx \\ &\quad + B(u^H, I^H e^H) - \sum_{K \in \mathcal{T}_H} \int_K f(I^H e^H) dx. \end{aligned} \tag{7.4.2}$$

From Lemma 53 we get, as  $I^H e^H \in V^p(\Omega, \mathcal{T}_H)$

$$\begin{aligned} B(u^H, I^H e^H) &= \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \cdot \nabla (I^H e^H) dx, \\ &= - \sum_{K \in \mathcal{T}_H} \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) (I^H e^H) dx \\ &\quad + \sum_{e \in \mathcal{E}_H} \int_e \left( \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} e(s) \right) (I^H e^H) ds. \end{aligned} \tag{7.4.3}$$

We insert (7.4.3) into (7.4.2) to obtain

$$\begin{aligned} &B_0(e^H, e^H) \\ &= \int_{\Omega} f^H (e^H - I^H e^H) dx + \int_{\Omega} (f - f^H) (e^H - I^H e^H) dx \\ &\quad - \sum_{e \in \mathcal{E}_H} \int_e \left( \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} e(s) \right) (e^H - I^H e^H) ds \\ &\quad + \sum_{K \in \mathcal{T}_H} \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) (e^H - I^H e^H) dx \\ &\quad + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla e^H dx. \end{aligned}$$

Let  $\phi^H := e^H - I^H e^H$ . Using the Cauchy-Schwarz inequality, we estimate

$$\begin{aligned} B_0(e^H, e^H) &\leq \left( \sum_{K \in \mathcal{T}_H} \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)} \|\phi^H\|_{L^2(K)} \right. \\ &\quad + \sum_{K \in \mathcal{T}_H} \|f - f^H\|_{L^2(K)} \|\phi^H\|_{L^2(K)} \\ &\quad + \sum_{e \in \mathcal{E}_H} \left\| \overline{\left[ \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \right]}_e(s) \right\|_{L^2(e)} \|\phi^H\|_{L^2(e)} \\ &\quad \left. + \sum_{K \in \mathcal{T}_H} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)} \|\nabla e^H\|_{L^2(K)} \right). \end{aligned}$$

With the Clément interpolation estimates (6.3.1) and (6.3.5), it follows that

$$\begin{aligned} B_0(e^H, e^H) &\leq C \left( \sum_{K \in \mathcal{T}_H} H_K \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)} \|\nabla e^H\|_{L^2(N(K))} \right. \\ &\quad + \sum_{K \in \mathcal{T}_H} H_K \|f - f^H\|_{L^2(K)} \|\nabla e^H\|_{L^2(N(K))} \\ &\quad + \sum_{e \in \mathcal{E}_H} H_e^{1/2} \left\| \overline{\left[ \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \right]}_e(s) \right\|_{L^2(e)} \|\nabla e^H\|_{L^2(N(e))} \\ &\quad \left. + \sum_{K \in \mathcal{T}_H} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)} \|\nabla e^H\|_{L^2(K)} \right). \end{aligned}$$

We finally use the finite overlapping property of the neighborhoods  $N(K_e)$  which allows us to estimate  $\sum_{K \in \mathcal{T}_H} \|\nabla e^H\|_{L^2(N(K))}^2 \leq C \sum_{K \in \mathcal{T}_H} \|\nabla e^H\|_{L^2(K)}^2$ , where  $C$  only depends on the shape regularity of the triangulation and the dimension  $d$ . With the triangle inequality and the coercivity of  $B_0(\cdot, \cdot)$ , we arrive at the upper *a posteriori* estimate

$$\begin{aligned} \|\nabla e^H\|_{L^2(\Omega)}^2 &\leq C \left( \sum_{K \in \mathcal{T}_H} H_K^2 \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)}^2 \right. \\ &\quad + \sum_{K \in \mathcal{T}_H} H_K^2 \|f - f^H\|_{L^2(K)}^2 \\ &\quad + \sum_{e \in \mathcal{E}_H} H_e \left\| \overline{\left[ \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \right]}_e(s) \right\|_{L^2(e)}^2 \\ &\quad \left. + \sum_{K \in \mathcal{T}_H} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)}^2 \right), \end{aligned}$$

where we used the coercivity of  $B_0(\cdot, \cdot)$ . Applying Poincaré inequality and using definitions (7.3.1) and (7.3.2) leads to

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

as stated in Theorem 55.

### 7.4.2 Lower bound (Proof of Theorem 56)

The derivation of the lower bound will involve bubble functions in a space  $\tilde{V}_H^p \supset V_H^p$ , defined over a refinement  $\tilde{\mathcal{T}}_H$  of  $\mathcal{T}_H$ . We choose the refinement  $\tilde{V}_H^p$  such that every  $K \in \mathcal{T}_H$  has an interior node  $\tilde{x}_K \in K \setminus \partial K$  in  $\tilde{\mathcal{T}}_H$  and every edge  $e$  which does not lie on the boundary  $\partial\Omega$  has an interior node in  $\tilde{\mathcal{T}}_H$ . We proceed by estimating the interior residual and the jump residual.

**Interior residual.** For any  $K \in \mathcal{T}_H$  we consider an interior bubble function, i.e., a function  $\psi_K \in \tilde{V}_H^1$  such that  $0 \leq \psi_K \leq 1$ ,  $\psi_K(\tilde{x}_K) = 1$  and  $\psi_K \equiv 0$  on  $\Omega \setminus K$ . We choose the appropriate test function  $v_K(x) \in H_0^1(\Omega)$  as  $v_K(x) := \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right]$ . We therefore have that

$$v_K(x) := \begin{cases} \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] & \text{in } K, \\ 0 & \text{in } \Omega \setminus K, \\ 0 & \text{on } \partial K. \end{cases}$$

We insert the test function  $v_K(x)$  into the representation formula (7.4.1) and obtain

$$\begin{aligned} & \int_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \left( \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \right) dx \\ &= B_0 \left( e^H, \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \right) \Big|_K \\ & \quad - \int_K (f - f^H) \left( \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \right) dx \\ & \quad - \int_K \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \cdot \nabla \left( \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \right) dx. \end{aligned}$$

First, the equivalence of norms on a finite dimensional space (see [19, Theorem 2.2] for details) yields

$$\begin{aligned} & C \left\| f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)}^2 \\ & \leq \int_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \left( \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^e \nabla u_{K_\ell}^h}(x) \right) \right] \right) dx. \end{aligned}$$

Second, we use the continuity of  $B_0(\cdot, \cdot)$ , the Cauchy-Schwarz inequality and the inverse inequality (6.3.4) to arrive at

$$\begin{aligned}
 & \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)}^2 \\
 & \leq C \left( \|\nabla e^H\|_{L^2(K)} \left\| \nabla \left( \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right] \right) \right\|_{L^2(K)} \right. \\
 & \quad + \|f - f^H\|_{L^2(K)} \left\| \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right] \right\|_{L^2(K)} \\
 & \quad \left. + \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)} \left\| \nabla \left( \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right] \right) \right\|_{L^2(K)} \right) \\
 & \leq C \left( H_K^{-1} \|\nabla e^H\|_{L^2(K)} + \|f - f^H\|_{L^2(K)} \right. \\
 & \quad \left. + H_K^{-1} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)} \right) \left\| \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right] \right\|_{L^2(K)}.
 \end{aligned}$$

Finally, we use  $\left\| \psi_K \left[ f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right] \right\|_{L^2(K)} \leq \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)}$ , which holds as  $0 \leq \psi_K \leq 1$ , in order to obtain the *interior residual*

$$\begin{aligned}
 & H_K^2 \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)}^2 \\
 & \leq C \left( \|\nabla e^H\|_{L^2(K)}^2 + H_K^2 \|f - f^H\|_{L^2(K)}^2 \right. \\
 & \quad \left. + \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)}^2 \right). \tag{7.4.4}
 \end{aligned}$$

**Jump residual.** Let  $e \in \mathcal{E}_H$  be an interior interface and let  $K_1 \in \mathcal{T}_H$  and  $K_2 \in \mathcal{T}_H$  be such that  $K_1 \cap K_2 = e$ . We rearrange (7.4.1) for  $\sum_{e \in \mathcal{E}_H} \int_e \overline{[\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}]} e(s) v ds$  and add and subtract  $\int_\Omega f^H v dx$ . The continuity of  $B_0(\cdot, \cdot)$  and the Cauchy-Schwarz inequality lead to

$$\begin{aligned}
 & \sum_{e \in \mathcal{E}} \int_e \overline{\llbracket \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \rrbracket}_e(s) v \, ds \\
 &= \int_\Omega \left( f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right) v \, dx - B_0(e^H, v) \\
 & \quad + \int_\Omega (f - f^H) v \, dx + \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right) \nabla v \, dx \\
 &\leq C \left( \sum_{K \in \mathcal{T}_H} \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)} \|v\|_{L^2(K)} \right. \\
 & \quad + \sum_{K \in \mathcal{T}_H} \|\nabla e^H\|_{L^2(K)} \|\nabla v\|_{L^2(K)} + \sum_{K \in \mathcal{T}_H} \|f - f^H\|_{L^2(K)} \|v\|_{L^2(K)} \\
 & \quad \left. + \sum_{K \in \mathcal{T}_H} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K)} \|\nabla v\|_{L^2(K)} \right). \tag{7.4.5}
 \end{aligned}$$

We then need a lifting operator as described in the following Lemma (see [33] for details and a proof).

**Lemma 60.** *Let  $\mathcal{P}_0^{p+d+1}(e)$  be the space of polynomials  $\mathcal{P}^{p+d+1}(e)$  that vanish on  $\partial e$ . Then for every element  $K \in \mathcal{T}_H$  and every interface  $e \subset \partial K$  there exists an operator  $\mathcal{R}_{K,e} : \varphi \mapsto \mathcal{R}_{K,e}(\varphi)$ ,  $\mathcal{P}_0^{p+d+1}(e) \rightarrow \mathcal{P}^{p+d+1}(K)$  such that for each function  $\varphi \in \mathcal{P}_0^{p+d+1}(e)$  we have*

$$i) \quad \mathcal{R}_{K,e}(\varphi) \equiv \varphi \text{ on } e$$

$$ii) \quad \mathcal{R}_{K,e}(\varphi) \equiv 0 \text{ on } \partial K \setminus e$$

and the following estimate holds

$$\|\nabla \mathcal{R}_{K,e}(\varphi)\|_{L^2(K)} + H_K^{-1} \|\mathcal{R}_{K,e}(\varphi)\|_{L^2(K)} \leq C H_e^{-1/2} \|\varphi\|_{L^2(e)}. \tag{7.4.6}$$

Let  $e$  be an interior interface of  $K_1$  and  $K_2$  and let  $x_e \in e$  be an interior node on that interface. Let  $\psi_e \in \tilde{V}^1(\Omega)$  be an edge bubble function in  $\omega_e = K_1 \cup K_2$  which satisfies  $\psi_e(x_e) = 1$ ,  $\psi_e|_{\partial \omega_e} = 0$ ,  $0 \leq \psi_e \leq 1$  and especially  $\text{supp}(\psi_e) = \omega_e$ . We choose the test function as

$$v_e(x) = \begin{cases} \mathcal{R}_{K_1,e} \left( \overline{\llbracket \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \rrbracket}_e(s) \psi_e \right) & \text{in } K_1 \\ \mathcal{R}_{K_2,e} \left( \overline{\llbracket \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \rrbracket}_e(s) \psi_e \right) & \text{in } K_2 \\ 0 & \text{in } \Omega \setminus (K_1 \cup K_2). \end{cases} \tag{7.4.7}$$

By definition,  $\psi_e$  vanishes on  $\partial e$ , therefore  $v_e(x)$  vanishes on  $\partial(K_1 \cup K_2)$ . A direct consequence of Lemma 60 are the following estimates

$$\|\nabla v_e\|_{L^2(K)} \leq C H_e^{-1/2} \left\| \overline{\llbracket \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \rrbracket}_e(s) \psi_e \right\|_{L^2(e)}$$

and

$$H_K^{-1} \|v_e\|_{L^2(K)} \leq CH_e^{-1/2} \left\| \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \psi_e \right\|_{L^2(e)}.$$

Inserting the test function (7.4.7) for any  $e$  into (7.4.5) and using  $0 \leq \psi_e \leq 1$ , we obtain

$$\begin{aligned} & \int_e \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) v_e ds \\ & \leq CH_e^{-1/2} \left( \sum_{K_1, K_2} H_{K_i} \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K_i)} \right. \\ & \quad + \sum_{K_1, K_2} \|\nabla e^H\|_{L^2(K_i)} + \sum_{K_1, K_2} H_{K_i} \|f - f^H\|_{L^2(K_i)} \\ & \quad \left. + \sum_{K_1, K_2} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K_i)} \right) \left\| \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \right\|_{L^2(e)}. \end{aligned} \quad (7.4.8)$$

From the equivalence of norms in finite dimensional spaces follows that

$$\begin{aligned} & C \left\| \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \right\|_{L^2(e)}^2 \\ & \leq \left\| \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \psi_e^{1/2} \right\|_{L^2(e)}^2 = \int_e \left( \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \right)^2 \psi_e ds, \end{aligned}$$

which we insert into (7.4.8) to arrive at the jump residual estimate

$$\begin{aligned} & H_e \left\| \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \right\|_{L^2(e)}^2 \\ & \leq C \left( \sum_{K_1, K_2} H_{K_i}^2 \left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K_i)}^2 \right. \\ & \quad + \sum_{K_1, K_2} \|\nabla e^H\|_{L^2(K_i)}^2 + \sum_{K_1, K_2} H_{K_i}^2 \|f - f^H\|_{L^2(K_i)}^2 \\ & \quad \left. + \sum_{K_1, K_2} \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(K_i)}^2 \right). \end{aligned} \quad (7.4.9)$$

**Combining interior and jump residuals.** We eliminate  $\left\| f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \right\|_{L^2(K)}$  from the jump residual (7.4.9) with the help of the interior residual (7.4.4) to obtain

$$\begin{aligned} H_e \left\| \left[ \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \right]_e(s) \right\|_{L^2(e)}^2 & \leq C \left( \|\nabla e^H\|_{L^2(\omega_e)}^2 + H_{\omega_e}^2 \|f - f^H\|_{L^2(\omega_e)}^2 \right. \\ & \quad \left. + \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^H \right\|_{L^2(\omega_e)}^2 \right), \end{aligned} \quad (7.4.10)$$

where  $H_{\omega_e} = \max_{i=1,2} H_{K_i}$ . Adding the interior residual (7.4.4) to (7.4.10) provides the desired lower bound

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

as stated in Theorem 56.

## 7.5 Proof of Theorem 57 (Estimation of the data approximation error)

In this Section we will estimate the data approximation error and obtain the results given in Theorem 57.

The term  $\|f^H - f\|_{L^2(K)}$  of the data approximation error (7.3.2) can be easily estimated using classical FEM techniques, see [44, 40, 39]. We obtain

$$\|f^H - f\|_{L^2(K)} \leq CH^{m+1}.$$

One typically chooses  $m = p - 1$  such that

$$\|f^H - f\|_{L^2(K)} \leq CH^p.$$

Let us introduce the tensor  $\bar{a}_K^0(x_{K_\ell})$ , which is defined as  $a_K^0(x_{K_\ell})$  given in (7.2.5) but where the corresponding functions  $\psi_{K_{\delta_\ell}}^i$  are found in the exact Sobolev space  $W(K_{\delta_\ell})$  instead of its FE approximation space  $S^q(K_{\delta_\ell}, \mathcal{T}_h)$ . We decompose the term

$$\left\| a^0(x) \nabla u^H - \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right\|_{L^2(K)}$$

of the data approximation error (7.3.2) as follows

$$\left\| a^0(x) \nabla u^H(x) - \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right\|_{L^2(K)} \tag{7.5.1}$$

$$= \left\| a^0(x) \nabla u^H(x) - \Pi_{a_{K_\ell}^0 \nabla u^H}(x) \right\|_{L^2(K)} \tag{7.5.2}$$

$$\leq \underbrace{\left\| a^0(x) \nabla u^H(x) - \Pi_{a^0(x_{K_\ell}) \nabla u_{K_\ell}^H}(x) \right\|_{L^2(K)}}_{err_{mac}} \tag{7.5.3}$$

$$+ \underbrace{\left\| \Pi_{a^0(x_{K_\ell}) \nabla u_{K_\ell}^H}(x) - \Pi_{\bar{a}_{K_\ell}^0 \nabla u_{K_\ell}^H}(x) \right\|_{L^2(K)}}_{err_{mod}} \tag{7.5.4}$$

$$+ \underbrace{\left\| \Pi_{\bar{a}_{K_\ell}^0 \nabla u_{K_\ell}^H}(x) - \Pi_{a_{K_\ell}^0 \nabla u_{K_\ell}^H}(x) \right\|_{L^2(K)}}_{err_{mic}}, \tag{7.5.5}$$



## 7.5 Proof of Theorem 57 (Estimation of the data approximation error)

where  $x_{K_\ell} \in K$  are the quadrature nodes of element  $K$ , and  $\Pi_{a^0(x_{K_\ell})\nabla u_{K_\ell}^H}(x)$ ,  $\Pi_{\bar{a}_{K_\ell}^0\nabla u_{K_\ell}^H}(x)$  and  $\Pi_{a_{K_\ell}^0\nabla u_{K_\ell}^H}(x)$  are the interpolating polynomials based on the interpolation nodes  $x_{K_\ell}$  with function values  $a^0(x_{K_\ell})\nabla u^H(x_{K_\ell})$ ,  $\bar{a}_{K_\ell}^0(x_{K_\ell})\nabla u^H(x_{K_\ell})$ , and  $a_{K_\ell}^0(x_{K_\ell})\nabla u^H(x_{K_\ell})$ , respectively.

In the above equation,  $err_{mac}$  and  $err_{mic}$  denote the macroscopic and microscopic error, respectively and their analysis depends on both the finescale and homogenized tensor. While the modeling error  $e_{mod}$  is independent of the discretization parameters such as  $H$  and  $h$ , it does depend on the structure of  $a^\varepsilon$  such as periodicity, etc.

### Assumptions

As in the piecewise linear case (see Remark 43), we need to have an appropriate regularity of the cell functions  $\psi_{K_{\delta_\ell}}^i$ . Here  $\psi_{K_{\delta_\ell}}^i \in W(K_{\delta_\ell})$  are the exact cell functions, where  $W(K_{\delta_\ell})$  is defined as in (2.4.10) or (2.4.11). The exact cell functions  $\psi_{K_{\delta_\ell}}^i$  are defined as the discrete cell functions  $\psi_{K_{\delta_\ell}}^{i,h}$ , solutions of equation (7.2.4), but are chosen in  $W(K_{\delta_\ell})$  instead of  $S_h^q(K_{\delta_\ell}, \mathcal{T}_h)$ . We need further regularity and growth conditions on the small scale tensor  $a^\varepsilon$ , which read as follows for Dirichlet micro boundary conditions (see (2.4.11)) and periodic micro boundary conditions (see (2.4.10)). We refer to [16, Section 3] for a detailed discussion.

**Dirichlet coupling**,  $W(K_{\delta_\ell}) = H_0^1(K_{\delta_\ell})$ . We assume that  $a_{ij}^\varepsilon|_K \in W^{1,\infty}(K) \forall K \in \mathcal{T}_H$  and that

$$\left| a_{ij}^\varepsilon \right|_{W^{1,\infty}(K)} \leq C\varepsilon^{-1}. \text{ For the case } q = 1 \text{ we can use classical } H^2 \text{ regularity results by applying the second fundamental inequality for elliptic operators (see [72, Chap. 2.6]) to the cell problem. We find that we have the regularity } \left| \psi_{K_{\delta_\ell}}^i \right|_{H^2(K_{\delta_\ell})} \leq C\varepsilon^{-1} \sqrt{|K_{\delta_\ell}|} \text{ provided that } \psi_{K_{\delta_\ell}}^i \in H^2(K_{\delta_\ell}).$$

**Periodic coupling**,  $W(K_{\delta_\ell}) = H_{per}^1(K_{\delta_\ell})$ . We assume that  $a_{ij}^\varepsilon$  is locally periodic and that  $\delta/\varepsilon \in \mathbb{N}$ . We further assume that the slow variable  $x$  is collocated at  $x_K$  in each sampling domain, i.e., we use a tensor of form  $a(x_K, y) = a(x_K, \frac{x}{\varepsilon})$ , where  $y$  is  $Y$ -periodic and where the map  $x \mapsto a_{ij}(x, \cdot)$  is Lipschitz continuous and bounded from  $\bar{\Omega} \times \mathbb{R}$  into  $W_{per}^{1,\infty}(Y)$ . If  $a^\varepsilon$  is sufficiently smooth, then higher regularity  $\left| \psi_{K_{\delta_\ell}}^i \right|_{H^{q+1}(K_{\delta_\ell})} \leq C\varepsilon^{-q} \sqrt{|K_{\delta_\ell}|}$  of  $\psi_{K_{\delta_\ell}}^i$  can be shown (see [34, Chap. 3]).

It is more convenient, though, to assume the regularity conditions directly on the cell functions instead for both Dirichlet and periodic coupling.

**Assumption 61.** We assume that the cell functions  $\psi_{K_{\delta_\ell}}^i \in W(K_{\delta_\ell})$  satisfy

$$\left| \psi_{K_{\delta_\ell}}^i \right|_{H^{q+1}(K_{\delta_\ell})} \leq C\varepsilon^{-q} \sqrt{|K_{\delta_\ell}|}, \quad (7.5.6)$$

where  $C$  is independent of  $\varepsilon$ , the quadrature node  $x_{K_\ell}$  and the domain  $K_{\delta_\ell}$  for all  $i = 1, \dots, d$ .

For simplicity, we furthermore make the following assumption.

**Assumption 62.** We assume that  $a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y)$  is  $Y$ -periodic in  $y$  and

$$a(\cdot, y)|_K \text{ is constant} \quad (7.5.7)$$

for  $K \in \mathcal{T}_H$ .

Assumption 62 implies that the homogenized tensor  $a^0$  is constant in any  $K \in \mathcal{T}_H$ .

### 7.5.1 Macro error

Due to Assumption 62, we have that  $a^0(x) = \text{const}$  in any  $K \in \mathcal{T}_H$ . Therefore the interpolation (7.2.7) is exact, i.e.  $\Pi_{a^0(x_{K_\ell})\nabla u_{K_\ell}^H}(x) \equiv a^0(x)\nabla u^H(x)$  for  $x \in K$ , and we have for the macro error

$$\text{err}_{mac} = 0.$$

### 7.5.2 Modeling and micro error

For the estimation of the modeling error  $\|\Pi_{a^0(x_{K_\ell})\nabla u_{K_\ell}^H}(x) - \Pi_{\bar{a}_{K_\ell}^0\nabla u_{K_\ell}^H}(x)\|_{L^2(K)}$  and micro error  $\|\Pi_{\bar{a}_{K_\ell}^0\nabla u_{K_\ell}^H}(x) - \Pi_{a_{K_\ell}^0\nabla u_{K_\ell}^H}(x)\|_{L^2(K)}$  we relate the error of two interpolating polynomials to the error of the function values at the corresponding interpolation nodes. This relation will allow us to use the estimates of the data approximation error obtained for the piecewise linear adaptive FE-HMM presented in Section 6.3.5.

**Lemma 63.** For  $v, \bar{v} \in C^0(K)$  we consider the interpolation polynomials  $\Pi_{v_\ell}(x) \in \mathcal{P}^{p-1}(K)$  and  $\Pi_{\bar{v}_\ell}(x) \in \mathcal{P}^{p-1}(K)$ , respectively, which are defined by (7.2.3) and are assumed to satisfy (Q2) (see (2.4.5)), i.e.,

$$\Pi_{v_\ell}(x_{K_\ell}) = v(x_{K_\ell}), \quad \ell = 1, \dots, \mathcal{L}$$

and

$$\Pi_{\bar{v}_\ell}(x_{K_\ell}) = \bar{v}(x_{K_\ell}), \quad \ell = 1, \dots, \mathcal{L}.$$

We assume that the interpolation nodes  $x_{K_\ell}$  are chosen to be the quadrature nodes of a quadrature formula  $\{x_{K_\ell}, \omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$  that satisfies condition (7.2.2). Then

$$\|\Pi_{v_\ell}(x) - \Pi_{\bar{v}_\ell}(x)\|_{L^2(K)} \leq C \sqrt{|K|} \max_{1 \leq \ell \leq \mathcal{L}} |v_\ell - \bar{v}_\ell|,$$

where the constant  $C$  only depends on the dimensionality  $d$ , the polynomial degree  $p$ , the shape regularity constant  $\gamma$ , the quadrature formula  $\{\hat{x}_\ell, \hat{\omega}_\ell\}_{\ell=1}^{\mathcal{L}}$  and the basis  $\{\hat{q}_j\}_{j=1}^{\mathcal{L}}$  of  $\mathcal{P}^{p-1}(\hat{K})$  on the reference element  $\hat{K}$ .

*Proof.* From (7.2.3) follows that we can introduce an interpolating polynomial  $\Pi_{v_\ell - \bar{v}_\ell}(x) \in \mathcal{P}^{p-1}(K)$  s.t.

$$\Pi_{v_\ell - \bar{v}_\ell}(x_{K_\ell}) = v(x_{K_\ell}) - \bar{v}(x_{K_\ell}), \quad \ell = 1, \dots, \mathcal{L}. \quad (7.5.8)$$

## 7.5 Proof of Theorem 57 (Estimation of the data approximation error)

We use the transformation  $F_K$  between  $K$  and the reference element  $\hat{K}$  and have

$$\Pi_{v_\ell - \tilde{v}_\ell}(x) = \Pi_{v_\ell - \tilde{v}_\ell}(F_K(\hat{x})) = \widehat{\Pi_{v_\ell - \tilde{v}_\ell}}(\hat{x}). \quad (7.5.9)$$

It follows that [39]

$$\|\Pi_{v_\ell - \tilde{v}_\ell}(x)\|_{L^2(K)} \leq C |\det(\partial F_K)|^{1/2} \|\widehat{\Pi_{v_\ell - \tilde{v}_\ell}}(\hat{x})\|_{L^2(\hat{K})}. \quad (7.5.10)$$

Let  $\{\hat{q}_j\}_{j=1}^{\mathcal{L}}$  be a basis of  $\mathcal{P}^{p-1}(\hat{K})$ , in which we expand  $\widehat{\Pi_{v_\ell - \tilde{v}_\ell}}(\hat{x})$  as

$$\widehat{\Pi_{v_\ell - \tilde{v}_\ell}}(\hat{x}) = \sum_{\ell=1}^{\mathcal{L}} \hat{\beta}_\ell \hat{q}_\ell(\hat{x}) \quad (7.5.11)$$

with coefficients  $\hat{\beta}_\ell \in \mathbb{R}$ ,  $\ell = 1, \dots, \mathcal{L}$ . We evaluate (7.5.9) at  $\hat{x}_\ell = F_K^{-1}(x_{K_\ell})$  and use (7.5.8) to get

$$\sum_{j=1}^{\mathcal{L}} \hat{A}_{\ell j} \hat{\beta}_j = \sum_{j=1}^{\mathcal{L}} \hat{\beta}_j \hat{q}_j(\hat{x}_\ell) = \widehat{\Pi_{v_\ell - \tilde{v}_\ell}}(\hat{x}_\ell) = \Pi_{v_\ell - \tilde{v}_\ell}(x_{K_\ell}) = v(x_{K_\ell}) - \tilde{v}(x_{K_\ell}), \quad (7.5.12)$$

where  $\hat{A}_{\ell j} = \hat{q}_j(\hat{x}_\ell)$ . We write equation (7.5.12) as  $\hat{A}\hat{\beta} = \delta\mathbf{v}$ , where  $(\hat{A})_{ij} = \hat{A}_{ij}$ ,  $(\hat{\beta})_i = \hat{\beta}_i$  and  $(\delta\mathbf{v})_\ell = v(x_{K_\ell}) - \tilde{v}(x_{K_\ell})$ . From Proposition 50 follows that (7.5.12) has a unique solution  $\hat{\beta}$  and that  $\hat{A}$  is invertible. Then we can estimate (7.5.10) further as follows using (7.5.11) and (7.5.12):

$$\begin{aligned} \|\Pi_{v_\ell - \tilde{v}_\ell}(x)\|_{L^2(K)} &\leq C |\det(\partial F_K)|^{1/2} \|\widehat{\Pi_{v_\ell - \tilde{v}_\ell}}(\hat{x})\|_{L^2(\hat{K})} \\ &\leq C \sqrt{H_K} \left\| \sum_{\ell=1}^{\mathcal{L}} \hat{\beta}_\ell \hat{q}_\ell(\hat{x}) \right\|_{L^2(\hat{K})} \\ &\leq C \sqrt{H_K} \sum_{\ell=1}^{\mathcal{L}} |\hat{\beta}_\ell| \|\hat{q}_\ell(\hat{x})\|_{L^2(\hat{K})} \\ &\leq C \sqrt{H_K} \sum_{\ell=1}^{\mathcal{L}} \|\hat{q}_\ell(\hat{x})\|_{L^2(\hat{K})} \|\hat{\beta}\|_\infty \\ &\leq C \sqrt{H_K} \|\hat{A}^{-1}(\delta\mathbf{v})\|_\infty \\ &\leq C \sqrt{H_K} \max_{1 \leq \ell \leq \mathcal{L}} |v(x_{K_\ell}) - \tilde{v}(x_{K_\ell})|, \end{aligned}$$

where  $\|\cdot\|_\infty$  is the supremum matrix and vector norm, respectively, and the constant  $C$  only depends on the dimensionality  $d$ , the polynomial degree  $p$ , the shape regularity constant  $\gamma$ , the quadrature formula  $\{\hat{x}_\ell, \hat{\omega}_\ell\}_{\ell=1}^{\mathcal{L}}$  and the basis  $\{\hat{q}_j\}_{j=1}^{\mathcal{L}}$  of  $\mathcal{P}_d^{p-1}(\hat{K})$  (via  $\|\hat{q}_\ell(\hat{x})\|_{L^2(\hat{K})}$  and  $\|\hat{A}^{-1}\|_\infty$ ) on the reference element  $\hat{K}$ .  $\square$

**Modeling error.** Applying Lemma 63 to (7.5.1) and using the inverse inequality on the space  $L^\infty(K)$  (see [44, Thm. 17.2])

$$\sqrt{|K|} \|\nabla u^H(x)\|_{L^\infty(K)} \leq C \|\nabla u^H(x)\|_{L^2(K)}$$

leads to the following estimate for the modeling error

$$\begin{aligned} err_{mod} &= \left\| \Pi_{a^0(x_{K_\ell})} \nabla u_{K_\ell}^H(x) - \Pi_{\bar{a}_{K_\ell}^0} \nabla u_{K_\ell}^H(x) \right\|_{L^2(K)} \\ &\leq C \sqrt{|K|} \max_{1 \leq \ell \leq \mathcal{L}} \left\| (a^0(x_{K_\ell}) - \bar{a}_{K_\ell}^0(x_{K_\ell})) \nabla u^H(x_{K_\ell}) \right\|_2 \\ &\leq C \sqrt{|K|} \left( \max_{1 \leq \ell \leq \mathcal{L}} \left\| (a^0(x_{K_\ell}) - \bar{a}_{K_\ell}^0(x_{K_\ell})) \right\|_2 \right) \left( \max_{1 \leq \ell \leq \mathcal{L}} \left\| \nabla u^H(x_{K_\ell}) \right\|_2 \right) \end{aligned} \quad (7.5.13)$$

$$\leq C \left( \max_{1 \leq \ell \leq \mathcal{L}} \left\| (a^0(x_{K_\ell}) - \bar{a}_{K_\ell}^0(x_{K_\ell})) \right\|_F \right) \sqrt{|K|} \left\| \nabla u^H(x) \right\|_{L^\infty(K)} \quad (7.5.14)$$

$$\leq C \left( \max_{1 \leq \ell \leq \mathcal{L}} \left\| a^0(x_{K_\ell}) - \bar{a}_{K_\ell}^0(x_{K_\ell}) \right\|_F \right) \left\| \nabla u^H(x) \right\|_{L^2(K)}, \quad (7.5.15)$$

where  $\|\cdot\|_2$  is the Euclidean norm, and where the constant  $C$  only depends on the dimensionality  $d$ , the polynomial degree  $p$ , the shape regularity constant  $\gamma$ , the quadrature formula  $\{\hat{x}_\ell, \hat{\omega}_\ell\}_{\ell=1}^{\mathcal{L}}$  and the basis  $\{\hat{q}_{jj}\}_{j=1}^{\mathcal{L}}$  of  $\mathcal{P}_d^{p-1}(\hat{K})$  (via  $\|\hat{q}_\ell(\hat{x})\|_{L^2(\hat{K})}$  and  $\|\hat{A}^{-1}\|_\infty$ ) on the reference element  $\hat{K}$ . The term (7.5.15) the maximal modeling error over all sampling domains  $K_{\delta_\ell}$  within  $K$ , and can be estimated as described in Section 6.3.5.2.

**Micro error.** For the micro error we similarly get

$$\begin{aligned} err_{mic} &= \left\| \Pi_{\bar{a}_{K_\ell}^0} \nabla u_{K_\ell}^H(x) - \Pi_{a_{K_\ell}^0} \nabla u_{K_\ell}^H(x) \right\|_{L^2(K)} \\ &\leq C \max_{1 \leq \ell \leq \mathcal{L}} \left\| (\bar{a}_{K_\ell}^0(x_{K_\ell}) - a_{K_\ell}^0(x_{K_\ell})) \nabla u^H(x_{K_\ell}) \right\|_2 \\ &\leq C \left( \max_{1 \leq \ell \leq \mathcal{L}} \left\| \bar{a}_{K_\ell}^0(x_{K_\ell}) - a_{K_\ell}^0(x_{K_\ell}) \right\|_F \right) \left\| \nabla u^H(x) \right\|_{L^2(K)}. \end{aligned}$$

We can use the results obtained in Section 6.3.5 in context of the piecewise linear FE-HMM, but here under the Assumption 61, to estimate

$$err_{mic} \leq C \max_{1 \leq \ell \leq \mathcal{L}} \left\| \bar{a}_{K_\ell}^0(x_{K_\ell}) - a_{K_\ell}^0(x_{K_\ell}) \right\|_F \left\| \nabla u^H(x) \right\|_{L^2(K)} \leq C \left( \frac{h}{\varepsilon} \right)^{2q}.$$

## 7.6 Numerical experiments

We reproduce the experiment presented in Section 6.4.2, however we now use piecewise quadratic instead of piecewise linear macro FE.

The numerical experiments were implemented in Matlab and are based on the FE-HMM code presented in [11]; for the element bisection we use code that is based in part on *i*FEM (see [41]). We will follow the notation presented in Section 6.4.

### 7.6.1 Crack problem

We consider a crack problem, which exhibits a singularity on the macro domain. This problem features a two-dimensional, highly oscillating tensor and is based upon [82, Example 5.2]. We investigated the crack problem in the piecewise linear case in Section 7.6.1 (see also [12]).

Here we use piecewise quadratic finite element functions in the macro problem and piecewise linear functions in the micro problems. We consider the problem

$$\begin{aligned} -\nabla \cdot (a^\varepsilon(x) \nabla u^\varepsilon) &= 1 & \text{in } \Omega \\ u^\varepsilon &= g_D & \text{on } \Gamma_D = \partial\Omega \end{aligned}$$

on a domain  $\Omega = \{|x_1| + |x_2| < 1\} \setminus \{0 \leq x_1 \leq 1, x_2 = 0\}$  with a crack along the positive  $x_1$ -axis (see Figure 7.6.1a). The conductivity tensor is given by

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

and has coefficients chosen such that the homogenized tensor coincides with the unit tensor  $I_2$  (see [70, Chap. 1.2]). The exact homogenized solution  $u^0$  of the problem is given (in polar coordinates) by

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

where  $x_1 = r \cos(\vartheta)$ ,  $x_2 = r \sin(\vartheta)$ . We use Dirichlet boundary conditions  $g_D = u^0$  matching the exact homogenized solution  $u^0$ .

We choose  $\varepsilon = \delta = 10^{-3}$  and periodic boundary conditions for the micro problems. Furthermore we let Dörfler's bulk-chasing strategy (Marking Strategy E, see Section 5.4) with a parameter of  $\theta = 0.3$  drive the marking and choose the coupling between micro and macro refinement as  $h_K = H_K$ , where we start with an initial mesh of  $\hat{h}_K = \frac{1}{8}$ . We refine the mesh twice in every iteration (i.e. after two iterations, two elements can be divided in up to eight).

*Remark 64.* If we strictly follow the refinement strategy  $h_K = H_K$  then as soon as in iteration 20 a very large micro problem with  $(4, 194, 304)^2$  DOF is required. At the same time, at iteration 8 the highest resolved numerically homogenized tensor  $a_K^0(x_{K_\ell})$  computed with  $(1024)^2$  DOF is exact up to  $\|a_K^0(x_{K_\ell}) - a^0(x_{K_\ell})\|_\infty < 9.2 \cdot 10^{-7}$ , which is much smaller than the macro error. We therefore keep this micro-resolution and do not refine the micro mesh any further for every macro element already refined at least 8 times. *A posteriori* estimates on the micro problem could be used to choose the micro-macro coupling rate adaptively. See also the Remark at the end of Section 7.3.1.

In Figure 7.6.1 we show the error in the  $H^1$ -norm. There,  $N_{mac}$  denotes the macro degrees of freedom. Our experiments confirm the theoretical asymptotic rate of  $\mathcal{O}(N_{mac}^{-2/d})$ . As a comparison we also plot the error when using  $h_K = \sqrt{H_K}$  for coupling micro and macro meshes. We get a convergence rate that is not asymptotically optimal, therefore confirming our ideal coupling rate of Remark 58. We remark that, despite being expensive, the coupling  $h_K \propto H_K$  is required for optimal  $\mathcal{O}(N_{mac}^{-2/d})$  convergence.

In Table 7.6.1 we list various quantities that reveal the quality of the adaptive method, where we use the notation introduced in Section 6.4. After a few iterations, the error estimator and the error itself converge with the same convergence rate, which matches the asymptotic

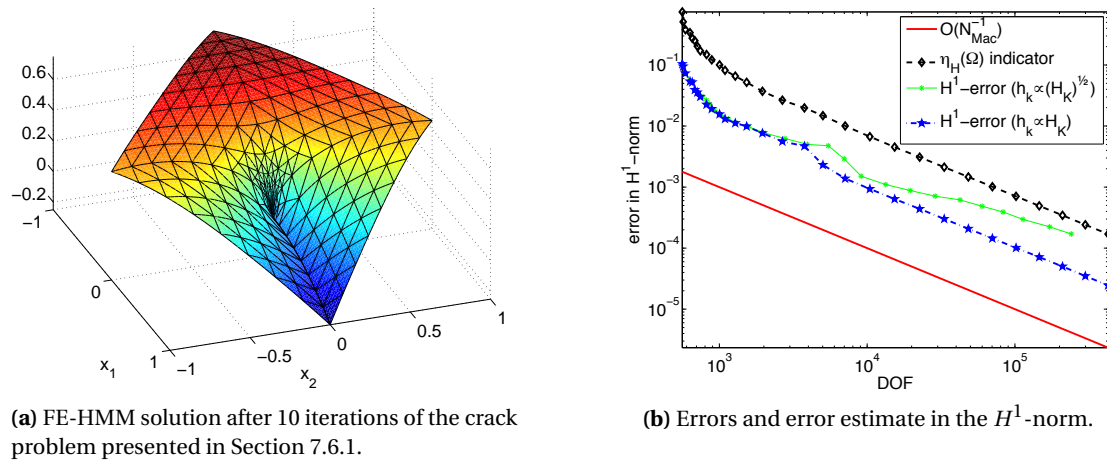


Figure 7.6.1: FE-HMM solution and errors for the crack problem described in Section 7.6.1.

rate we would expect theoretically. Furthermore, the effectivity remains nearly constant which implies that our adaptive method is both efficient and effective.

## 7.6 Numerical experiments

| Iteration | #el    | $\ u^0 - u^{HMM}\ _{H^1(\Omega)}$ | $\eta_H(\Omega)$ | $EOC(e^H)$ | $EOC(\eta_H)$ | $Z_e$ | Eff := $\frac{\eta^H}{e^H}$ |
|-----------|--------|-----------------------------------|------------------|------------|---------------|-------|-----------------------------|
| 1         | 256    | 1.04e-01                          | 7.42e-01         |            |               |       | 0.140                       |
| 2         | 260    | 9.47e-02                          | 5.09e-01         | 11.79      | 48.64         | 0.913 | 0.186                       |
| 3         | 270    | 7.31e-02                          | 3.80e-01         | 13.69      | 15.46         | 0.772 | 0.192                       |
| 4         | 292    | 5.39e-02                          | 3.39e-01         | 7.79       | 2.93          | 0.737 | 0.159                       |
| 5         | 302    | 5.22e-02                          | 2.75e-01         | 1.92       | 12.29         | 0.968 | 0.189                       |
| 6         | 316    | 3.92e-02                          | 2.50e-01         | 12.61      | 4.36          | 0.751 | 0.157                       |
| 7         | 328    | 3.56e-02                          | 2.04e-01         | 5.23       | 10.95         | 0.907 | 0.175                       |
| 8         | 344    | 3.02e-02                          | 1.71e-01         | 6.87       | 7.41          | 0.849 | 0.177                       |
| 9         | 382    | 2.25e-02                          | 1.49e-01         | 5.61       | 2.62          | 0.746 | 0.151                       |
| 10        | 416    | 1.90e-02                          | 1.22e-01         | 3.92       | 4.74          | 0.846 | 0.157                       |
| 11        | 472    | 1.56e-02                          | 1.01e-01         | 3.12       | 2.99          | 0.821 | 0.155                       |
| 12        | 516    | 1.31e-02                          | 8.19e-02         | 3.97       | 4.61          | 0.838 | 0.160                       |
| 13        | 609    | 1.12e-02                          | 6.58e-02         | 1.91       | 2.65          | 0.854 | 0.170                       |
| 14        | 729    | 9.88e-03                          | 5.20e-02         | 1.39       | 2.62          | 0.883 | 0.190                       |
| 15        | 939    | 7.65e-03                          | 3.71e-02         | 2.02       | 2.67          | 0.774 | 0.206                       |
| 16        | 1293   | 5.59e-03                          | 2.66e-02         | 1.96       | 2.08          | 0.731 | 0.211                       |
| 17        | 1834   | 4.69e-03                          | 1.98e-02         | 1.01       | 1.67          | 0.838 | 0.236                       |
| 18        | 2444   | 2.31e-03                          | 1.48e-02         | 4.92       | 2.04          | 0.494 | 0.156                       |
| 19        | 3463   | 1.38e-03                          | 1.00e-02         | 2.96       | 2.23          | 0.597 | 0.138                       |
| 20        | 5122   | 9.32e-04                          | 6.70e-03         | 2.01       | 2.07          | 0.674 | 0.139                       |
| 21        | 7565   | 6.39e-04                          | 4.53e-03         | 1.94       | 2.00          | 0.685 | 0.141                       |
| 22        | 11196  | 4.39e-04                          | 3.10e-03         | 1.91       | 1.94          | 0.688 | 0.142                       |
| 23        | 16440  | 3.01e-04                          | 2.12e-03         | 1.97       | 1.97          | 0.686 | 0.142                       |
| 24        | 24002  | 2.08e-04                          | 1.47e-03         | 1.97       | 1.95          | 0.689 | 0.142                       |
| 25        | 34936  | 1.45e-04                          | 1.01e-03         | 1.93       | 1.96          | 0.697 | 0.143                       |
| 26        | 50537  | 1.01e-04                          | 7.05e-04         | 1.93       | 1.97          | 0.700 | 0.143                       |
| 27        | 72674  | 7.12e-05                          | 4.92e-04         | 1.94       | 1.98          | 0.703 | 0.145                       |
| 28        | 104440 | 5.00e-05                          | 3.45e-04         | 1.95       | 1.96          | 0.702 | 0.145                       |
| 29        | 149224 | 3.49e-05                          | 2.42e-04         | 2.01       | 1.99          | 0.699 | 0.144                       |
| 30        | 212688 | 2.46e-05                          | 1.70e-04         | 1.98       | 1.98          | 0.704 | 0.144                       |

**Table 7.6.1:** Iteration number, number of macro elements,  $H^1$  error, error indicator, experimental order of convergence for the error and the indicator, reduction factor and effectivity for the Crack problem presented in Section 7.6.1.

### 7.7 Summary

In the previous Chapter 6 we derived the *a posteriori* error analysis for the FE-HMM in the energy norm for piecewise linear macro FE. In this Chapter we extended the analysis and the design of the adaptive method to higher order macro (and micro) FE. The extended analysis is based on the construction of the *higher order multiscale flux* and *multiscale jump*; this construction allowed us to largely follow the strategy and analysis of residual-based *a posteriori* error estimators developed for single-scale problems.

These are the first rigorous *a posteriori* results for the FE-HMM with higher order FE derived in the energy norm of the physical variables. The efficiency and reliability of our adaptive strategy is confirmed by numerical results.

The *a posteriori* error analysis for higher order FE-HMM clears the way to introduce *a posteriori* estimates for the FE-HMM in terms of errors in quantities of interest, as they rely on finding the approximation of a so-called dual solution in a space of higher order FE than the space where the original, primal solution was found. We will derive *a posteriori* error estimates in quantities of interest and design a goal-oriented adaptive FE-HMM in Chapter 8.



## 8 Goal-Oriented Adaptive FE-HMM

The residual-based adaptive FE-HMM we have seen so far was designed in order to adapt the macro mesh according to potential singularities (such as reentrant corners). We are now concerned with the problem of adapting the mesh to control the error in a desired quantity of interest. These quantities of interest are obtained with the help of a linear, bounded functional  $J(u)$ . The dual weighted residual (DWR) FE-HMM is based upon the division of the problem into the *primary* (standard) FE-HMM problem and an auxiliary, *dual* FE-HMM problem. The dual problem enables us to “extract” information required to express the error in terms of quantities of interest. We construct an exact error representation using macroscopic *error estimators* that allow us to find an approximation of the error in the quantity of interest. We further design localized macroscopic *refinement indicators*, which are used to drive an adaptive mesh refinement. These refinement indicators and error estimators depend on information from the microscopic quantities. As they are not readily available, we recover the required information on the fly. We will again use the relation between the special FE-HMM components such as the multiscale fluxes and jumps to their equivalent classical, single-scale FEM counterparts, which we derived in Section 7.2 in the context of higher order residual-based FE-HMM. We therefore can largely follow the derivation of single-scale dual weighted residual FE methods.

In Section 8.2 we state the main results of the *a posteriori* analysis for the DWR FE-HMM, and present details of the algorithm in Section 8.3. We provide a proof of our *a posteriori* error estimates in Section 8.4 and estimate the data approximation error in Section 8.5. Finally, numerical experiments in Section 8.6 demonstrate the capabilities of the goal-oriented adaptive FE-HMM.

The results of this Chapter will be published in [13].

## 8.1 Model problem

Let  $\mathcal{T}_H$  denote a conformal mesh and let  $\mathcal{E}_H$  be the set of interfaces. Given  $f \in L^2(\Omega)$  and a domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , we consider the second-order elliptic multiscale equation

$$\begin{aligned} -\nabla \cdot (a^\varepsilon \nabla u^\varepsilon) &= f \quad \text{in } \Omega, \\ u^\varepsilon &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{8.1.1}$$

where  $a^\varepsilon$  is symmetric, satisfies  $a^\varepsilon(x) \in (L^\infty(\Omega))^{d \times d}$  and is uniformly elliptic and bounded. We consider in what follows only homogeneous Dirichlet boundary conditions, but emphasize that more general boundary conditions (such as Neumann or Robin) can be considered by following the lines of the results presented in this Chapter.

## 8.2 Main results

We want to remind the reader once more of the terminology we use. The word *error estimate* denotes a quantity that is used to approximate the unknown error; *upper* and *lower bound* denote quantities that are *always* larger and smaller, respectively, than the actual error. Therefore, error estimates are close to the actual error, but cannot be guaranteed. In contrast, the upper and lower bound might be more inaccurate, but are always guaranteed. We furthermore use the word (*local*) *refinement indicator* (typically obtained from a localization of the error estimator) to denote local quantities on an element, which are used to drive an adaptive mesh refinement.

*Notation 65.* In what follows we will frequently use the primal and dual Galerkin orthogonality property. To emphasize the difference between the FE-HMM solution and a general function in  $V^H = V^p(\Omega, \mathcal{T}_H)$ , we will thus write  $u^{HMM}$  instead of  $u^H$  for the FE-HMM solution of problem (2.4.12) and  $v^H$  and  $\psi^H$  will refer to *any* function in  $V^H = V^p(\Omega, \mathcal{T}_H)$ . Here,  $V^p(\Omega, \mathcal{T}_H)$  is the macro higher-order FE space (7.2.1). We further use  $V = H_0^1(\Omega)$ .

We will only use simplicial elements in what follows, as quadrature formulae for quadrilaterals or hexahedra do not meet condition (7.2.2); they still can be used in the implementation of the method, but the proof involves additional technicalities. We let  $f^H$  be an element-wise  $\mathcal{P}^m(K)$  approximation of  $f$ , we will find out how to choose  $m$  appropriately in Section 8.5.

Let  $J : V \rightarrow \mathbb{R}$  be a linear, bounded functional. We seek the error  $e^H := u^0 - u^{HMM}$  in quantities of interest, i.e.

$$J(e^H) = J(u^0) - J(u^{HMM}).$$

As we mentioned, the DWR method relies on the solution of a primal and a dual problem, which are given as follows.

**Primal problem.** The primal problem for the homogenized solution  $u^0$  is given by (2.1.7), i.e.

$$B_0(u^0, v) = \int_{\Omega} f v \, dx \quad \forall v \in H_0^1(\Omega). \quad (8.2.1)$$

We approximate  $u^0$  by the FE-HMM solution  $u^{HMM} \in V^H$ , where

$$V^H := V^P(\Omega, \mathcal{T}_H), \quad (8.2.2)$$

which is the solution of

$$B(u^{HMM}, v^H) = \int_{\Omega} f v^H \, dx \quad \forall v^H \in V^H(\Omega, \mathcal{T}_H),$$

where  $B(\cdot, \cdot)$  refers to the FE-HMM bilinear form (2.4.6).

**Dual problem.** Then let the *dual solution*  $z^0 \in H_0^1(\Omega)$  be the solution of the *dual problem*

$$B_0(\varphi, z^0) = J(\varphi) \quad \forall \varphi \in H_0^1(\Omega). \quad (8.2.3)$$

The dual solution necessary for the DWR method must be found in a space that is richer than the original FE space, see Remark 20 in Chapter 5.2. We cannot simply solve for a dual solution in  $V^H$ . We choose the richer space

$$V^{\mathcal{H}} := V^{\tilde{P}}(\Omega, \mathcal{T}_H) \quad (8.2.4)$$

to consist of polynomials with  $\tilde{p} > p$ . The discretized dual problem (using the FE-HMM) is then

$$B(\varphi^{\mathcal{H}}, z^{\mathcal{H}EMM}) = J(\varphi^{\mathcal{H}}) \quad \forall \varphi^{\mathcal{H}} \in V^{\mathcal{H}}, \quad (8.2.5)$$

where  $B(\cdot, \cdot)$  is the FE-HMM bilinear form.

*Remark 66.* We want to remind the reader again that when replacing  $z^0$  by  $z^{\mathcal{H}EMM}$  we get reliable error estimates *only* if we can neglect the approximation error  $\|z^0 - z^{\mathcal{H}EMM}\|$ . We refer to Remark 21 for a discussion.

### 8.2.1 Exact error representation

**Definition 67.** We define the *interior* and *jump residuals*  $R_{I,H}$  in  $K$  and  $R_{J,H}$  on  $e$ , respectively as

$$\begin{aligned} R_{I,H}(x)|_K &= f^H + \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \\ R_{J,H}(s)|_e &= -\frac{1}{2} \llbracket \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \rrbracket_e(s), \end{aligned}$$

where the multiscale flux  $\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x)$  is defined in (7.2.7) and the higher order multiscale jump  $\llbracket \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \rrbracket_e(s)$  is defined in (7.2.9) in Section 7.2.2.

We now state the DWR FE-HMM error representation, which we will derive in Section 8.4.

**Theorem 68** (Exact DWR FE-HMM error representation). *The error in quantities of interest is given by the exact representation*

$$J(u^0 - u^{\text{HMM}}) = \sum_{K \in \mathcal{T}_H} \eta_H(K) + \xi_H(K),$$

where the local error estimator is given by

$$\eta_H(K) := \int_K R_{I,H}(x) \left( z^{\mathcal{HMM}} - \psi^H \right) dx + \int_{\partial K} R_{J,H}(x) \left( z^{\mathcal{HMM}} - \psi^H \right) ds, \quad (8.2.6)$$

and the data approximation error  $\xi_H(K)$  is defined as in Definition 69.

*Proof.* We give the proof in Section 8.4. □

**Definition 69** (Data approximation error). We define the data<sup>1</sup> approximation error  $\xi_H(K)$  on an element  $K$  as

$$\xi_H(K) = \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^{\text{HMM}} \right) \cdot \nabla \left( z^{\mathcal{HMM}} - \psi^H \right) dx \quad (8.2.7a)$$

$$+ B_{0,K} \left( u^0 - u^{\text{HMM}}, z^0 - z^{\mathcal{HMM}} \right) \quad (8.2.7b)$$

$$+ B_{0,K} \left( u^0 - u^{\text{HMM}}, \psi^H \right) \quad (8.2.7c)$$

$$- \int_K (f^H - f) \left( z^{\mathcal{HMM}} - \psi^H \right) dx, \quad (8.2.7d)$$

where the subscript  $K$  in  $B_{0,K}(\cdot, \cdot)$  indicates the restriction of  $B_0(\cdot, \cdot)$  onto the element  $K$ . The function  $\psi^H \in V^H$  is an *arbitrary* function; there are various strategies for the choice of  $\psi^H$ , see also Section 5.3.1. We will investigate the choice of  $\psi^H$  in Section 8.5.1.

*Remark 70.* The components of the data approximation error (8.2.7) can be understood as follows.

(8.2.7a) Approximation error caused by using an inaccurate tensor (implicitly computed within the micro problem) and numerical quadrature instead of using the exact, continuous  $a^0(x)$  in the estimator, i.e. the macro, micro and modeling error.

(8.2.7b) Approximation error for using  $z^{\mathcal{HMM}}$  instead of  $z^0$ .

---

<sup>1</sup>In contrast to the residual-based approach,  $\xi_H(K)$  here not only involves errors due to approximation of *data* but also an additional error introduced by using  $z^{\mathcal{HMM}}$  instead of  $z^0$ .

(8.2.7c) Approximation error caused by the introduction of  $\psi^H$  and the lack of Galerkin orthogonality due to the discrepancy between  $u^{HMM}$  and  $u_{ex}^{0,H}$ . Here,  $u_{ex}^{0,H}$  is the exact FEM solution of problem (8.2.1).

(8.2.7d) Approximation error caused by approximating the right-hand side  $f(x)$  in the estimator.

**Theorem 71** (Estimation of the data approximation error). *Under the assumption 78 and 79 and given a quadrature formula  $\{x_{K_\ell}, \omega_{K_\ell}\}_{\ell=1}^{\mathcal{L}}$  that satisfies condition (7.2.2), we can estimate the data approximation error  $\xi_H(\Omega)$  as follows.*

$$\begin{aligned}
 |\xi_H(\Omega)| \leq & C \left( \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left\| \nabla \left( z^{HMM} - \psi^H \right) \right\|_{L^2(\Omega)} \\
 & + C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left( H^{\bar{p}} + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod,dual} \right) \\
 & + C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left\| \nabla \psi^H \right\|_{L^2(\Omega)} \\
 & + C H^{m+1} \left\| \nabla \left( z^{HMM} - \psi^H \right) \right\|_{L^2(\Omega)}
 \end{aligned} \tag{8.2.8}$$

*Proof.* The proof follows in Section 8.5. □

*Remark 72.* In Section 8.5.1 we will evaluate the estimates of Theorem 71 for two choices of  $\psi^H$ , namely  $\psi^H \equiv 0$  and  $\psi^H \equiv z^{HMM}$ , where  $z^{HMM}$  is the dual solution found in  $V^H$  instead of  $V^{\mathcal{H}}$  of problem (8.2.5). We will usually choose  $m = p - 1$ .

Error estimators involving the Cauchy-Schwarz inequality do not allow for cancelations of errors among different elements of the domain and therefore typically overestimate the error significantly; many explicit error estimators further involve unspecific constants. In contrast, the DWR method allows for cancelation of errors among elements and does not depend on unspecific constants. Cancelation over elements, however, would not be possible if we had a positive local estimator on every element and then summed over all elements, as it was done in the residual-based case. Instead, we need to distinguish between the (global) error representation formula (Theorem 68) and local refinement indicators  $\bar{\eta}_H(K)$ . The (global) error representation formula consists of local (signed) error estimators  $\eta_H(K)$  and is used for estimating the global error of the approximate solution  $u^{HMM}$ , where we want to have cancelation of errors. The *local refinement indicators*  $\bar{\eta}_H(K)$  presented in Definition 73 are used to drive the adaptive mesh refinement procedure. We thus take the (positive) refinement indicator to drive the adaptive scheme, i.e. to refine those elements whose indicator is “largest” in some sense to be defined; see the different marking schemes presented in Section 5.4.

**Definition 73** (*A posteriori* DWR refinement indicators). We define the following *local refinement indicators*

$$\bar{\eta}_H(K) := \left| \int_K R_{I,H}(x) \left( z^{HMM} - \psi^H \right) dx + \int_{\partial K} R_{J,H}(s) \left( z^{HMM} - \psi^H \right) ds \right|.$$

**Corollary 74** (A posteriori DWR upper bound). *We have an a posteriori upper bound given by*

$$|J(u^0 - u^{HMM})| \leq \sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K) + |\xi_H(K)|. \quad (8.2.9)$$

*Proof.* Follows immediately from Definition 73.  $\square$

The term (8.2.9) can be further estimated as in the following Proposition (see [28]).

**Proposition 75** (A posteriori error estimate). *We have an alternative a posteriori upper bound given by*

$$|J(u^0 - u^{HMM})| \leq \sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K) + |\xi_H(K)| \quad (8.2.10)$$

$$\leq \sum_{K \in \mathcal{T}_H} \rho_K \omega_K + \bar{\xi}_H(K), \quad (8.2.11)$$

where the cell-residuals  $\rho_K$  and the weights  $\omega_K$  are given by

$$\rho_K := \left( \|R_{I,H}(x)\|_{L^2(K)}^2 + \frac{1}{H_K} \|R_{J,H}(s)\|_{L^2(\partial K)}^2 \right)^{1/2}$$

$$\omega_K := \left( \|z^{\mathcal{HMM}} - \psi^H\|_{L^2(K)}^2 + H_K \|z^{\mathcal{HMM}} - \psi^H\|_{L^2(\partial K)}^2 \right)^{1/2},$$

and where  $\bar{\xi}_H(K)$  is given by (8.2.7b), but where we sum over the norms (using Cauchy-Schwarz inequality for the integrals and bilinear forms, together with boundedness) of every single component.

By separating the error into cell residuals and weights, we not only eliminate cancellation over elements but also eliminate cancellation of data approximation errors between the element and cell residuals.

We remark as mentioned in Section 5.3 that since going from (8.2.10) to (8.2.11) does not add any additional information on either the approximate dual solution  $z^{\mathcal{HMM}}$  or on the approximation of the tensor  $a^0(x)$ , the expression  $\sum_{K \in \mathcal{T}_H} \rho_K \omega_K + \bar{\xi}_H(K)$  will be bigger than  $\sum_{K \in \mathcal{T}_H} \bar{\eta}_H(K) + |\xi_H(K)|$  without increased reliability (see [87]). We will therefore use (8.2.10) in what follows.

### 8.3 Algorithm

Our adaptive algorithm for the DWR FE-HMM again follows the framework of the adaptive algorithm for standard FEM and consists of loops of the form

Solve  $\rightarrow$  Estimate  $\rightarrow$  Mark  $\rightarrow$  Refine.

However, due to the multiscale nature of the problem and due to the primal-dual aspect of the goal-oriented method, we need to modify the algorithm accordingly.

**Algorithm 76** (Adaptive DWR-FE-HMM).

**Solve.** For the macro and micro meshes obtained by REFINE, compute the micro solutions (only for the refined macro elements) and the macro solutions  $u^{\text{HMM}}$  and  $z^{\text{HMM}}$  for the primal and the dual problem, respectively. Compute  $\psi^H$  if appropriate. Compute and store the high-order multiscale fluxes  $\Pi_{a^\varepsilon \nabla \varphi_{K_\ell}^h}(x)$  (based on the FE basis functions  $\varphi^H$ ) for the refined elements during the macro assembly process.

**Estimate.** Reconstruct the full multiscale fluxes  $\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x)$  and jumps  $\llbracket \overline{\Pi_{a^\varepsilon \nabla u_{K_\ell}^h}} \rrbracket_e(s)$ . Estimate the error in the quantity of interest  $J(u^0 - u^{\text{HMM}})$  by computing the estimate  $\sum_{K \in \mathcal{T}_H} \eta_H(K)$ . (This can serve as a stopping criterion using  $|\sum_{K \in \mathcal{T}_H} \eta_H(K)| < \text{tol}$ , for a certain prescribed tolerance  $\text{tol}$ .)

**Mark.** Mark the elements on a subset  $\widetilde{\mathcal{T}}_H$  of  $\mathcal{T}_H$  based on the refinement indicators  $\bar{\eta}_H(K)$  following the strategy described below.

**Refine.** Refine the marked elements (and some neighbors for mesh conformity) and update the mesh for the primal and the dual problem. Update the micro meshes of the sampling domains corresponding to the refined macro elements (see below for a discussion about an optimal coupling).

**Mesh adaption strategy** Depending on the problem to be solved, a suitable mesh adaption strategy has to be chosen. See also Section 5.4. For the marking strategy we will follow the *maximum marking strategy* (used in [96]) and refine the elements which contribute the most to  $J(u^0 - u^{\text{HMM}})$ . An element  $K$  of the mesh is refined if

$$\frac{\eta_H(K)}{\max_{\tilde{K} \in \mathcal{T}_H} \eta_H(\tilde{K})} \geq \vartheta,$$

where  $0 < \vartheta < 1$  is a user-defined parameter.

As in the case of the adaptive, residual-based FE-HMM we will re-use the solution of the micro problems obtained in previous iterations, see Remark 38. Since the polynomial degree  $p$  we use for the primal FE problem is different from the polynomial degree  $\tilde{p}$  we use for the dual FE problem, typically also different quadrature rules will be used. For example in the case  $p = 1$ ,  $\tilde{p} = 2$ , we have one quadrature node, and therefore one micro problem per macro element in the primal problem, but three in the dual problem, all at different locations. A naive approach would be to compute all four micro problems. As the computation of the micro problems is computationally very expensive, we should only compute the three micro problems for the assembly of the dual stiffness matrix and use these result for the assembly of the primal stiffness matrix. An elegant way to do this is described in the following remark.

*Remark 77.* The following choice of FEM shape functions will reduce the computational cost of the primal and dual problem. We can use the (somewhat non-standard) choice of piecewise quadratic shape functions as in [30], where out of the six shape functions per element, three

are piecewise linear (and coincide with the standard piecewise linear shape functions) and three are quadratic. Then, when assembling the dual stiffness matrix, the computation of every entry of the primal stiffness matrix with piecewise linear functions is already done and comes for free. This way, we need to compute only three instead of four micro problems per macro element (for  $p = 1$ ,  $\tilde{p} = 2$ ). Furthermore, we then have also in the primal problem a quadrature formula which is correct up to piecewise quadratic tensors  $a^0(x)$  instead of piecewise linear ones.

However, the downside of this choice of basis functions is that we no longer have a nodal basis.

### 8.3.1 Macro-micro coupling

In the residual-based adaptive FE-HMM case, we found that – under appropriate assumptions – the ideal coupling for the data approximation error of the estimator is given by  $h_K \propto (H_K)^{\frac{p}{2q}}$ . Here  $p$  and  $q$  refer to the polynomial degree of the (primal) macro and micro FE space, respectively. The polynomial degree of the dual macro FE space is denoted as  $\tilde{p}$ . The coupling rate suggested for the estimator coincides with the coupling rate for the solver suggested by the *a priori* results for the error in the  $H^1$ -norm. In our residual-based FE-HMM implementation, we therefore use the same coupling for both the solver (to obtain  $u^{HMM}$ ) as well as for the estimator (to obtain  $\eta_H(K)$ ). In the general DWR FE-HMM case, the optimal coupling for the solution  $u^{HMM}$  does not necessarily coincide with the optimal coupling for the estimator  $\eta_H(K)$ . We only require the functional  $J(\cdot)$  to be linear and bounded. It is not straightforward to obtain *a priori* estimates for errors in *general* quantities of interest that could be used to extract an (optimal) macro-micro coupling rate. We know *a priori* coupling rates for the error  $J(e^H)$  for a few quantities of interest, such as (presented for the case of piecewise linear elements)

- for the global  $H^1$ -norm we expect a convergence rate of  $J(e^H) \propto H$ ,
- for the global  $L^2$ -norm we expect a convergence rate of  $J(e^H) \propto H^2$ ,
- for the point-wise error we have a very rough estimate of  $J(e^H) \propto H^3$ , see [28, Section 3.3].

For each case, the data approximation error of the estimator has to be studied in order to find an optimal coupling rate for the estimator.

If we simply follow the *a priori* estimates for the  $H^1$ -norm to obtain the coupling for the primal and dual solver, we obtain the following macro-micro coupling scheme: refine the micro mesh in each sampling domain  $K_{\delta_\ell}$  according to

$$h_K \propto (H_K)^{\frac{p}{2q}} \text{ for the primal problem}$$

and

$$h_{\mathcal{X}} \propto (H_K)^{\frac{\tilde{p}}{2q}} \text{ for the dual problem,}$$



## 8.4 Proof of Theorem 68 (Exact DWR FE-HMM error representation)

---

where  $H_K$  is the size of the macro element  $K \in \mathcal{T}_H$  and where the subscript  $K$  and  $\mathcal{K}$  indicate the primal and dual coupling, respectively. We emphasize, that for reasons of computational efficiency, all of the micro problems should be solved with one specific coupling and the results should be used for the primal and dual solver and for the estimator.

For the *error estimator*, we will derive upper bounds in Section 8.5, which will lead to equations from which we can derive a coupling scheme. We emphasize though that to obtain these results, the cancelations of errors amongst the elements is ignored, which may lead to a significant overestimation of the data approximation error.

In the numerical experiment in Section 8.6.1 we will further analyze different macro-micro couplings.

### 8.3.2 Approximation of the linear functional

In our analysis we ignore the approximation error originating from the (numerical) approximation of the functional  $J(\cdot)$ . In general, we have to approximate the functional in some way, which introduces a discretization error. For example, as mentioned in Remark 22, we might use *mollification* in the case of point-wise directional derivative errors and approximate  $J(u) = \partial_n u(x^*)$  at a point  $x^*$  by an approximation

$$J_\epsilon(u) := \int_{\Omega} \partial_n u(x) k_\epsilon(x - x^*) dx,$$

where  $u$  stands for the exact or FE-HMM solution of the primal problem. In an adaptive algorithm,  $\epsilon$  should also be chosen adaptively. Difficulties might not only arise from point-wise quantities of interest. For example, consider a quantity of interest such as the average over a certain area  $J(u) := \int_S u dx$ ,  $S \subset \Omega$  as  $J(u) = \int_{\Omega} \Theta(x) u dx$  with a step function

$$\Theta(x) := \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{otherwise,} \end{cases}$$

i.e. the dual problem in its strong form

$$-\nabla \cdot (a^0(x) \nabla z^0(x)) = \Theta(x)$$

has a non-smooth right hand side, which can impair the regularity of the dual solution  $z^0$ . When using regularization or mollification, the regularity can be improved, but this comes at the cost of introducing an additional *modeling error* for the functional. We will not take this kind of modeling error into account for our analysis.

## 8.4 Proof of Theorem 68 (Exact DWR FE-HMM error representation)

We prove Theorem 68, the exact DWR FE-HMM error representation. On the way, we will justify Definition 69 of the data approximation error.

We proceed in three steps. First, we use duality arguments to state the error in quantities of interest, but still depending on the unknown dual function  $z^0$ . Second, we introduce a function  $\psi^H$ . Third, we replace all numerically unavailable quantities, such as  $z^0$  or  $f$ , by their computable, discrete approximations such as  $z^{HEMM}$  and  $f^H$ , respectively in order to obtain a computable error estimator. All unknown discrepancy terms (e.g.  $f^H - f$ ) are put into the data approximation error.

**Step I: error in quantities of interest through duality arguments.** Let  $J(u^{HMM})$  denote the quantity of interest, where  $J: V \rightarrow \mathbb{R}$  is a linear and bounded error functional in the dual space  $V'$ . We want to find the exact error representation for the error  $e^H := u^0 - u^{HMM}$  in quantities of interest  $J(\cdot)$

$$J(e^H) = J(u^0) - J(u^{HMM}).$$

The dual problem (8.2.3)

$$B_0(\varphi, z^0) = J(\varphi)$$

holds for all  $\varphi \in V$ . We therefore choose  $\varphi = e^H$  to obtain

$$J(e^H) = B_0(e^H, z^0). \quad (8.4.1)$$

**Step II: Introduction of  $\psi^H$  and dealing with the lack of Galerkin orthogonality.** In classical FEM, we have Galerkin orthogonality of the homogenized bilinear form for the primal problem

$$B_0(u^0 - u_{ex}^{0,H}, \psi^H) = 0 \quad \forall \psi^H \in V^H, \quad (8.4.2)$$

and the dual problem

$$B_0(\varphi^H, z^0 - z_{ex}^{0,H}) = 0 \quad \forall \varphi^H \in V^H,$$

where  $u_{ex}^{0,H} \in V^H$  is the FEM solution of the homogenized bilinear form  $B_0(\cdot, \cdot)$  (without quadrature, as indicated by the subscript  $_{ex}$ ) (8.2.3) and  $z_{ex}^{0,H} \in V^H$  is defined likewise.

While the classical FEM without quadrature is free of variational crimes, and therefore Galerkin orthogonality holds, the FE-HMM bilinear form  $B(\cdot, \cdot)$  is based by construction on quadrature. Therefore, we commit variational crimes in the FE-HMM and in general, Galerkin orthogonality  $B_0(e^H, \psi^H) \neq 0$  does not hold for the FE-HMM solution  $u^{HMM}$ .

We now introduce an arbitrary  $\psi^H \in V^H$  in equation (8.4.1), which we will later choose with the aim of increasing the accuracy or reduce the computational cost of the error estimator, see Section 8.5.1. We will add the discrepancy terms to the data approximation error term  $\xi_H(K)$ . We introduce  $\psi^H$  and obtain

$$\begin{aligned}
 & J(u^0 - u^{HMM}) \\
 &= B_0(u^0 - u^{HMM}, z^0) \\
 &= B_0(u^0 - u^{HMM}, z^0 - \psi^H) + B_0(u^0 - u^{HMM}, \psi^H). \tag{8.4.3}
 \end{aligned}$$

**Step III: Making the estimate computable.** In (8.4.1), the dual solution  $z^0$  is the exact solution of (8.2.3). Furthermore, the bilinear form  $B_0(\cdot, \cdot)$  involves the exact homogenized tensor, and the primal solution  $u^0$  of (8.2.1) depends on the exact right hand side  $f$  of (8.1.1). These are all quantities, which in general cannot be evaluated exactly in the *a posteriori* error estimate. We therefore replace them in what follows step by step by computable quantities, e.g. we replace  $f$  by an approximation  $f^H$ , and add the introduced disparity terms, such as  $f - f^H$ , to the data approximation error  $\xi_H(K)$ .

We start by replacing the exact dual solution  $z^0$  by the computable FE-HMM dual solution  $z^{\mathcal{HMM}}$  in (8.4.3). We obtain

$$\begin{aligned}
 & J(u^0 - u^{HMM}) \\
 &= B_0(u^0 - u^{HMM}, z^0 - \psi^H) + B_0(u^0 - u^{HMM}, \psi^H). \\
 &= \underbrace{B_0(u^0 - u^{HMM}, z^{\mathcal{HMM}} - \psi^H)}_{(*)} + B_0(u^0 - u^{HMM}, z^0 - z^{\mathcal{HMM}}) + B_0(u^0 - u^{HMM}, \psi^H). \tag{8.4.4}
 \end{aligned}$$

The second term in (8.4.4) corresponds to using the approximation  $z^{\mathcal{HMM}}$  instead of the exact  $z^0$  and the third term arises from the introduction of  $\psi^H$  and the lack of Galerkin orthogonality between  $e^H$  and  $\psi^H$ . We now only consider expression (\*) of eq. (8.4.4). We split it into two parts, which are

$$\begin{aligned}
 (*) &= B_0(u^0 - u^{HMM}, z^{\mathcal{HMM}} - \psi^H) \\
 &= \int_{\Omega} f(z^{\mathcal{HMM}} - \psi^H) dx - B_0(u^{HMM}, z^{\mathcal{HMM}} - \psi^H), \tag{8.4.5}
 \end{aligned}$$

where we used (8.2.1). We replace the exact  $f$  in the first term of (8.4.5) by its approximation  $f^H$  and get

$$\int_{\Omega} f(z^{\mathcal{HMM}} - \psi^H) dx = \int_{\Omega} f^H(z^{\mathcal{HMM}} - \psi^H) dx - \int_{\Omega} (f^H - f)(z^{\mathcal{HMM}} - \psi^H) dx. \tag{8.4.6}$$

We then substitute the flux  $a^0 \nabla u^{HMM}(x)$  involving the exact homogenized tensor  $a^0(x)$  by the

computable multiscale flux (7.2.7) in the second term of (8.4.5). This leads to

$$\begin{aligned}
 & B_0 \left( u^{HMM}, z^{\mathcal{H}MM} - \psi^H \right) \\
 &= \sum_{K \in \mathcal{T}_H} \int_K \left( a^0(x) \nabla u^{HMM} \right) \cdot \nabla \left( z^{\mathcal{H}MM} - \psi^H \right) dx \\
 &= \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \cdot \nabla \left( z^{\mathcal{H}MM} - \psi^H \right) dx \\
 &\quad - \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^{HMM} \right) \cdot \nabla \left( z^{\mathcal{H}MM} - \psi^H \right) dx.
 \end{aligned} \tag{8.4.7}$$

We divide the estimate into element terms and jump terms. We therefore integrate (8.4.7) by parts to get

$$\begin{aligned}
 & B_0 \left( u^{HMM}, z^{\mathcal{H}MM} - \psi^H \right) \\
 &= - \sum_{K \in \mathcal{T}_H} \int_K \nabla \cdot \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right) \left( z^{\mathcal{H}MM} - \psi^H \right) dx \\
 &\quad + \frac{1}{2} \sum_{K \in \mathcal{T}_H} \sum_{e \subset \partial K} \int_e \overline{\left[ \Pi_{a^\varepsilon \nabla u_{K_\ell}^h} \right]}_e(s) \left( z^{\mathcal{H}MM} - \psi^H \right) ds \\
 &\quad - \sum_{K \in \mathcal{T}_H} \int_K \left( \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^{HMM} \right) \cdot \nabla \left( z^{\mathcal{H}MM} - \psi^H \right) dx,
 \end{aligned} \tag{8.4.8}$$

where we used the definition of the multiscale jump (7.2.9). We plug (8.4.5), (8.4.6) and (8.4.8) into (8.4.4) and obtain the error representation formula

$$J(u^0 - u^{HMM}) = \sum_{K \in \mathcal{T}_H} \eta_H(K) + \xi_H(K),$$

where  $\eta_H(K)$  and  $\xi_H(K)$  are defined in (8.2.6) and (8.2.7), respectively, and the proof is complete.

## 8.5 Proof of Theorem 71 (Estimation of the data approximation error)

Up to this point we did not make any specific spatial assumptions on the oscillating tensor, such as periodicity in the micro scale, to derive our error estimate. Furthermore, we kept both sampling domain size and boundary conditions of the micro problem for the FE-HMM rather general. We now find estimates of the data approximation error (8.2.7) by using the results we obtained in Section 7.5, where we analyzed the data approximation error of the residual-based FE-HMM.

### Assumptions.

Similarly to our analysis in the higher-order residual-based FE-HMM, make the following assumptions. See Section 7.5 and [16, Section 3] for details.

## 8.5 Proof of Theorem 71 (Estimation of the data approximation error)

**Assumption 78.** We assume that the cell functions  $\psi_{K_{\delta_\ell}}^i \in W(K_{\delta_\ell})$  satisfy

$$\left| \psi_{K_{\delta_\ell}}^i \right|_{H^{q+1}(K_{\delta_\ell})} \leq C \varepsilon^{-q} \sqrt{|K_{\delta_\ell}|}, \quad (8.5.1)$$

where  $C$  is independent of  $\varepsilon$ , the quadrature node  $x_{K_\ell}$  and the domain  $K_{\delta_\ell}$  for all  $i = 1, \dots, d$ .

For simplicity, we furthermore make the following assumption.

**Assumption 79.** We assume that  $a^\varepsilon = a(x, x/\varepsilon) = a(x, y)$  is  $Y$ -periodic in  $y$  and

$$a(\cdot, y)|_K \text{ is constant} \quad (8.5.2)$$

for  $K \in \mathcal{T}_H$ .

Assumption 79 implies that the homogenized tensor  $a^0$  is constant in any  $K \in \mathcal{T}_H$ .

### Decomposition and estimation of the data approximation error.

The data approximation error (8.2.7) consists of four components and is given by

$$\xi_H(K) = \int_K \left[ \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^{HMM} \right] \cdot \nabla (z^{\mathcal{HMM}} - \psi^H) dx \quad (8.5.3a)$$

$$+ B_{0,K} \left( u^0 - u^{HMM}, z^0 - z^{\mathcal{HMM}} \right) \quad (8.5.3b)$$

$$+ B_{0,K} \left( u^0 - u^{HMM}, \psi^H \right) \quad (8.5.3c)$$

$$- \int_K (f^H - f) (z^{\mathcal{HMM}} - \psi^H) dx. \quad (8.5.3d)$$

We will estimate some components using the *global a priori* estimates for the FE-HMM that were obtained in [2] and that we presented in (2.4.15). The *a priori* estimates are given for the primal problem as

$$\| u^0 - u^{HMM} \|_{H^1(\Omega)} \leq C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} \right) + err_{mod}, \quad (8.5.4)$$

and similarly we have for the dual problem

$$\| z^0 - z^{\mathcal{HMM}} \|_{H^1(\Omega)} \leq C \left( H^{\tilde{p}} + \left( \frac{h}{\varepsilon} \right)^{2q} \right) + err_{mod,dual}, \quad (8.5.5)$$

where we assume the same micro mesh size for both primal and dual problem.

We further use the results we obtained in context of the data approximation error for the higher order FE-HMM in Section 7.5, where we found that

$$\left\| a^0(x) \nabla u^{HMM}(x) - \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) \right\|_{L^2(K)} \leq C \left( \left( \frac{h}{\varepsilon} \right)^{2q} \right) + err_{mod}.$$

We estimate each of the four different components as follows. For the approximation error caused by using inaccurate tensor instead of  $a^0(x)$  (due to micro and modeling errors) in the estimator, which is given by (8.5.3a), we obtain

$$\begin{aligned}
 & \left| \int_K \left[ \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^{HMM} \right] \cdot \nabla \left( z^{\mathcal{HMM}} - \psi^H \right) dx \right| \\
 & \leq \left\| \Pi_{a^\varepsilon \nabla u_{K_\ell}^h}(x) - a^0(x) \nabla u^{HMM} \right\|_{L^2(K)} \left\| \nabla \left( z^{\mathcal{HMM}} - \psi^H \right) \right\|_{L^2(K)} \\
 & \leq C \left( \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left\| \nabla \left( z^{\mathcal{HMM}} - \psi^H \right) \right\|_{L^2(K)}. \tag{8.5.6}
 \end{aligned}$$

The dual solution approximation error (8.5.3b) for using  $z^{\mathcal{HMM}}$  instead of  $z^0$  reads

$$\begin{aligned}
 & \left| B_{0,K} \left( u^0 - u^{HMM}, z^0 - z^{\mathcal{HMM}} \right) \right| \\
 & \leq C \left\| \nabla \left( u^0 - u^{HMM} \right) \right\|_{L^2(K)} \left\| \nabla \left( z^0 - z^{\mathcal{HMM}} \right) \right\|_{L^2(K)}
 \end{aligned}$$

where we used boundedness of  $B_0(\cdot, \cdot)$ . We can further find the *global* estimate

$$\begin{aligned}
 & \left| B_0 \left( u^0 - u^{HMM}, z^0 - z^{\mathcal{HMM}} \right) \right| \\
 & \leq C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left( H^{\bar{p}} + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod,dual} \right), \tag{8.5.7}
 \end{aligned}$$

where we used the *a priori* estimates (8.5.4) and (8.5.5) for the primal and dual solution, respectively. The approximation error caused by the introduction of  $\psi^H$  and the lack of Galerkin orthogonality (8.5.3c) can be estimated as

$$\begin{aligned}
 & \left| B_{0,K} \left( u^0 - u^{HMM}, \psi^H \right) \right| \\
 & \leq C \left\| \nabla \left( u^0 - u^{HMM} \right) \right\|_{L^2(K)} \left\| \nabla \psi^H \right\|_{L^2(K)},
 \end{aligned}$$

and we again can find the global estimate

$$\begin{aligned}
 & \left| B_0 \left( u^0 - u^{HMM}, \psi^H \right) \right| \\
 & \leq C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left\| \nabla \psi^H \right\|_{L^2(\Omega)}. \tag{8.5.8}
 \end{aligned}$$

And finally the approximation error (8.5.3d) for approximating the (sufficiently regular) right hand side  $f(x)$  with its piecewise  $\mathcal{P}^m(K)$  interpolation  $f^H$  is

$$\begin{aligned}
 & \left| \int_K \left( f^H - f \right) \left( z^{\mathcal{HMM}} - \psi^H \right) dx \right| \\
 & \leq C \left\| f - f^H \right\|_{L^2(K)} \left\| \nabla \left( z^{\mathcal{HMM}} - \psi^H \right) \right\|_{L^2(K)} \\
 & \leq C H^{m+1} \left\| \nabla \left( z^{\mathcal{HMM}} - \psi^H \right) \right\|_{L^2(K)}. \tag{8.5.9}
 \end{aligned}$$

Summing up equations (8.5.6), (8.5.7), (8.5.8), and (8.5.9) completes the proof.

### 8.5.1 Choice of $\psi^H$

As explained in Section 5.3.1, we can choose the arbitrary  $\psi^H \in V^H$  in many different ways. While Bangerth and Rannacher [28] aim at reducing the computational effort to obtain the estimator, we aim at reducing its data approximation error. We consider two different choices,  $\psi^H \equiv 0$ , which will act as a benchmark we compare other choices to, and  $\psi^H \equiv z^{HMM}$ , the dual solution  $z^{HMM} \in V^H$ . The motivation behind choosing  $\psi^H \equiv z^{HMM}$  is that it is expected to reduce the norm  $\left\| \nabla \left( z^{\mathcal{HMM}} - z^{HMM} \right) \right\|_{L^2(\Omega)}$  since  $z^{\mathcal{HMM}} \approx z^{HMM}$  and that we can find an upper bound for the norm, similar to our choice  $\psi^H \equiv I^H e^H$  with the Clément interpolant  $I^H$  in context of the residual-based estimates.

$\psi^H \equiv 0$ . With this choice, we obtain

$$\begin{aligned} |\xi_H(\Omega)| \leq & C \left( \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left\| \nabla z^{\mathcal{HMM}} \right\|_{L^2(\Omega)} \\ & + C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left( H^{\tilde{p}} + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod,dual} \right) \\ & + C H^{m+1} \left\| \nabla z^{\mathcal{HMM}} \right\|_{L^2(\Omega)} \end{aligned} \quad (8.5.10)$$

$\psi^H \equiv z^{HMM}$ . We choose  $\psi^H$  to be solution  $z^{HMM} \in V^H$  of the dual problem. Then we can further estimate

$$\begin{aligned} \left\| \nabla \left( z^{\mathcal{HMM}} - z^{HMM} \right) \right\|_{L^2(\Omega)} & \leq \left\| \nabla \left( z^0 - z^{\mathcal{HMM}} \right) \right\|_{L^2(\Omega)} + \left\| \nabla \left( z^0 - z^{HMM} \right) \right\|_{L^2(\Omega)} \\ & \leq C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} \right) + err_{mod} + err_{mod,dual}, \end{aligned}$$

where we used estimates *a priori* estimates (8.5.4) and (8.5.5) and that  $H^p > H^{\tilde{p}}$  as  $p < \tilde{p}$ . We get

$$\begin{aligned} |\xi_H(\Omega)| \leq & C \left( \left( \frac{h}{\varepsilon} \right)^2 + err_{mod} \right) \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} + err_{mod,dual} \right) \\ & + C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left( H^{\tilde{p}} + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod,dual} \right) \\ & + C \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} \right) \left\| \nabla z^{HMM} \right\|_{L^2(\Omega)} \\ & + C H^{m+1} \left( H^p + \left( \frac{h}{\varepsilon} \right)^{2q} + err_{mod} + err_{mod,dual} \right). \end{aligned} \quad (8.5.11)$$

Looking at (8.5.10) and (8.5.11) we see that the convergence rate may be considerably limited by  $\left\| \nabla z^{\mathcal{HMM}} \right\|_{L^2(\Omega)}$  and  $\left\| \nabla z^{HMM} \right\|_{L^2(\Omega)}$ , respectively. We therefore expect the data approximation error for  $\psi^H \equiv 0$  to be of similar size as for  $\psi^H \equiv z^{HMM}$ . In general, with a macro or modeling error, it is not straightforward to say which choice of  $\psi^H$  will be more beneficial.

Further investigation is required to explain the influence of  $\psi^H$  on the data approximation error. We emphasize that the error estimator of the DWR method exploits cancelation of errors over different elements and we therefore expect the above estimates to significantly overestimate the real data approximation error.

## 8.6 Numerical Experiments

The numerical experiments were carried out using an implementation in Matlab. Our implementation is based on the FE-HMM code presented in Chapter 3 and in [11]. The code for the mesh bisection is based in part on *i*FEM (see [41]).

### 8.6.1 Point-wise directional derivative

In this example (which is chosen similarly to the one-dimensional problems presented in [96]) we consider a square domain, where we are interested in the directional derivative error at a certain point  $x^*$

$$J(u) := \nabla u(x^*) \cdot n, \quad (8.6.1)$$

with the normed vector  $n$ . We consider the problem

$$\begin{aligned} -\nabla \cdot (a^\varepsilon(x) \nabla u^\varepsilon) &= f \quad \text{in } \Omega \\ u^\varepsilon &= g_D \quad \text{on } \Gamma_D = \partial\Omega \end{aligned}$$

on a domain  $\bar{\Omega} = [0, 1]^2$ . We choose  $\varepsilon = 10^{-5}$  and set the tensor to

$$a\left(x, \frac{x}{\varepsilon}\right) := \tilde{a}^\varepsilon\left(\frac{x}{\varepsilon}\right) a^0(x),$$

where the exact homogenized tensor  $a^0(x)$  is given as

$$a^0(x) := \left(9e^{-1000(x_1-0.5)^2-1000(x_2-0.5)^2} + 1\right) I_2. \quad (8.6.2)$$

Here

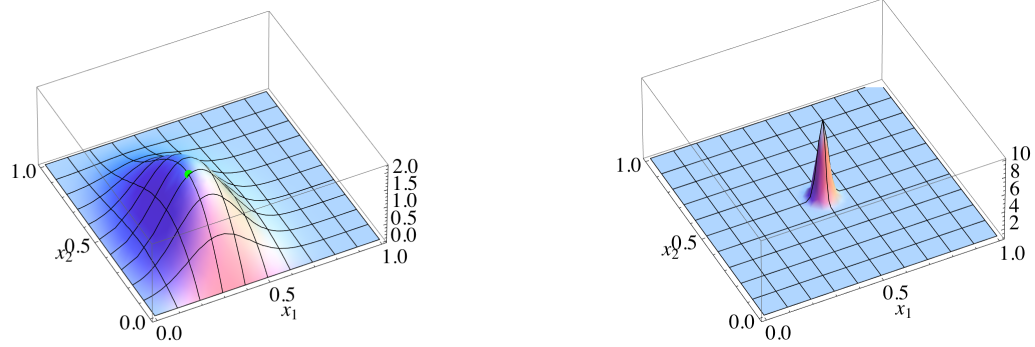
$$\tilde{a}^\varepsilon(x) := \frac{16}{15} \left( \sin\left(2\pi \frac{x_1}{\varepsilon}\right) + \frac{5}{4} \right) \left( \cos\left(2\pi \frac{x_2}{\varepsilon}\right) + \frac{5}{4} \right),$$

whose coefficients are chosen in such a way that they give the unit tensor when homogenized. We define the exact solution as

$$u^0(x) := 100(1-x_1)^2 x_1 (1-x_2)^2 x_2 e^{-20(x_1-\frac{1}{3})^2 - (x_2-\frac{1}{4})^2},$$

which determines the Dirichlet boundary conditions as  $g_D = u^0$  on  $\partial\Omega$ , and choose the right hand side  $f$  accordingly. An illustration of the exact solution and the homogenized tensor are given in Figure 8.6.1. We use piecewise linear macro FE for the solution  $u^{HMM}$  of the primal problem and piecewise quadratic FE for the solution  $z^{HEMM}$  of the macro dual problem. The





(a) Exact homogenized solution  $u^0(x)$ . We are interested in the directional derivative error at the point  $x^* = (0.3, 0.3)$ , which is marked in green.

(b) Exact homogenized tensor  $a^0(x)$ .

**Figure 8.6.1:** Exact solution and tensor of the goal-oriented problem of Section 8.6.1.

micro problems are solved with piecewise linear FE.

As a measure of the quality of our error estimator, we define the effectivity indices

$$\text{Eff} := \left| \frac{\eta_H(\Omega)}{J(u^0 - u^{HMM})} \right|, \quad (8.6.3)$$

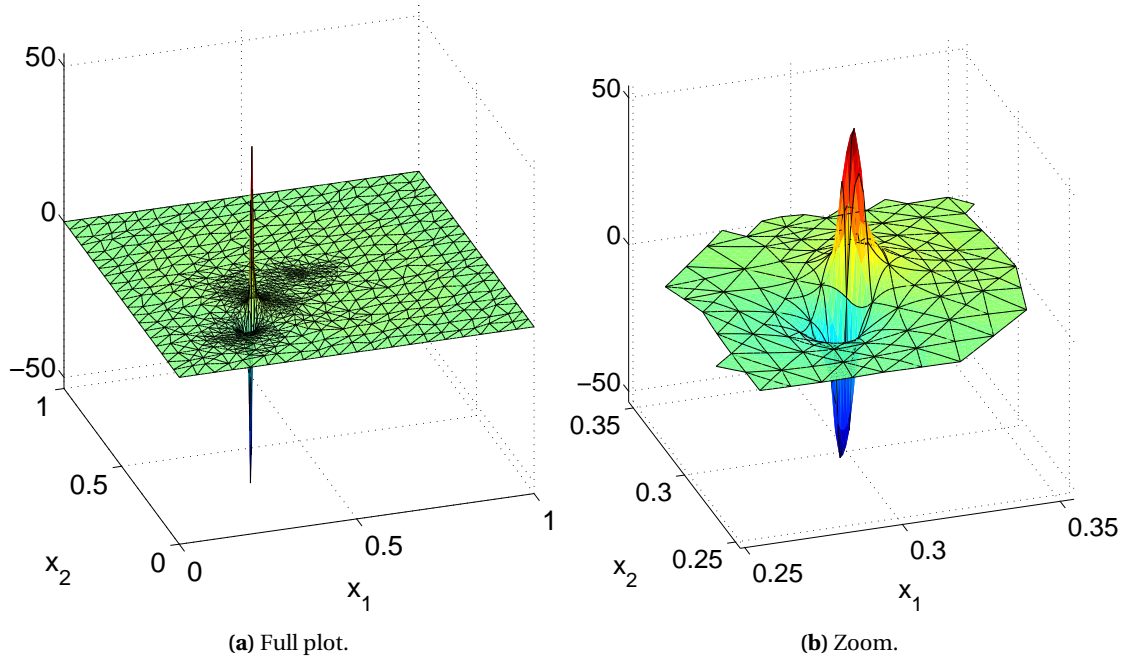
which ideally are equal to one. Furthermore we inspect the data approximation error

$$|\zeta_H(\Omega)| = \left| J(u^0 - u^{HMM}) - \sum_{K \in \mathcal{T}_H} \eta_H(K) \right|. \quad (8.6.4)$$

Let  $\hat{h} := (N_{mic})^{-1/d}$  be the scaled (i.e. independent of  $\varepsilon$ ) micro mesh size, where  $N_{mic}$  denotes the degrees of freedom of the micro problems in one micro sampling domain corresponding to a certain macro element. We choose our quantity of interest to be the point-wise directional derivative error at the point  $x^* = (0.3, 0.3)$  (slightly off the peak of the “bump”) in direction  $n = (1/\sqrt{2}, 1/\sqrt{2})$ . The exact solution is given by  $J = 3.25819$ . As the point-wise derivatives of FE solutions are in general not defined on edges of elements (see also Remark 22), we approximate the functional (8.6.1) at point  $x^* \in \Omega$  with a regularized output functional  $J(u) = \frac{1}{|S^\varepsilon|} \int_{S^\varepsilon} \nabla u \cdot n \, dx$ , where the small domain  $S^\varepsilon$  is an  $\varepsilon$ -ball centered around the point  $x^*$ , see [28]. To evaluate the integral in the numerical implementation we choose a node of the macro mesh to coincide with the point  $x^*$  and take the average of the gradient on the supporting elements of the corresponding hat function evaluated at point  $x^*$ <sup>2</sup>.

For the macro mesh we start from a mesh with 441 DOF. We do not choose a uniform grid but one where the points are slightly distorted compared to uniform mesh, see Figure 8.6.10. The distortion will resemble to more realistic meshes originating from practical applications and reduce aliasing-effects. We use an initial mesh size for the micro mesh of  $\hat{h} = \frac{1}{8}$  and

<sup>2</sup>See the code corresponding to exercise 3.3 provided with [28].



**Figure 8.6.2:** Numerical dual solution  $z^{HMM}$  for the goal-oriented problem after 4 iterations. We are interested in the point-wise directional derivative error at  $x^* = (0.3, 0.3)$ .

$\delta = \varepsilon = 10^{-5}$  with periodic boundary conditions for the micro problems. We use the *Maximum* marking scheme [104, 96] described in Section 8.3 and choose a parameter of  $\vartheta = 0.25$ .

Since we are interested in multiscale problems and therefore in finding how the numerical homogenization, and thus the approximate homogenized conductivity tensor, influences the convergence rate and quality of the indicator, we choose  $f^H$  to be a piecewise  $\mathcal{P}^4$  approximation. The approximation error is then  $\|f - f^H\|_{L^2} \approx \mathcal{O}(H^5)$ , which ensures that the influence of the approximation  $f^H$  is negligible in comparison to the whole data approximation error. We remark that, although the numerical evaluation of  $f$  is computationally much cheaper than the solution of the micro problems, the polynomial degree of the approximation  $f^H$  should be chosen such that the convergence rate of the approximation error of  $f$  coincides with the convergence rate of the approximation error of the tensor. In Figure 8.6.2 we present the numerical dual solution  $z^{HMM}$  after 4 iterations, which is very localized around  $x^*$  and has very small support.

For a first test, we use the macro-micro coupling rate  $\hat{h} \propto H_K$  and set  $\psi_H \equiv 0$ . In Figure 8.6.3 we compare the exact quantity of interest  $J(u^0)$  with the numerically obtained  $J(u^{HMM})$ . As we have given a (signed) error estimator  $\eta_H(\Omega) \approx J(u^0) - J(u^{HMM})$  we can compute an estimated confidence interval where we predict – based only upon our numerical solution – the exact quantity of interest to be. To take the data approximation error  $|\xi_H(\Omega)|$  into account, we *assume* that it is of similar size as the estimated error  $|\eta_H(\Omega)|$  and thus we define the confidence interval to be the area between  $J(u^{HMM}) - \eta_H(\Omega)$  and  $J(u^{HMM}) + \eta_H(\Omega)$ . The

confidence interval is therefore rather an indication than a reliable bound, see also Remark 23. Estimating the data approximation error  $\xi_H(K)$  itself is rather difficult, computationally expensive and problem dependent. The data approximation error might be, in some cases, significantly larger than the indicator. Our asymptotic analysis presented in 8.5 gives an asymptotic upper bound of the data approximation error; we remark again that due to the cancelation of errors, this upper bound might significantly overestimate the true error.

The estimated confidence interval is marked with the shaded blue area in Figure 8.6.3. We see that starting from the iteration four, the exact quantity of interest lies within our predicted confidence interval. We further notice that  $J(u^{HMM})$  converges to  $J(u^0)$  and the confidence interval width converges to zero. However, in the first three iterations, the exact quantity of interest is outside the predicted confidence interval and therefore the interval is – as mentioned in Remark 23 – not unconditionally reliable.

We emphasize again that a confidence interval as given in the DWR method is difficult to construct for the residual-based error estimate, as we can quantify the residual-based estimate only up to an unspecific constant (even when neglecting the data approximation error). This unspecific constant is in general very difficult to estimate.

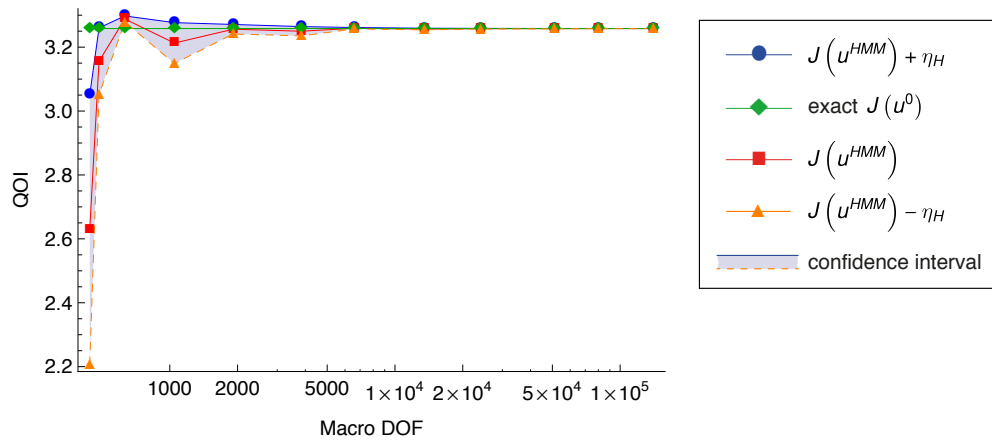
In Figure 8.6.4 we compare the point-wise derivative value at  $x^*$  of our goal-oriented refinement scheme with a residual-based adaptive refinement. While the adaptive strategy for the goal-oriented refinement optimizes the mesh for a minimal point-wise derivative error at  $x^*$ , the residual-based refinement optimizes the mesh for a minimal error in the energy norm. Therefore, we see that the DWR-FE-HMM has an error in the quantity of interest, which is significantly smaller than the error for the residual-based method. We now turn to a more rigorous study of the DWR FE-HMM.

### Main observations.

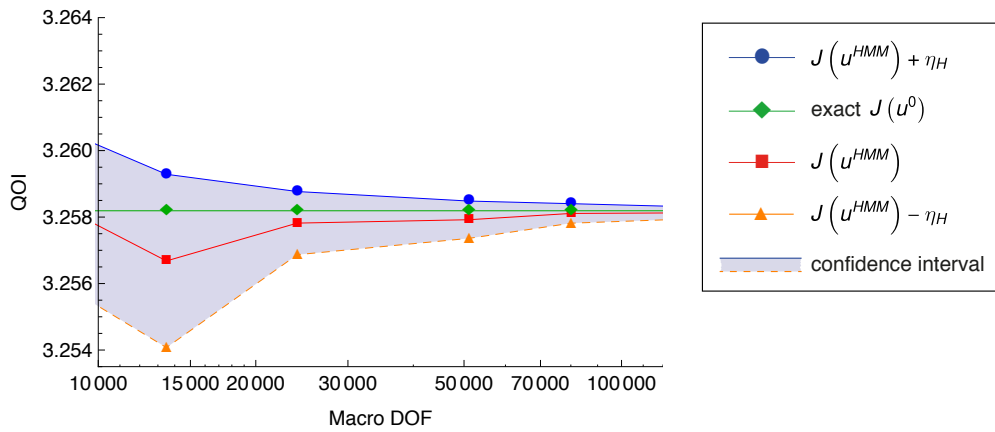
In what follows we will find four main observations from our numerical experiment:

1. The convergence rate of the error in the quantity of interest can be significantly better for the goal-oriented DWR FE-HMM than for the residual-based or uniform FE-HMM.
2. We can get an effectivity index close to one, depending on the choice of parameters ( $\psi^H$ , the optimal macro-micro coupling rate for the primal and dual problem); *how* the parameters influence the effectivity index is not fully understood.
3. A good refinement indicator might not be a good local estimator. A bad local estimator might not be a bad refinement indicator.
4. Cancelation of errors is essential for an accurate global estimator.

We now study different aspects of the numerical experiment, which will lead to the main observations above. As there are various sources of errors, which influence the convergence rate and the effectivity index, we proceed in two steps.

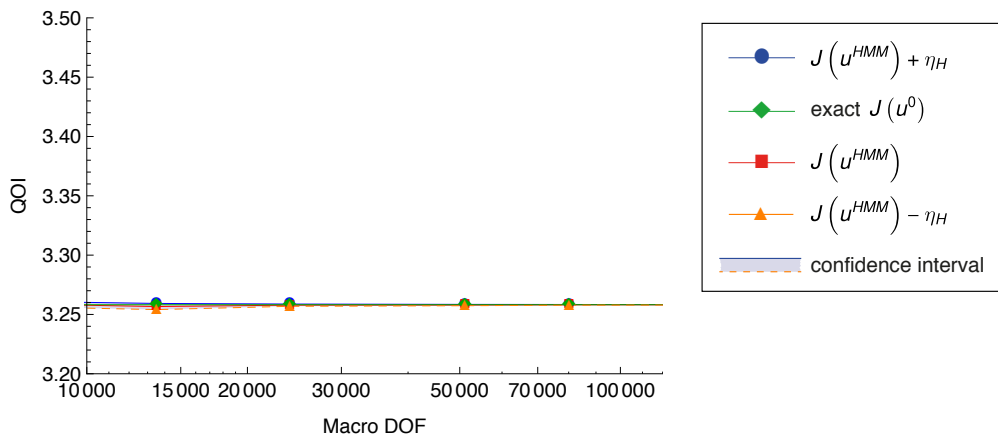


(a) Full plot

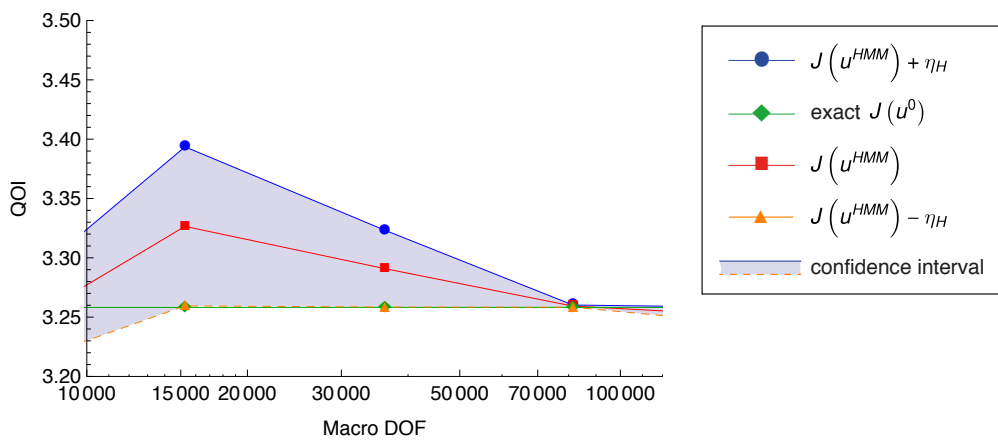


(b) Zoom

**Figure 8.6.3:** Plot of the point-wise directional derivative at a point  $x^*$  vs. the macro degrees of freedom. The green line marks the exact value  $J(u^0) = \partial_n u^0(x)|_{x=x^*}$ , which the engineer or scientist is interested in. The red line indicates the numerically obtained quantity of interest  $J(u^{HMM})$ . The DWR method allows to specify – by introducing upper and lower error approximation estimates  $J(u^{HMM}) + \eta_H(\Omega)$  and  $J(u^{HMM}) - \eta_H(\Omega)$ , respectively – a confidence interval (shaded blue), where we expect the exact solution  $J(u^0)$  (green line) to be in.



(a) Adaptive, goal-oriented refinement.



(b) Adaptive global residual-based refinement.

**Figure 8.6.4:** Plot of the point-wise directional derivative at a point  $x^*$  vs the macro degrees of freedom. The green line marks the exact value  $J(u^0) = \partial_n u^0(x)|_{x=x^*}$ , which the engineer or scientist is interested in. The red line indicates the numerically obtained quantity of interest  $J(u^{HMM})$ . The DWR method allows to specify – by introducing upper and lower error approximation estimates  $J(u^{HMM}) + \eta_H(\Omega)$  and  $J(u^{HMM}) - \eta_H(\Omega)$ , respectively – a confidence interval (shaded blue), where we expect the exact solution  $J(u^0)$  (green line) to be in. Comparison of the DWR-FE-HMM (a) with a global residual-based refinement scheme (b).

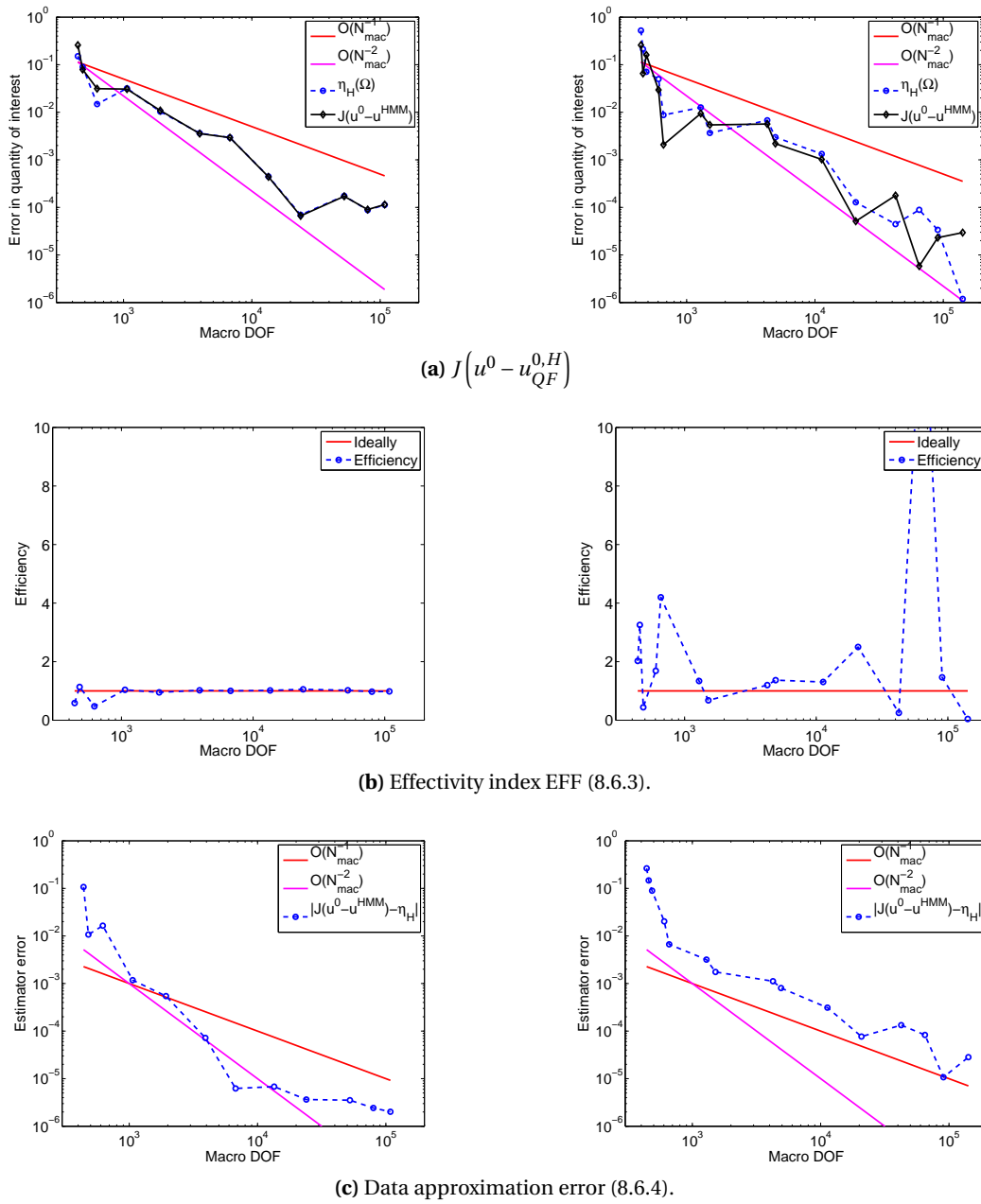
1. We will first consider the case where we assume that we have an exact homogenized tensor given; we will use the explicitly available tensor (8.6.2) (of course, in general, this tensor is unknown). Thus, both the micro and the modeling errors vanish and we simply have a standard DWR FEM with quadrature.
2. We will use the FE-HMM to approximate the conductivity tensor. Thus, we have an additional error from the micro-problems and (in general) a modeling error.

**The DWR FEM case.** We start with the standard DWR FE method and the primal FEM (with quadrature) solution  $u_{QF}^{0,H}$  of (8.2.1), where we evaluate the exact homogenized tensor  $a^0(x_\ell)$  numerically at the quadrature nodes  $x_\ell$ , thus the subscript  $QF$ . In this case, both the micro and the modeling error vanish. We first set  $\psi^H \equiv 0$ . In Figure 8.6.5 we see that the error in the quantity of interest converges with a rate between  $\mathcal{O}(N^{-1})$  and  $\mathcal{O}(N^{-2})$  (for the  $H^1$ -norm we would obtain  $\mathcal{O}(N^{-1/2})$ ). The effectivity index is very close to one. The data approximation error converges with a rate between  $\mathcal{O}(N^{-1})$  and  $\mathcal{O}(N^{-2})$ , it is comparable to the error  $J(e^H)$  itself.

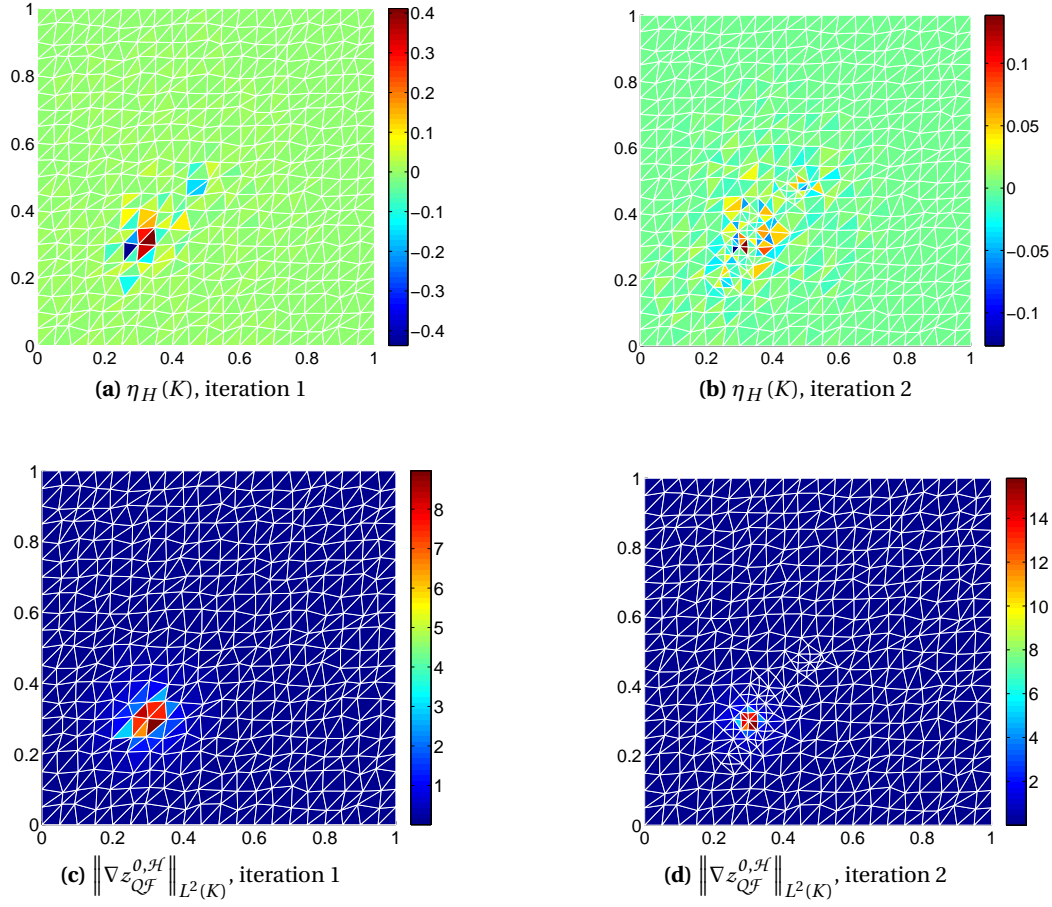
We then choose  $\psi^H \equiv z_{QF}^{0,H}$ ; where  $z_{QF}^{0,H} \in V^H$  is the piecewise linear FEM solution of the dual problem (8.2.3), where we use numerical quadrature but the exact  $a^0(x_\ell)$ . The error in the quantity of interest again converges with a rate between  $\mathcal{O}(N^{-1})$  and  $\mathcal{O}(N^{-2})$ . But the effectivity index varies between 0 and 14 and the convergence rate data approximation error deteriorates to  $\mathcal{O}(N^{-1})$ .

The error estimates for the data approximation obtained in Section 8.5 do not take the cancelation of the error among elements into account. Hence, we inspect how the cancelation of errors influences the error estimator and therefore the data approximation error. As the data approximation error (8.2.8) may involve the terms  $\left\| \nabla z_{QF}^{0,\mathcal{H}} \right\|_{L^2(K)}$  (or  $\left\| \nabla (z_{QF}^{0,\mathcal{H}} - \psi^H) \right\|_{L^2(K)}$ ), where  $z_{QF}^{0,\mathcal{H}}$  is the piecewise quadratic dual FEM solution of (8.2.3), we also want to have a closer look at how these terms.

**Inter-element cancelations and gradient of the dual solution.** We set again  $\psi^H \equiv 0$ . In Figure 8.6.6 (top) we look at the element-wise error estimator  $\eta_H(K)$  (notice the different range of the color bars). We see that the error estimators have similar positive and negative values, therefore we expect the error cancelation among the elements to be highly effective. Moreover, we compare in Figure 8.6.7(a) the error estimator  $\sum_{K \in \mathcal{T}_H} \eta_H(K)$  with cancelation to the error estimator  $\sum_{K \in \mathcal{T}_H} |\eta_H(K)|$  without cancelation. We see that  $\sum_{K \in \mathcal{T}_H} |\eta_H(K)| \approx \mathcal{O}(1)$ . We emphasize that while the error estimator without cancelation stays nearly constant with increasing DOE, the error estimator with cancelation converges and features good accuracy. Furthermore, although the error estimator without cancelation is inappropriate for quantifying the error in the quantity of interest, the local component  $\bar{\eta}(K) = |\eta_H(K)|$  is very suitable for driving the adaptive mesh refinement, i.e. for indicating which elements have to be refined.



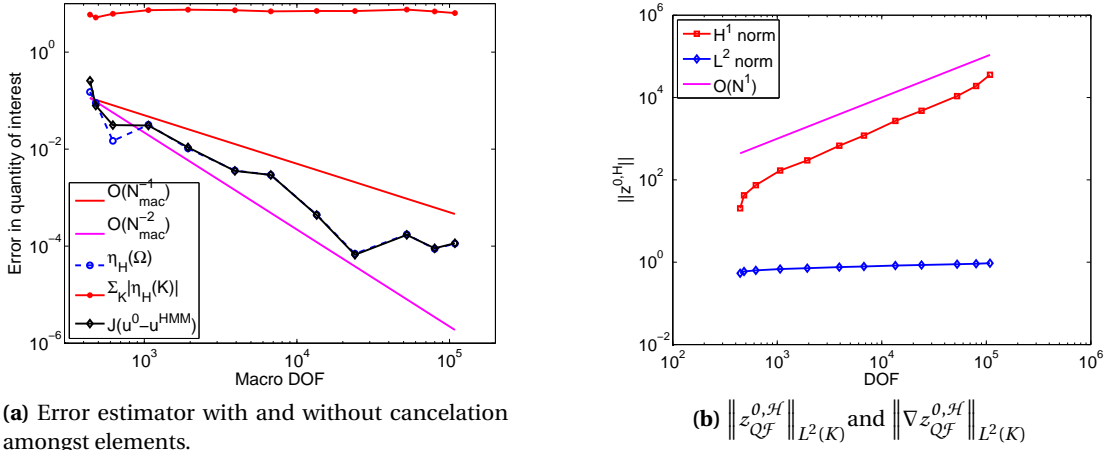
**Figure 8.6.5:** Error and error indicator, effectivity index and approximation error (from top to bottom) for the DWR FEM solution with an *exact homogenized tensor*  $a^0(x_\ell)$  for the goal-oriented problem of Section 8.6.1. Comparison of the choices  $\psi^H \equiv 0$  (left) and  $\psi^H \equiv z_{QF}^{0,H}$  (right). The jagged convergence plots are a consequence of the cancellation of errors, see Remark 80.



**Figure 8.6.6:**  $\eta_H(K)$  and  $\|\nabla z_{Q\mathcal{F}}^{0,\mathcal{H}}\|_{L^2(K)}$  for the point-wise directional derivative single-scale DWR FEM.

We now look at the gradient of  $z_{Q\mathcal{F}}^{0,\mathcal{H}}$ , where we use an exact  $a^0(x_\ell)$  with numerical quadrature. We remark that in our analysis of the data approximation error in Section 8.5 is based on Assumption 79. As the analysis hence did not cover the general case  $a^\varepsilon(x) = a(x, \frac{x}{\varepsilon})$ , we cannot directly compare the experiment with the theoretical results. In Figure 8.6.6 (bottom) we show  $\|\nabla z_{Q\mathcal{F}}^{0,\mathcal{H}}\|_{L^2(K)}$  for iteration 1 and 2. The norm is highly localized around the point of interest  $x^*$  and it increases from iteration 1 to 2. This is due to the numerical implementation of the functional  $J(\cdot)$ , as described above. With increasing DOF, the support for the shape-function, which is used to evaluate the average point-wise directional derivative, shrinks. Hence, its numerically evaluated derivative and also  $\|\nabla z_{Q\mathcal{F}}^{0,\mathcal{H}}\|_{L^2(K)}$  increases. In Figure 8.6.7(b) we show that we have an increasing *global* norm  $\|z_{Q\mathcal{F}}^{0,\mathcal{H}}\|_{H^1(\Omega)} \approx \mathcal{O}(N)$ . A regularization or mollification could be used in order to obtain a more regular dual solution, as mentioned in Remark 22 and Section 8.3.2. This would come with a modeling error for the functional  $J(\cdot)$ . We remark that this behavior is not particular to point-wise errors, but will also be apparent e.g. for the





**Figure 8.6.7:** Estimated error and norms  $\|z_{Q^F}^{0,\mathcal{H}}\|_{L^2(\Omega)}$  and  $\|\nabla z_{Q^F}^{0,\mathcal{H}}\|_{L^2(\Omega)}$  for the point-wise directional derivative single-scale DWR FEM.

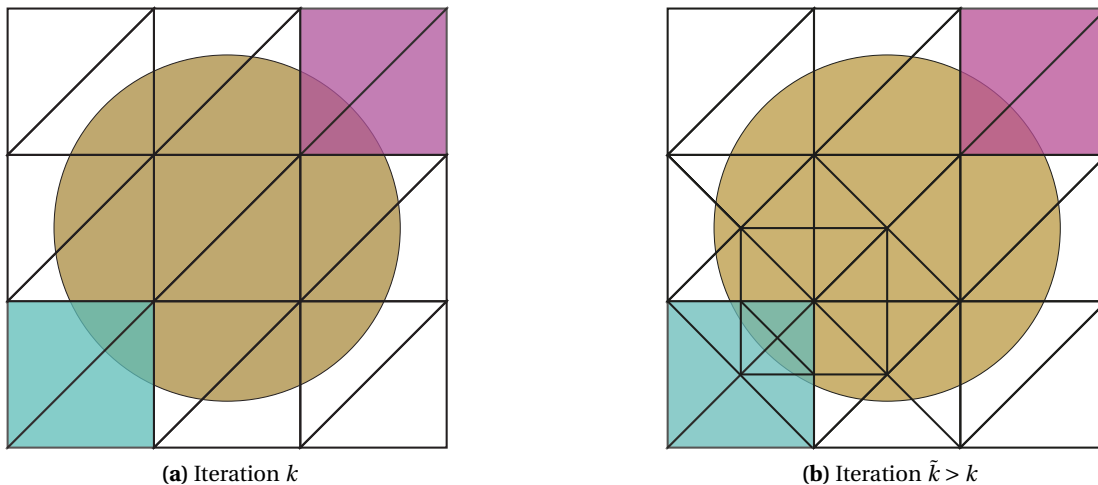
average over a domain  $S \subset \Omega$ , see Section 8.3.2.

*Remark 80* (Error cancelations). In explicit error estimates, such as the residual-based error estimates, the *a posteriori* error estimates provide estimates on the norm of the error. In contrast, the dual-weighted residual error estimates (and all implicit-type error estimates) provide rather an error function. In the DWR case, errors (and data approximation errors) on different elements can cancel each other out, leading to very accurate estimates. The error convergence curve, however, may not be a smooth line, but may be rather jagged. This is caused, as the mesh might not always take the symmetry into account that allowed for error cancelations on different elements. See Figure 8.6.8 for an illustration. We emphasize that this is an effect that shows both in the single-scale DWR FEM as well as in the DWR FE-HMM.

**The DWR FE-HMM case.** Up until now we looked at the *single scale* DWR FE method. We now consider the *multiscale* DWR FE-HMM. When moving from the DWR FEM to the DWR FE-HMM, additional sources of errors, such as the micro and modeling errors are introduced; we analyze how they affect the point-wise derivative error, the effectivity index and data approximation error.

We start with comparing in Figure 8.6.9 the error and effectivity index when using a uniform, a global residual-based (as in Chapter 7) and the DWR goal-oriented approach. We set  $\psi^H \equiv 0$  and choose a coupling of  $\hat{h} \propto H_k$  for both the solution of the primal and dual problem and for the evaluation of the *a posteriori* estimator.

In Figure 8.6.9(a) we see that the point-wise derivative error  $J(e^H)$  converges with an order of  $\mathcal{O}(N_{mac}^{-1/2})$  for the uniform refinement and for the adaptive, residual-based refinement scheme. For the DWR FE-HMM we obtain a (mean) convergence rate of approximately  $\mathcal{O}(N_{mac}^{-3/2})$ , which varies between  $\mathcal{O}(N_{mac}^{-1})$  and  $\mathcal{O}(N_{mac}^{-2})$ . We notice that the convergence rate is much larger than the  $\mathcal{O}(N_{mac}^{-1/2})$  we would have in the global energy-norm. In Figure 8.6.9(b)



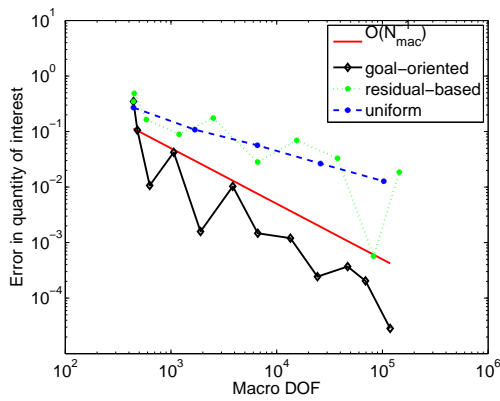
**Figure 8.6.8:** Schematic of error cancellation in the DWR method. The brown disk symbolizes a conductivity tensor or quantity of interest which is such that – due to symmetry – the error as well as the data approximation error among the elements marked in green and purple cancel each other (left figure). After a few iterations, the mesh might not take the symmetry into account anymore (right figure), and the error and data approximation error in the elements marked in green (with a significantly reduced error) do not cancel with the error in the unrefined element marked in purple anymore.

we see that the effectivity index for the DWR FE-HMM varies between 0.5 and 8, whereas the efficiencies for the residual-based and uniformly refined FE-HMM are very close to 1. We will further examine the effectivity index in the DWR FE-HMM case below. In Figure 8.6.10 we see the corresponding mesh after 4 iterations for both the goal-oriented and the residual-based adaptive FE-HMM. While the residual-based scheme mostly refines around the peak of the conductivity tensor, the goal-oriented *additionally* refines around the point of interest  $x^*$ .

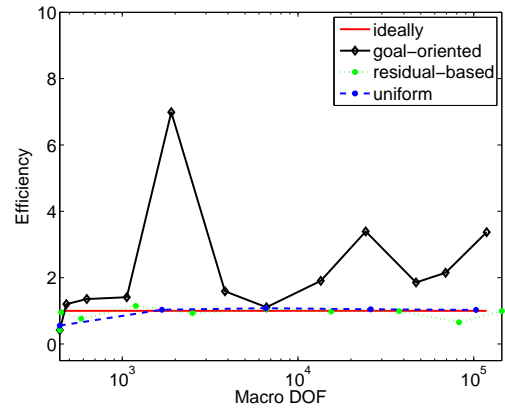
We now have a closer look at the error  $J(e^H)$ , the effectivity index and the data approximation error and their dependence on several parameter choices.

**Macro-micro coupling.** As it is not clear what convergence rate we should expect *a priori* for this quantity of interest (see Section 8.3), we test the coupling  $\hat{h} \propto H_k$  and  $\hat{h} \propto H_k^{3/2}$  for all primal and dual problem and the estimator. In Figure 8.6.11 we show the error and error indicator, the effectivity index and the data approximation error for the different coupling rates and choices of  $\psi^H$ .

The error  $J(e^H)$  lies for all parameter choices between  $\mathcal{O}(N_{mac}^{-1})$  and  $\mathcal{O}(N_{mac}^{-2})$ , with small variations. The data approximation error follows the asymptotic rate of  $\mathcal{O}(N_{mac}^{-1})$  and therefore does not converge as fast as  $J(e^H)$ . The quality of the effectivity index depends strongly on the choice of parameters. While the effectivity index improves as expected from Figure (a) to (b) due to a smaller micro error and is very close to 1 in Figure (b), further investigation is required to understand the deterioration of the effectivity index from Figure (c) to (d). Finally,

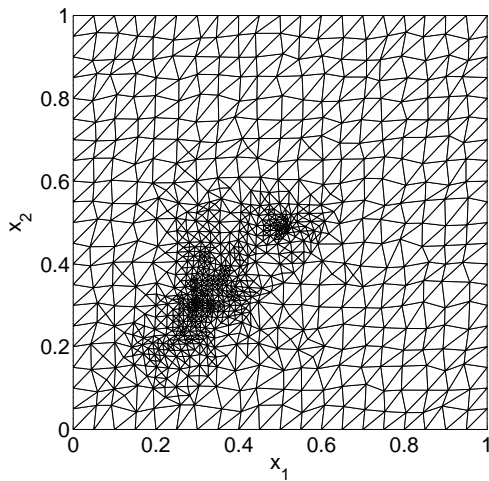


(a) Error for uniform, global residual-based and goal-oriented refinement.

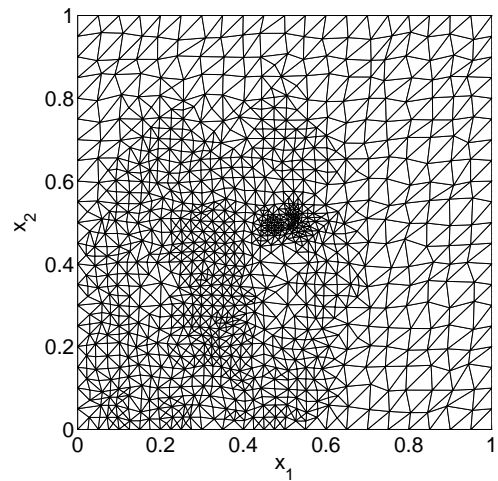


(b) Effectivity index for the uniform, global residual-based and goal-oriented refinement.

**Figure 8.6.9:** Errors and effectivity index for the goal-oriented problem described in Section 8.6.1.



(a) Goal-oriented refinement (1041 DOF,  $J(u^0 - u^{HMM}) = 0.034$ ).



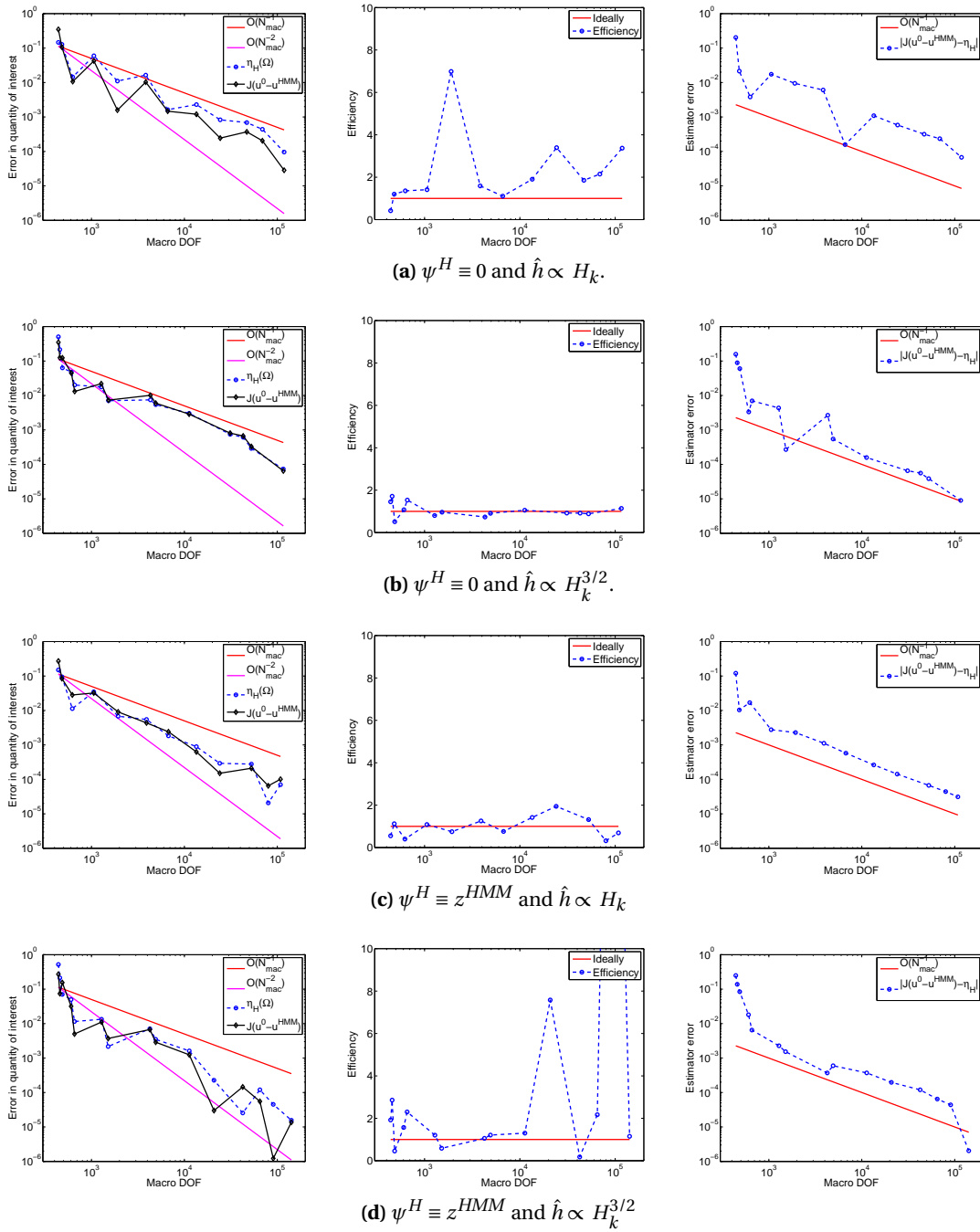
(b) Global residual-based refinement (1209 DOF,  $J(u^0 - u^{HMM}) = 0.113$ ).

**Figure 8.6.10:** Mesh after 4 iterations for global residual-based and goal-oriented refinement for the point-wise directional derivative problem.

we see that the indicator  $|\eta_H(\Omega)|$  is of comparable magnitude as the data approximation error  $|\xi_H(\Omega)|$  (although the asymptotic convergence rates are slightly different), which justifies our choice of the confidence interval chosen in Figures 8.6.3 and 8.6.4.

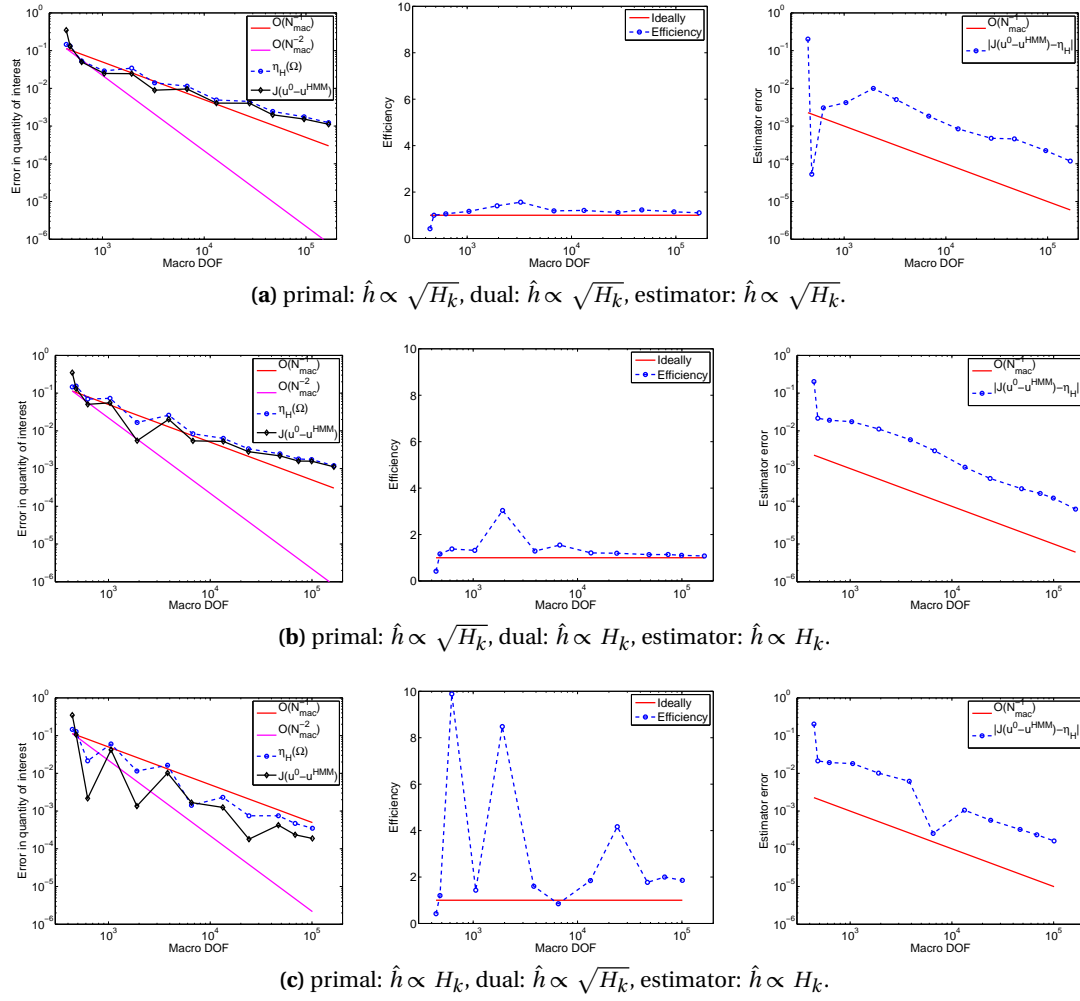
In Figure 8.6.12 we finally choose different coupling schemes for the primal and dual problem and the estimator. We see from Figure (a) and (b) that the coupling  $\hat{h} \propto \sqrt{H_k}$  for the solution of the primal problem deteriorates the convergence rate of  $J(e^H)$  to approximately  $\mathcal{O}(N_{mac}^{-1})$ , even if we select the coupling  $\hat{h} \propto H_k$  for the dual problem (and estimator). In Figure (c) we see that the convergence rate of the error  $J(e^H)$  is indeed less affected by choosing  $\hat{h} \propto \sqrt{H_k}$  for the dual problem than for the primal problem. Further investigation is required to fully understand this behavior. We remark that while here we chose different coupling schemes for the different components, this is usually not done in practice. For reasons of computational effectivity index, the micro problems are computed once (with a specific coupling scheme) and the multiscale fluxes are stored for later use. Therefore one specific coupling scheme is used for both the solution of the primal and dual problem and the estimator.

From Figures 8.6.11 and 8.6.12 it is evident that while the effectivity index is highly sensitive to the choice of parameters such as the function  $\psi^H$  or the coupling scheme, the refinement indicators  $\bar{\eta}(K)$  are still suitable for driving the adaptive mesh refinement.



**Figure 8.6.11:** Choice of  $\psi^H$  and coupling rate. Error and error indicator, effectivity index and the data approximation error (from left to right) of the four parameter choices. The jagged convergence plots are a consequence of the cancelation of errors, see Remark 80; we emphasize that this effect also appears in the single-scale DWR FEM case.

## Chapter 8. Goal-Oriented Adaptive FE-HMM



**Figure 8.6.12:** Different macro-micro coupling for the primal and dual solution  $u^{HMM}$  and  $z^{HMM}$ , respectively and the estimator  $\eta_H(K)$ . Error and error indicator, effectivity index and the data approximation error (from left to right) of the three parameter choices for the DWR-FE-HMM.

## 8.7 Summary

We have derived the first *a posteriori* error estimates for the FE-HMM for errors in quantities of interest (often needed to make a proper engineering design decision). The techniques we have developed in Chapter 6 and Chapter 7, which relate the special multiscale FE-HMM components to the equivalent classical, single-scale counterparts, allowed us to follow the analysis of classical, single scale goal-oriented adaptive FEM. Except for a data approximation term, the *a posteriori* error representation formula was obtained without special assumptions on the structure of the oscillating tensor.

We derived and analyzed the adaptive, goal-oriented FE-HMM, which is a multiscale counterpart of the classical dual-weighted residual method by Bangerth and Rannacher [28]. The efficiency of the method was demonstrated in numerical experiments with a quantity of interest chosen to be the point-wise directional derivative.





## 9 Conclusion and Outlook of Part II

In this thesis we have proposed *adaptive* and *efficient* numerical methods for *multiscale problems* that will provide engineers with *reliable* data that allow them to make proper engineering *design decisions*.

We derived the first rigorous *a posteriori* results for the FE-HMM in the energy norm of the physical variables. The *a posteriori* error analysis allowed us to find estimates on the accuracy (and therefore the reliability) of a numerical FE-HMM solution. The error analysis further allowed us to derive explicit localized error indicators that we use to drive a robust and reliable adaptive mesh refinement. Since these error indicators depend on macroscopic data (such as the macroscopic diffusion tensor) that are not readily available, we constructed error indicators that only depend on the available macro and micro FE solutions (computed previously). The crucial components for this construction are the so-called *multiscale bilinear forms, fluxes and jumps*. We showed that it is possible (thanks to a representation formula) to relate our multiscale strategy to adaptive strategies for single-scale adaptive FEM [20, 113]. Furthermore, the derived *a posteriori* estimates are consistent with classical explicit residual-based *a posteriori* error estimators applied to the homogenized problems in the case of periodic tensors and resolved micro calculations.

Up to a data approximation term, upper and lower bounds were obtained without specific structure assumptions (as periodicity, random stationarity) on the oscillating tensor of the elliptic problem. The error indicator provides information that is used to refine the macro mesh. As singularities in the micro problems could only arise in the micro scale of the conductivity tensor, we used a uniform refinement for the micro problems (there, standard adaptive FEM could be used). We proposed a strategy to couple the refinement of the micro mesh to the refinement of the corresponding macro element and demonstrated – under appropriate assumptions – that our strategy leads to an optimal convergence. Our numerical results demonstrated that the adaptive strategy is both reliable and efficient. We furthermore showed that through adaptivity, we can significantly improve the *efficiency* of multiscale methods.

We extended our adaptive, residual-based FE-HMM to support higher order FEs in the macro and micro spaces. The extension to higher order FE is a crucial component for the development of our goal-oriented adaptive FE-HMM, where higher-order dual solutions are

required.

In order to provide error estimates in quantities of interest (often needed for appropriate engineering design decisions), we derived the first *a posteriori* results for the FE-HMM in quantities of interest. These quantities of interest are expressed by bounded linear functionals. Our analysis is based on the framework we developed for the adaptive, residual-based FE-HMM. Except for a data approximation term, the *a posteriori* error representation formula was obtained without special assumptions on the structure of the oscillating tensor. The goal-oriented FE-HMM, a multiscale counterpart of the classical dual-weighted residual method by Bangerth and Rannacher [28], was derived and analyzed. Extensive numerical experiments were reported, which illustrate the efficiency and versatility of the method.

### 9.1 Outlook

There are still many areas of improvement that can provide the adaptive FE-HMM with better efficiency, an applicability to a broader set of problems and with an increased reliability of the error estimates. We summarize the most important suggestions in what follows.

**Efficiency.** In order to increase the efficiency of the adaptive FE-HMM, it is strongly suggested to use a reduced basis technique for the solution of the micro problems (see [101, 8]). The conductivity tensor of the micro problems often does not change significantly from one macro element to its neighbors (provided regularity w.r.t. the slow variable), and therefore this approach can lead to a remarkable reduction in computational cost. Furthermore, for the goal-oriented FE-HMM, we could employ interpolation techniques as described in [28] to obtain an approximate solution of the dual problem that avoids the computation of the full dual solution. An alternative is to reduce the degrees of freedom of the discretized dual problem by using specially designed FE spaces, e.g., involving bubble basis functions [96] or solve the dual solution on a refined mesh instead of using higher order polynomials with respect to the primal problem.

Our strategy for the macro-micro mesh refinement is based on *a priori* estimates. It is, however, not straightforward to obtain *a priori* estimates for errors in *general* quantities of interest that could be used to extract an (optimal) macro-micro coupling rate. We notice that the coupling rate could be steered adaptively using additional *a posteriori* error estimates on the micro problems.

**Applicability.** The adaptive FE-HMM has been developed for elliptic problems. It would be of interest to extend this multiscale adaptive framework for problems in linear elasticity [3] or for non-linear elliptic problems (see [15] for non-linear problems of non-monotone type) or non-linear quantities of interest.

**Reliability.** Arguably the most important area for improvement is however the reliability of the adaptive FE-HMM. For both the residual-based adaptive and the goal-oriented adaptive FE-HMM, we neither estimate the micro error nor the modeling error *a posteriori*. An *a*

---

*a posteriori* estimation of the micro error could easily be integrated in our adaptive algorithm using standard *a posteriori* error estimates on the micro problems. A rigorous adaptive strategy for an *a posteriori* estimate of modeling errors, though, is still to be derived.

For macro-to-micro multiscale methods as considered in this thesis, it is desirable to change the physical model in those regions of the computational domain where the size of the macro elements becomes smaller than the corresponding sampling domain. Combining model adaptivity [90] with the adaptive strategies proposed in this thesis is therefore of high interest. Furthermore, it is also desirable to adaptively select the size of the micro sampling domains (or the domain size when using oversampling techniques) based on *a posteriori* modeling error estimates, similar to the adaptive selection of the size of patches in [73, 74] in the context of the variational multiscale method.

For the adaptive FE-HMM in quantities of interest, there exist further challenges that concern the reliability of the estimates. We introduced confidence intervals that give an estimate of the *exact* solution in the quantity of interest. By doing so, we *assume* that the data approximation error is of the same order as the (estimated) error in the quantity of interest. Therefore, the confidence interval is rather an indication than a reliable bound. Finding a reliable estimate on the data approximation error (that in general involves a macro, micro and modeling error) that is inexpensive to compute is however not straightforward.

One of the most important challenges when it comes to the reliable estimation of error in quantities of interest is the substitution of the (unknown) exact dual solution by a numerical approximation. It was shown in [87] in the context of the single-scale DWR FE method, that replacing the exact dual solution by a numerical approximation (in a higher order FE space) can lead to a *significant underestimation* of the approximation error, thus causing the adaptive algorithm to stop even though the accuracy of the solution in the quantity of interest might be insufficient to make a proper engineering decision. In [87] a safeguarded DWR FEM is proposed. There has furthermore been very recent effort to develop guaranteed and fully computable bounds on the error in quantities of interest [21]. It could be of interest to extend these results in our multiscale context.



## Bibliography

- [1] A. Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM Journal on Scientific Computing*, 23(6):2041–2054, 2002.
- [2] A. Abdulle. On A Priori Error Analysis of Fully Discrete Heterogeneous Multiscale FEM. *SIAM Multiscale Modeling & Simulation*, 4(2):447–459, 2005.
- [3] A. Abdulle. Analysis of a Heterogeneous Multiscale FEM for Problems in Elasticity. *Math. Mod. Meth. Appl. Sci. (M3AS)*, 6:615–635, 2006.
- [4] A. Abdulle. Heterogeneous multiscale method with quadrilateral elements. *Numerical mathematics and advanced applications, Springer, Berlin*, pages 743–751, 2006.
- [5] A. Abdulle. Multiscale method based on discontinuous Galerkin methods for homogenization problems. *Comptes rendus-Mathématique*, 346(1-2):97–102, 2008.
- [6] A. Abdulle. The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. *GAKUTO International Series, Math. Sci. Appl.*, 31:135–181, 2009.
- [7] A. Abdulle. Discontinuous Galerkin Finite Element Heterogeneous Multiscale Method for Elliptic Problems with Multiple Scales. *To appear in Math. Comp.*, 2011.
- [8] A. Abdulle and Y. Bai. A reduced basis heterogeneous multiscale method. *In preparation*, 2011.
- [9] A. Abdulle and B. Engquist. Finite element heterogeneous multiscale methods with near optimal computational complexity. *SIAM Multiscale Modeling & Simulation*, 6(4):1059–1084, 2007.
- [10] A. Abdulle and A. Nonnenmacher. A posteriori error analysis of the heterogeneous multiscale method for homogenization problems. *Comptes Rendus Mathématique*, 347(17-18):1081–1086, 2009.
- [11] A. Abdulle and A. Nonnenmacher. A short and versatile finite element multiscale code for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 198(37-40):2839–2859, 2009.

## Bibliography

---

- [12] A. Abdulle and A. Nonnenmacher. Adaptive finite element heterogeneous multiscale method for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 200:2710–2726, 2011.
- [13] A. Abdulle and A. Nonnenmacher. Adaptive finite element heterogeneous multiscale method for homogenization problems in quantities of interest. *In preparation*, 2011.
- [14] A. Abdulle and Ch. Schwab. Heterogeneous Multiscale FEM for Diffusion Problems on Rough Surfaces. *SIAM Multiscale Modeling & Simulation*, 3(1):195–220, 2005.
- [15] A. Abdulle and G. Vilmart. Analysis of the finite element heterogeneous multiscale method for nonmonotone elliptic homogenization problems. *Submitted for publication*, 2011.
- [16] A. Abdulle and G. Vilmart. On micro macro errors of the heterogeneous multiscale method for parabolic problems. *Submitted for publication*, 2011.
- [17] F.F. Abraham, R. Walkup, H. Gao, M. Duchaineau, T. Diaz De La Rubia, and M. Seager. Simulating materials failure by using up to one billion atoms and the world's fastest computer: Work-hardening. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9):5783–5787, 2002.
- [18] S. Agmon. *Lectures on elliptic boundary value problems*. Chelsea Pub Co, 2010.
- [19] M. Ainsworth and J.T. Oden. A posteriori error estimation in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 142(1-2):1–88, 1997.
- [20] M. Ainsworth and T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley, New York, 2000.
- [21] M. Ainsworth and R. Rankin. Guaranteed computable bounds on quantities of interest in finite element computations. *submitted to the International Journal of Numerical Methods in Engineering*, 2011.
- [22] J. Alpert, C. Carstensen, and S.A. Funken. Remarks around 50 lines of Matlab: short finite element implementation. *Numerical Algorithms*, 20(2):117–137, 1999.
- [23] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23:1482, 1992.
- [24] I. Babuška, G. Caloz, and J.E. Osborn. Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM Journal on Numerical Analysis*, 31(4):945–981, 1994.
- [25] I. Babuška and J.E. Osborn. Generalized finite element methods: their performance and their relation to mixed methods. *SIAM Journal on Numerical Analysis*, 20(3):510–536, 1983.

- 
- [26] I. Babuška and W.C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15(4):736–754, 1978.
- [27] I. Babuška and T. Strouboulis. *The finite element method and its reliability*. Oxford University Press, USA, 2001.
- [28] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations*. Birkhäuser Verlag, Basel, 2003.
- [29] R.E. Bank, A.H. Sherman, and A. Weiser. Some refinement algorithms and data structures for regular local mesh refinement. *Scientific Computing, Applications of Mathematics and Computing to the Physical Sciences*, 1983.
- [30] S. Bartels, C. Carstensen, and A. Hecht. P2Q2Iso2D= 2D isoparametric FEM in Matlab. *Journal of Computational and Applied Mathematics*, 192(2):219–250, 2006.
- [31] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta numerica*, 10:1–102, 2001.
- [32] A. Bensoussan, J.L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures (Studies in mathematics and its applications)*. Elsevier North-Holland, 1978.
- [33] Ch. Bernardi, Y. Maday, and F. Rapetti. *Discrétisations Variationnelles De Problèmes Aux Limites Elliptiques*. Springer, Berlin, 2004.
- [34] L. Bers, F. John, and M. Schechter. Partial differential equations. *Lectures in Applied Mathematics, Proceedings of the Summer Seminar, Boulder, CO*, 1957.
- [35] M.J. Biercuk, M.C. Llaguno, M. Radosavljevic, J.K. Hyun, A.T. Johnson, and J.E. Fischer. Carbon nanotube composites for thermal management. *Applied Physics Letters*, 80:2767–2769, 2002.
- [36] A. Bourgeat and A. Piatnitski. Approximations of effective coefficients in stochastic homogenization. *Annales de l'Institut Henri Poincaré/Probabilités et statistiques*, 40(2):153–165, 2004.
- [37] S. Boyaval. Application of reduced basis approximation and a posteriori error estimation to homogenization theory. *SIAM Multiscale Modeling and Simulation*, 2007.
- [38] S. Boyaval, C. Le Bris, Y. Maday, C. Nguyen, and A. Patera. A Reduced Basis Approach for Variational Problems with Stochastic Parameters: Application to Heat Conduction with Variable Robin Coefficient. 2008.
- [39] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 2007.
- [40] S. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, Berlin, 2008.

## Bibliography

---

- [41] L. Chen. *iFEM: an innovative finite element methods package in MATLAB*. In *Preparation*, 2008.
- [42] L. Chen. Short implementation of bisection in MATLAB. In *Recent advances in computational sciences: selected papers from the International Workshop on Computational Sciences and its Education, Beijing, China, 29-31 August 2005*, pages 318–333. World Scientific Pub Co Inc, 2008.
- [43] L. Chen and C.S. Zhang. AFEM@ matlab: a Matlab package of adaptive finite element methods. Technical report, 2006.
- [44] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial Mathematics, 2002.
- [45] D. Cioranescu and P. Donato. *An Introduction to Homogenization (Oxford Lecture Series in Mathematics and Its Applications, 17)*. Oxford University Press, USA, 2000.
- [46] P. Clément. Approximation by finite element functions using local regularization. *Rev. Francaise Automat. Informat. Recherche Operationnelle Ser. Rouge Anal. Numer*, 9:77–84, 1975.
- [47] R. Cools. Constructing cubature formulae: the science behind the art. *Acta Numerica*, 6:1–54, 1997.
- [48] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM Journal on Numerical Analysis*, pages 1106–1124, 1996.
- [49] W. E. *Principles of Multiscale Modeling*. Cambridge University Press, To appear.
- [50] W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci*, 1:87–132, 2003.
- [51] W. E, B. Engquist, and Z. Huang. Heterogeneous multiscale method: a general methodology for multiscale modeling. *Physical Review B*, 67(9):092101, 2003.
- [52] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. The Heterogeneous Multiscale Method: A Review. *Commun Comput Phys*, 2:367–450, 2007.
- [53] W. E, P Ming, and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *American Mathematical Society*, 18(1):121–156, 2005.
- [54] Y. Efendiev and T. Y. Hou. *Multiscale Finite Element Methods: Theory and Applications*. Springer, Berlin, 2009.
- [55] A. Ern and J.L. Guermond. *Theory and practice of finite elements*. Springer Verlag, 2004.
- [56] R. Ewing, O. Iliev, R. Lazarov, I. Rybak, and J. Willems. A simplified method for upscaling composite materials with high contrast of the conductivity. *SIAM journal on scientific computing*, 31(4):2568–2586, 2009.



- 
- [57] M.G.D Geers and V.G. Kouznetsova. Scale transitions in solid mechanics based on computational homogenization. CISM Lecture Notes, 2010.
- [58] M.B. Giles and E. Süli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numerica*, 11:145–236, 2002.
- [59] T. Grätsch and K.J. Bathe. A posteriori error estimation techniques in practical finite element analysis. *Computers & Structures*, 83(4-5):235–265, 2005.
- [60] P. Henning and M. Ohlberger. A-posteriori error estimate for a heterogeneous multi-scale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift. Preprint 09/09 - N, FB 10 , Universität Münster, 2009.
- [61] P. Henning and M. Ohlberger. The heterogeneous multiscale finite element method for elliptic homogenization problems in perforated domains. *Numerische Mathematik*, 113, Issue 4 (2009):601 – 629, 2009.
- [62] P. Henning and M. Ohlberger. A-posteriori error estimation for a heterogeneous multi-scale method for monotone operators and beyond a periodic setting. Preprint 01/11 - N, FB 10 , Universität Münster, 2011.
- [63] V.H. Hoang and Ch. Schwab. High-dimensional finite elements for elliptic problems with multiple scales. *Multiscale Modeling & Simulation*, 3:168–194, 2005.
- [64] T.Y. Hou and X.H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134(1):169–189, 1997.
- [65] T.Y. Hou, X.H. Wu, and Z. Cai. Convergence of a Multiscale Finite Element Method for Elliptic Problems with Rapidly Oscillating Coefficients. *Mathematics of Computation*, 68(227):913–943, 1999.
- [66] A. Huerta, P. Díez, A. Rodríguez-Ferran, and G. Pijaudier-Cabot. Error estimation and adaptive finite element analysis of softening solids. *Studies in Applied Mechanics*, 47:333–347, 1998.
- [67] T.J.R. Hughes. Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods\* 1. *Computer methods in applied mechanics and engineering*, 127(1-4):387–401, 1995.
- [68] T.J.R. Hughes, G.R. Feijoo, L. Mazzei, and J.B. Quincy. The variational multiscale method—a paradigm for computational mechanics. *Computer methods in applied mechanics and engineering*, 166(1-2):3–24, 1998.
- [69] B. Jakobsen and F. Rosendahl. The Sleipner platform accident. *Structural Engineering International*, 4(3):190–193, 1994.

## Bibliography

---

- [70] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag Telos, 1994.
- [71] V. Kouznetsova, W.A.M. Brekelmans, and F.P.T. Baaijens. An approach to micro-macro modeling of heterogeneous materials. *Computational Mechanics*, 27:37–48, 2001.
- [72] O.A. Ladyzhenskaya. *Boundary value problems of mathematical physics*. Springer-Verlag New York, 1985.
- [73] M.G. Larson and A. Målqvist. Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems. *Multiscale Methods in Science and Engineering*, pages 181–193, 2005.
- [74] M.G. Larson and A. Målqvist. Adaptive variational multiscale methods based on a posteriori error estimation: Energy norm estimates for elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 196(21-24):2313–2324, 2007.
- [75] F. Larsson and K. Runesson. Adaptive bridging of scales in continuum modeling based on error control. *International Journal for Multiscale Computational Engineering*, 6(4), 2008.
- [76] A.M. Matache, I. Babuška, and Ch. Schwab. Generalized p-FEM in homogenization. *Numerische Mathematik*, 86(2):319–375, 2000.
- [77] J. Melenk. *hp-Finite Element Methods for Singular Perturbations*. Springer, Berlin, 2002.
- [78] C. Miehe, J. Schröder, and C. Bayreuther. On the homogenization analysis of composite materials based on discretized fluctuations on the micro-structure. *Acta mechanica*, 155(1):1–16, 2002.
- [79] P. Ming and X. Yue. Numerical methods for multiscale elliptic problems. *Journal of Computational Physics*, 214(1):421–445, 2006.
- [80] W.F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Transactions on Mathematical Software (TOMS)*, 15(4):326–347, 1989.
- [81] W.F. Mitchell. Optimal multilevel iterative methods for adaptive grids. *SIAM Journal on Scientific and Statistical Computing*, 13:146–167, 1992.
- [82] P. Morin, R.H. Nochetto, and K.G. Siebert. Convergence of adaptive finite element methods. *SIAM review*, 44(4):631–658, 2002.
- [83] N. Neuss, W. Jäger, and G. Wittum. Homogenization and multigrid. *Computing*, 66(1):1–26, 2001.
- [84] G. Nguetseng. A general convergence result for a functional related to the theory of homogenization. *SIAM Journal on Mathematical Analysis*, 20:608, 1989.

- 
- [85] R. H. Nochetto. Lecture notes in Adaptive Finite Element Methods for Elliptic PDE. Technical report, 2006 CNA Summer School, Probabilistic and Analytical Perspectives in Contemporary PDE, 2006.
- [86] R.H. Nochetto. Adaptive finite element methods for elliptic PDEs. Lecture Notes, 2007.
- [87] R.H. Nochetto, A. Veiser, and M. Verani. A safeguarded dual weighted residual method. *IMA journal of Numerical Analysis*, 29(1):126–140, 2009.
- [88] J.T. Oden and S. Prudhomme. Goal-oriented error estimation and adaptivity for the finite element method. *Computers & Mathematics with Applications*, 41(5-6):735 – 756, 2001.
- [89] J.T. Oden and S. Prudhomme. Estimation of modeling error in computational mechanics. *Journal of Computational Physics*, 182(2):496–515, 2002.
- [90] J.T. Oden, S. Prudhomme, A. Romkes, and P.T. Bauman. Multiscale modeling of physical phenomena: Adaptive control of models. *SIAM Journal on Scientific Computing*, 28(6):2359–2392, 2006.
- [91] J.T. Oden and J.N. Reddy. *An introduction to the mathematical theory of finite elements*, volume 79. Wiley New York, 1976.
- [92] M. Ohlberger. A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems. *SIAM Multiscale Modeling & Simulation*, 4:88–114, 2005.
- [93] I. Özdemir, W.A.M. Breckelmanns, and M.G.D. Geers. Computational homogenization for heat conduction in heterogeneous solids. *International Journal for Numerical Methods in Engineering*, 73(2):185–204, 2008.
- [94] G.A. Pavliotis and A.M. Stuart. *Multiscale Methods: Averaging and Homogenization*. Springer, 2008.
- [95] C. Prud’homme, D.V. Rovas, K. Veroy, L. Machiels, Y. Maday, A.T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *Journal of Fluids Engineering*, 124:70–80, 2002.
- [96] S. Prudhomme and J.T. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Computer Methods in Applied Mechanics and Engineering*, 176(1-4):313–331, 1999.
- [97] S. Prudhomme, J.T. Oden, T. Westermann, J. Bass, and M.E. Botkin. Practical methods for a posteriori error estimation in engineering applications. *Int. J. Numer. Methods Eng*, 56(8):1193–1224, 2003.
- [98] A. Quarteroni. *Numerical approximation of partial differential equations*. Springer-Verlag, Berlin New York, 1994.

## Bibliography

---

- [99] T. Richter. Funktionalorientierte Gitteroptimierung für die Finite Elemente Methode. Master's thesis, University of Heidelberg, 2001.
- [100] M.C. Rivara. Design and data structure of fully adaptive, multigrid, finite-element software. *ACM Transactions on Mathematical Software (TOMS)*, 10(3):242–264, 1984.
- [101] G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):1–47, 2007.
- [102] Sandia National Labs. Cubit, <http://cubit.sandia.gov>.
- [103] H.J. Schmid. On cubature formulae with a minimal number of knots. *Numerische Mathematik*, 31(3):281–297, 1978.
- [104] A. Schmidt and K. Siebert. *Design of Adaptive Finite Element Software: the Finite Element Toolbox Alberta*. Springer, Berlin, 2005.
- [105] Ch. Schwab. *p- and hp-Finite Element Methods*. Clarendon Press, Oxford, 1998.
- [106] E. Süli. Review of the book *Adaptive finite element methods for differential equations by W. Bangerth and R. Rannacher*. *Math. Comp.*, 74:1033–1052, 2004.
- [107] T. Tao. *An epsilon of room, I, real analysis : pages from year three of a mathematical blog*. American Mathematical Society, Providence, R.I, 2010.
- [108] CPS Technologies. [http://www.alsic.com/\\_blog/CPS\\_Technologies\\_Blog/post/Introduction\\_What\\_is\\_ALSiC/](http://www.alsic.com/_blog/CPS_Technologies_Blog/post/Introduction_What_is_ALSiC/), March 15, 2011.
- [109] K. Terada and N. Kikuchi. A class of general algorithms for multi-scale analyses of heterogeneous media. *Computer methods in applied mechanics and engineering*, 190(40-41):5427–5464, 2001.
- [110] TOP500 project. [www.top500.org](http://www.top500.org), 2011.
- [111] S. Torquato. *Random Heterogeneous Materials*. Springer, 2005.
- [112] R. Verfürth. A posteriori error estimation and adaptive mesh-refinement techniques. *Proceedings of the fifth international conference on Computational and applied mathematics table of contents*, pages 67–83, 1994.
- [113] R. Verfürth. *A Review of A Posteriori Error Estimation & Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, 1996.
- [114] R. Verfürth. *Adaptive Finite Element Methods*, Lecture Notes Winter Term 2010/11, 2010.
- [115] T.C. Wallstrom, S. Hou, M.A. Christie, L.J. Durlofsky, and D.H. Sharp. Accurate scale up of two phase flow using renormalization and nonuniform coarsening. *Computational Geosciences*, 3(1):69–87, 1999.

- [116] D. Werner. *Funktionalanalysis*. Springer, Berlin, 2007.
- [117] Q. Yu and J. Fish. Multiscale asymptotic homogenization for multiphysics problems with multiple spatial and temporal scales: a coupled thermo-viscoelastic example problem. *International journal of solids and structures*, 39(26):6429–6452, 2002.
- [118] X. Yue and W. E. The local microscale problem in the multiscale modeling of strongly heterogeneous media: Effects of boundary conditions and cell size. *Journal of Computational Physics*, 222(2):556–572, 2007.
- [119] O. Zienkiewicz. *The Finite Element Method*. Elsevier/Butterworth-Heinemann, Amsterdam, 2005.
- [120] O. Zienkiewicz and J. Zhu. A simple error estimator and adaptive procedure for practical engineering analysis. *International Journal for Numerical Methods in Engineering*, 24(2):337–357, 1987.
- [121] C. Zweben. Advances in composite materials for thermal management in electronic packaging. *JOM Journal of the Minerals, Metals and Materials Society*, 50(6):47–51, 1998.



# Curriculum Vitae

## Personal Data

Date of birth June 28, 1980  
Nationality German

## Education

### Ph.D. in Mathematics

- 2009-2011 **Ph.D. in Mathematics**, *EPFL*, Lausanne, Switzerland.  
In the Chair of Numerical Analysis and Computational Mathematics (ANMC).  
Advisor: Prof. A. Abdulle.
- 2007-2008 **Ph.D. candidate in Mathematics**, *University of Edinburgh*, UK.  
Advisor: Prof. A. Abdulle.

### M.Sc. in Physics

- 2000-2007 **Diploma (M.Sc.) in Physics**, *Eberhard Karls Universität Tübingen*, Germany.  
Specialization on Scientific Computing.  
M.Sc. thesis: *Dynamical low-rank Approximation*. Advisor: Prof. Ch. Lubich.
- 2003-2004 **Visiting student**, *University of Maryland*, College Park, USA.

## Publication List

- [1] A. Abdulle and A. Nonnenmacher. A posteriori error analysis of the heterogeneous multiscale method for homogenization problems. *Comptes Rendus Mathematique*, 347(17-18):1081–1086, 2009.
- [2] A. Abdulle and A. Nonnenmacher. A short and versatile finite element multiscale code for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 198(37-40):2839–2859, 2009.
- [3] A. Abdulle and A. Nonnenmacher. Adaptive finite element heterogeneous multiscale method for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 200:2710–2726, 2011.
- [4] A. Abdulle and A. Nonnenmacher. Adaptive finite element heterogeneous multiscale method for homogenization problems in quantities of interest. *In preparation*, 2011.
- [5] A. Nonnenmacher and C. Lubich. Dynamical low-rank approximation: applications and numerical experiments. *Mathematics and Computers in Simulation*, 79(4):1346–1357, 2008.