EPFL

Thèse n° 7750

Novel corrector problems with exponential decay of the resonance error for numerical homogenization

Présentée le 2 octobre 2020

à la Faculté des sciences de base
Chaire d'analyse numérique et mathématiques computationnelles
Programme doctoral en mathématiques

pour l'obtention du grade de Docteur ès Sciences

par

Edoardo PAGANONI

Acceptée sur proposition du jury

Prof. F. Eisenbrand, président du jury
Prof. A. Abdulle, directeur de thèse
Dr F. Legoll, rapporteur
Prof. O. Runborg, rapporteur
Prof. A. Buffa, rapporteuse

École polytechnique fédérale de Lausanne

2020

Possunt,
quia posse videntur.
— *Aeneid*, Virgil

Ever tried.
Ever failed.
No matter.
Try again.
Fail again.
Fail better.
— *Worstward Ho*, Samuel Beckett

Acknowledgements

I believe that being a doctoral student is a unique opportunity to deal with new unexplored scientific challenges. Along this path, I not only learnt new mathematical concepts or methods but, most importantly, I matured as researchers and grew as a person. The critical sense, the self-discipline, and the ability to understand scientific works, which I have developed during the PhD, will always be part of my personal skillset.

I owe to my supervisor, Prof. Assyr Abdulle, the opportunity of undertaking this journey. I wish to thank him for teaching me to persevere tenaciously for my goals, and I will remind this lesson in my future work and life.

I owe a lot to my collaborator and mentor, Dr. Doghonay Arjmand, for having spent countless hours on blackboard discussions to solve many research questions. Without his precious feedback and motivating passion, many results in this thesis would not have been possible.

Furthermore, I wish to express my gratitude to the doctoral jury members, Prof. Annalisa Buffa, Prof. Frederic Legoll, and Prof. Olof Runborg, for having read the manuscript and for their positive feedback on my work. Equally, I thank the president of the doctoral committee, Prof. Friedrich Eisenbrandt, for mediating the discussion during the oral exam. I must thank as well Virginie Ledouble and Anna Dietler for managing all the paperwork in a very professional way.

These years at EPFL would not have been the same without all my friends and colleagues from the ANMC and MATHICSE groups. I am very thankful to all of them for the friendly and supportive environment that they created, retreat hikes, lively discussions on veganism, and cheerful moments in front of a pizza at Luigia.

Over the past years, all my dearest friends from Turin took different paths and spread across Europe. Nonetheless, our bonds did not get loose but they became even stronger. I thank Andrea, Enrico, Francesca, Marco, Mario and Umberto for all the pleasant moments we spent together.

A person in particular supported and pushed me more than anyone else during these years. Thanks to Chiara, *I knew myself, for in truth I had never known myself.*

Last, but not least, I thank my parents and my little sister for all the support they provided in all my choices.

Lausanne, September 12, 2020 Edoardo Paganoni

Abstract

Multiscale problems, such as modelling flows through porous media or predicting the mechanical properties of composite materials, are of great interest in many scientific areas. Analytical models describing these phenomena are rarely available, and one must recur to numerical simulations. This represents a great computational challenge, because of the prohibitive computational cost of resolving the small scales. Multiscale numerical methods are therefore necessary to solve multiscale problems within reasonable computational time and resources. In particular, numerical homogenization techniques aim to capture the macroscopic behaviour with equations whose coefficients are computed numerically from the solutions of corrector problems at the microscale. A lack of knowledge of the coupling conditions between the micro- and the macro-scales brings in the so-called resonance error, which affects the accuracy of all multiscale methods. This source of error often dominates the numerical discretization errors and increasing its rate of decay is crucial for improving the accuracy of multiscale methods.

In this work, we propose two novel upscaling schemes with arbitrarily high convergence rates of the resonance error to approximate the homogenized coefficients of scalar, linear second order elliptic differential equations. The first one is based on a parabolic equation, inspired by a model employed to compute the effective diffusive coefficients in stochastic diffusion processes. By using the approximation properties of smooth filtering functions, the homogenized coefficients can be approximated with arbitrary rates of accuracy. This claim is proved through an *a priori* convergence analysis, under the assumption of smooth periodic multiscale coefficients. Numerical experiments verify the expected convergence rates also under more general assumptions, such as discontinuous and random coefficients. The second method originates from the first by integrating the parabolic equation over a finite time interval. This method is referred to as the modified elliptic approach, because of the presence of a right-hand side which can be interpreted in terms of continuous semigroups and can be approximated numerically by Krylov subspace methods. The same convergence results as in the parabolic approach hold true. As a last step, a convergence analysis of the resonance error for the modified elliptic approach in the context of equations with random coefficients is performed. In this case, the resonance error is composed of a variance and a bias term, which can be bounded from above by a function decaying to zero. Numerical experiments reveal that the convergence rate of the resonance error for random coefficients is hampered, in comparison to the case of periodic coefficients, but the modified elliptic approach nevertheless outperforms standard methods.

Abstract

Key words: Multiscale problems, multiscale numerical methods, numerical homogenization, resonance error, corrector problems, stationary random processes, heterogeneous multiscale method, *a priori* convergence analysis.

Résumé

Les problèmes multi-échelles, comme la modélisation des écoulements dans les milieux poreux ou la prévision des propriétés mécaniques des matériaux composites, présentent un grand intérêt dans de nombreux domaines scientifiques. Des modèles analytiques permettant d'expliquer le comportement macroscopique sont rarement disponibles, et ces problèmes doivent être traités par des simulations numériques prenant en compte la présence d'échelles multiples. Cela représente un grand défi computationel, en raison du coût prohibitif de la résolution de toutes les échelles. Des méthodes numériques multi-échelles sont donc nécessaires pour résoudre les problèmes multi-échelles dans des délais et des ressources de calcul raisonnables. En particulier, les techniques d'homogénéisation numérique visent à résoudre des équations macroscopiques dont les coefficients sont calculés numériquement à partir des solutions des problèmes de correction à la micro-échelle. Un manque de connaissance des conditions de couplage entre les échelles microscopiques et macroscopiques entraîne ce que l'on appelle l'erreur de résonance, qui affecte toutes les méthodes multi-échelles. Cette erreur domine souvent celles dues à la discrétisation numérique et l'augmentation de son taux de décroissance est cruciale pour améliorer la précision des méthodes multi-échelles.

Dans ce travail, nous proposons deux nouveaux schémas de couplage entre les échelles microscopiques et macroscopiques avec des taux de convergence arbitrairement élevés de l'erreur de résonance pour l'approximation des coefficients homogénéisés des équations différentielles elliptiques du second ordre scalaires et linéaires. Le premier est basé sur une équation parabolique, inspirée par un modèle de calcul des coefficients effectifs de diffusion dans les processus de diffusion stochastique. En utilisant les propriétés d'approximation des fonctions de filtrage lisse, les coefficients homogénéisés peuvent être approximés avec des taux de précision arbitraires. Cette affirmation est prouvée par une analyse de convergence *a priori*, sous l'hypothèse de coefficients multi-échelles périodiques et lisses. Des expériences numériques vérifient les taux de convergence attendus également sous des hypothèses plus générales, comme des coefficients discontinus et aléatoires. La seconde méthode est issue de la précédente en intégrant l'équation parabolique sur un intervalle de temps fini. Cette méthode est appelée approche elliptique modifiée, en raison de la présence d'un terme dans le membre de droite qui peut être interprété à l'aide de semigroupes continus et qui peut être approximé numériquement par les méthodes de Krylov. Les mêmes résultats de convergence que dans l'approche parabolique restent valables. En dernier lieu, une analyse de convergence de l'erreur de résonance pour l'approximation elliptique modifiée dans le contexte des coefficients aléatoires est effectuée. Dans ce cas, l'erreur de résonance est composée d'un terme de

variance et d'un terme de biais, qui peuvent être borné par une fonction tendant vers zéro. Les expériences numériques révèlent que le taux de convergence de l'erreur de résonance pour les coefficients aléatoires est fortement ralenti par rapport au cas des coefficients périodiques, mais l'approche elliptique modifiée reste plus performante que des méthodes standards.

Mots clés : Problèmes multi-échelles, méthodes numériques multi-échelles, homogénéisation numérique, erreur de résonance, problèmes des correcteurs, processus aléatoires stationnaires, méthode multi-échelles hétérogène, analyse de convergence *a priori*.

Notation

We will use the following notations throughout the exposition:

- The Sobolev space $W^{k,p}(D)$ is defined as

$$W^{k,p}(D) := \{f : D^\gamma f \in L^p(D) \text{ for all multi-index } \gamma \text{ with } |\gamma| \leq k\}.$$

  The norm of a function $f \in W^{k,p}(D)$ is given by

$$\|f\|_{W^{k,p}(D)} := \begin{cases} \left(\sum_{|\gamma|\leq k} \int_D |D^\gamma f(x)|^p \, dx\right)^{1/p} & (1 \leq p < \infty) \\ \sum_{|\gamma|\leq k} \text{ess sup}_D |D^\gamma f| & (p = \infty). \end{cases}$$

- The space $H_0^1(D)$ is the closure in the $W^{1,2}$-norm of $C_c^\infty(D)$, the space of infinitely differentiable functions with compact support in D. The norm associated with $H_0^1(D)$ is

$$\|f\|^2_{H^1(D)} := \|f\|^2_{L^2(D)} + \|\nabla f\|^2_{L^2(D)}.$$

  Due to the Poincaré inequality, an equivalent norm in $H_0^1(D)$ is given by

$$\|f\|_{H_0^1(D)} := \|\nabla f\|_{L^2(D)}.$$

  We will use this second notation for the H_0^1-norm.

- We use the notation $\langle f, g\rangle_{L^2(D)} := \int_D fg \, dx$ to denote the L^2 inner product over D.

- $H^{-1}(D)$ is the dual space of $H_0^1(D)$. It can be characterized as the set of distributional derivatives of $L^2(D)$-functions: For any $F \in H^{-1}(D)$, there exist $f_0, f_1, \ldots, f_d \in L^2(D)$ such that

$$\langle F, v\rangle_{H^{-1}(D), H_0^1(D)} = \langle f_0, v\rangle_{L^2(D)} + \sum_{i=1}^d \langle f_i, D_i v\rangle_{L^2(D)}.$$

- The space $H_{div}(D)$ is defined as

$$H_{div}(D) := \{f : f \in [L^2(D)]^d \text{ and } \nabla \cdot f \in L^2(D)\}.$$

  The norm associated with $H_{div}(D)$ is

$$\|f\|^2_{H_{div}(D)} := \|f\|^2_{L^2(D)} + \|\nabla \cdot f\|^2_{L^2(D)}.$$

- The space $L^2_0(K)$ is defined as

$$L^2_0(K) = \left\{ f \in L^2(K) : \int_K f \, dx = 0 \right\}.$$

  It is an Hilbert space with respect to the L^2-inner product.

- The space $W^1_{per}(K)$ is defined as the closure of

$$\left\{ f \in C^\infty_{\text{per}}(K) : \int_K f \, dx = 0 \right\}$$

  for the H^1-norm. Thanks to the Poincaré-Wirtinger inequality, an equivalent norm in $W^1_{per}(K)$ is

$$\|f\|_{W^1_{per}(K)} = \|\nabla f\|_{L^2(K)}.$$

  Moreover, $W^1_{per}(K)$ is an Hilbert space for the inner product

$$(u, v)_{W^1_{per}(K)} = \int_K \nabla u \cdot \nabla v.$$

- The space $W^1_{per}(K)'$ is the dual space of $W^1_{per}(K)$. It can be identified with the space

$$\left\{ F \in H^{-1}(K) : \langle F, c \rangle_{W^1_{per}(K)', W^1_{per}(K)} = 0, \forall c \in \mathbb{R}^d \right\}.$$

- Let f belong to the Bochner space $L^p(0,T;X)$, where X is a Banach space. Then the norm associated with this space is defined as

$$\|f\|_{L^p(0,T;X)} := \left(\int_0^T \|f\|^p_X \, dt \right)^{\frac{1}{p}}.$$

- $\mathbf{L}^2(D)$ is the space of square-integrable vector functions $\mathbf{f} : D \mapsto \mathbb{R}^d$, $\mathbf{f} = (f_1, \ldots, f_d)$, with $f_i \in L^2(D)$. $\mathbf{L}^2_{loc}(\mathbb{R}^d)$ is the space of *locally* square-integrable vector functions.

- $\mathbf{L}^2_{pot}(\mathbb{R}^d)$ is the space of *vortex-free* vector fields:

$$\mathbf{L}^2_{pot}(\mathbb{R}^d) = \left\{ \mathbf{f} \in \mathbf{L}^2_{loc}(\mathbb{R}^d) : \int_{\mathbb{R}^d} f_i \frac{\partial \varphi}{\partial x_j} - f_j \frac{\partial \varphi}{\partial x_i} \, dx = 0, \quad \forall \varphi \in C^\infty_0(\mathbb{R}^d) \right\}.$$

  Any vortex-free (or *irrotational*) field can be expressed as the gradient of a potential function $u \in H^1_{loc}(\mathbb{R}^d)$: $\mathbf{f} = \nabla u$.

- $\mathbf{L}^2_{sol}(\mathbb{R}^d)$ is the space of *solenoidal* (i.e., divergence-free) vector fields:

$$\mathbf{L}^2_{sol}(\mathbb{R}^d) = \left\{ \mathbf{f} \in \mathbf{L}^2_{loc}(\mathbb{R}^d) : \int_{\mathbb{R}^d} f_i \frac{\partial \varphi}{\partial x_i} = 0, \quad \forall \varphi \in C^\infty_0(\mathbb{R}^d) \right\}.$$

- Cubes in $\mathbb{R}^d$ are denoted by $K_L := (-L/2, L/2)^d$. In particular, K is the unit cube $(-1/2, 1/2)^d$.
- We will use the notation $\fint_D f(x)\,dx$ to denote the average $\frac{1}{|D|}\int_D f(x)\,dx$ over a domain D.
- $\mathscr{M}(\alpha, \beta, D)$ denotes the class of *symmetric* matrix-valued function $a \in L^\infty(D, \mathbb{R}^{d\times d})$ such that

$$\alpha|\xi|^2 \le \left|\xi \cdot a^\varepsilon(x)\xi\right|, \quad \left|a^\varepsilon(x)\xi\right| \le \beta|\xi|, \quad \forall \xi \in \mathbb{R}^d, \text{ a.e. } x \in D, \tag{1}$$

  for $0 < \alpha \le \beta$. In the case $D = \mathbb{R}^d$ the short-hand notation $\mathscr{M}(\alpha, \beta)$ will be used.

Contents

1 Introduction

Multiscale phenomena are ubiquitous in many fields of science and engineering. In fact, several problems studied in physics, chemistry, biology, material sciences and engineering are characterized by the presence of multiple scales in space or time. Oftentimes, phenomena taking place at different scales are not independent, but mutually influence each other. As a consequence, studying a system at a given scale disregarding what happens at smaller (or larger) scales often leads to wrong physical interpretations or to the need of employing empirical models.

Multiscale problems can be very different in nature and each can even involve different physical models. For example, the mechanical behaviour of a solid can be influenced by its crystalline structure, hence a coupling between the continuum mechanical model and the discrete solid state physics is needed. The multiple scales can be present in space, but also in time, as in the case of chemical reactions where the concentrations of reagents differs of many orders of magnitude and processes evolve at different time scales [127]. The class of spatially multiscale problems can be separated into two sub-classes, depending on the local or global nature of the multiscale feature. Locally multiscale problems (denoted as type A problems in [53]) are those where the small scale heterogeneity is present only on a small portion of the domain. Typical examples are the evolution of cracks in elastic materials [128], the rise of boundary layers in fluid flows and small scale structures in fully developed turbulent flows [112]. This type of problems can be solved by locally refining the mesh or by using a local small scale model and, then, coupling this information with the macroscopic solution. Globally multiscale problems (denoted as type B problems in [53]) are those in which the microscale model is needed everywhere because the small scale pattern is repeated throughout the domain. For example, groundwater pollution through infiltration of a fluid in a porous medium [132], the effective properties of composite materials [93, 102], the use of small scale optical elements to produce meta-materials with negative refraction index [136] and the mechanical behaviour of porous media such as bones or soft tissue [69] are classical globally multiscale problems. In this case we cannot use the local refinement strategy and we have to rely on a macroscopic model in order to capture the solution. However, the macroscopic

constitutive relations may be missing and we need a microscopic model to recover them, together with a coupling scheme between the two models. This is the idea behind many multiscale numerical methods, like the Heterogeneous Multiscale Method (HMM) [54]. Other examples of multiscale and multiphysics models are collected in [4, 53, 64, 65].

The presence of several orders of magnitude between the scale of interest and small scale features represents a challenge under several aspects. From the analytical perspective, it may not be possible to derive a model describing the macroscopic behaviour because of the microscale effects. From the point of view of numerical computations, the need of resolving all the small scale features on a much larger computational domain dramatically increases the number of degrees of freedom, thus making the simulation of globally multiscale problems unfeasible even on modern supercomputers. Usually, the interest is not in capturing all the small scale variations of the phenomenon, but rather in capturing correctly the macroscopic behaviour. This is the motivation for developing novel multiscale numerical methods that could provide a reasonably accurate approximation of the solution with a number of degrees of freedom not constrained by the small scale. Most of the multiscale numerical methods suffer from the so-called resonance error, due to the coupling conditions between the micro and the macro scales, which eventually affects the reliability of the numerical upscaled solution. Therefore, many studies focused on the reduction of the resonance error and on the improvement of its convergence rate. The goal of the present study is to propose and study two novel micro-corrector equations with higher convergence rates of the resonance error.

1.1 General framework and multiscale problems

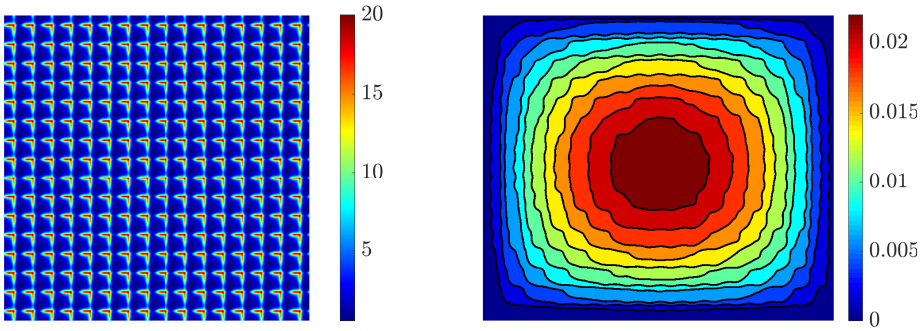
In this thesis we will consider the (globally) multiscale partial differential equation:

$$\begin{cases} -\nabla \cdot (a^\varepsilon(x)\nabla u^\varepsilon) = f & \text{in } D \subset \mathbb{R}^d, \\ u^\varepsilon = 0 & \text{on } \partial D, \end{cases} \tag{1.1}$$

where the multiscale structure is encoded in the diffusion coefficient a^ε, i.e. we assume that a^ε oscillates at a small scale $\varepsilon \ll |D|$, throughout the entire domain. In Figure 1.1a, an example of a multiscale coefficient field $a^\varepsilon(x) = a(x/\varepsilon)$ with oscillations only at the ε-scale is pictured. The solution u^ε has a *multiscale* behaviour in the sense that it varies both at a scale comparable to the size of D (the macroscale) and at a much smaller scale (the microscale), see Figure 1.1b. The multiscale behaviour is inherited from the microscale oscillations of the coefficient a^ε.

The problem of multiple scales: the example of locally periodic coefficients

In connection to the feature of scale separation is the number of scales present in the problem. In Figure 1.1a we pictured a coefficient field of the form $a^\varepsilon(x) = a(x/\varepsilon)$, with $\varepsilon = 1/16$. By construction, $a^\varepsilon(x)$ oscillates only at the ε-scale. However it is possible to enrich the set of



(a) $a^\varepsilon(x)$. (b) $u^\varepsilon(x)$.

Figure 1.1 – A multiscale diffusion field with the solution for the problem with $f = 1$.

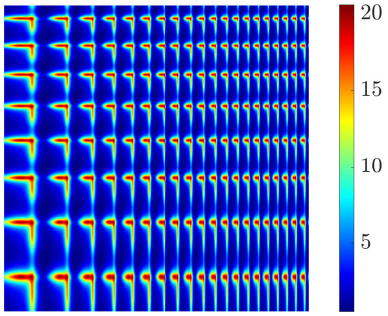


Figure 1.2 – A locally periodic coefficient field.

scales of oscillation for $a^\varepsilon(x)$, see, e.g., [29, 38]. In this case, we can assume, for example, that

$$a^\varepsilon(x) = a\left(x, \frac{x}{\varepsilon_1}, \ldots, \frac{x}{\varepsilon_N}\right).$$

When $\frac{\varepsilon_{i+1}}{\varepsilon_i} \ll 1$, we say that the scales of oscillations are *well separated*. Otherwise, we talk of a *continuum* of scales. An example of coefficients varying at multiple well separated scales is the case of *locally periodic* coefficients. These are functions of the form

$$a^\varepsilon(x) = a\left(x, \frac{x}{\varepsilon}\right), \quad \text{where } a(x, y) \text{ is periodic in the } y\text{-argument.}$$

An example of locally periodic coefficient is depicted in Figure 1.2.

We can take advantage of the scale separation and develop computational methods with reduced computational complexity. As we will see in Section 2.2, some multiscale numerical methods are more suitable for scale-separated problems, e.g. the HMM, while others are designed to solve problems with a continuum of scales, e.g. the Multiscale Finite Element Method. The present thesis will focus on methods to improve the convergence rate of the

resonance error in HMM, so it is mostly suitable for problems with scale separation.

Periodic, almost-periodic and stationary random structures

Multiscale problems with oscillating coefficients can be classified by the micro-structure of the coefficients. Periodic problems have micro-structures that are periodically repeated throughout the domain. Mathematically, it means that $a^\varepsilon(x) = a(x/\varepsilon)$, where $a(y)$ is periodic. Almost-periodic coefficients are defined as elements of the closure, in the L^∞-norm, of the space of functions

$$a(y) = \mathrm{Re}\left(\sum_k a_k e^{i\xi_k \cdot y}\right), \quad a_k \in \mathbb{C},\ \xi_k \in \mathbb{R}^d,$$

which provide a generalization of periodic functions, still retaining some of their most important properties. Finally, another possible structural assumption is that the coefficient is the realization of a stationary random field. Periodic coefficients will be assumed in the analysis of the two corrector equations of Chapters 4 and 5. Almost-periodic coefficients will only be employed in numerical tests. Instead, stationary random coefficients will be considered in Chapters 6 and 7.

1.2 Thesis outline and main contributions

As anticipated above, direct numerical approximations of the multiscale equation eq. (1.1) are not viable because the mesh size has to be chosen smaller than ε, which implies that the number of degrees of freedom blows up, since $\varepsilon \ll 1$. On the other hand, as $\varepsilon \to 0$, the fast-oscillating solution u^ε converges to a smoother, non-oscillating function, which is denoted by u^0. The function u^0 represents the macroscopic component of u^ε and solves the so-called *homogenized* equation

$$\begin{cases} -\nabla \cdot \left(a^0(x)\nabla u^0\right) = f & \text{in } D \subset \mathbb{R}^d, \\ u^0 = 0 & \text{on } \partial D, \end{cases} \tag{1.2}$$

whose coefficients a^0_{ij} are not known *a priori*. The homogenized equation is independent of ε and can be discretized with a freely chosen mesh size. As a first step, before the discretization of eq. (1.2), we must approximate the value of a^0. The homogenized coefficients are found by solving local *corrector* equations that allow to couple the microscale structure to the macroscopic behaviour of the original problem. Except for the case of periodic coefficients where the exact period is known, a resonance error appears in the numerical computation of the homogenized coefficients. As a consequence, the homogenized coefficients are only approximated up to an error, which is known as *resonance* error. The resonance error eventually affects the accuracy of the numerical approximation of u^0 and can lead to very inaccurate simulations.

The main contribution of this thesis to the question of high accuracy approximation of the homogenized coefficients consists in the development and the analysis of two novel corrector

problems to approximate a^0 with high order convergence rates of the resonance error, thus allowing to couple accurately the micro- to the macroscale. In particular, we focus on analysing the corrector problems at the microscale, disregarding the macroscopic behaviour. For this reason, we have only considered coefficients varying at a single scale, $a^\varepsilon(x) = a(x/\varepsilon)$, ignoring other possible coefficient structures, such as locally periodic coefficients. However, the reader should bear in mind that the proposed novel approaches can straightforwardly be applied to those cases: it suffices to apply them to upscale the coefficients $b^\varepsilon_{\bar{x}}(x) := a(\bar{x}, x/\varepsilon)$, at each given point $\bar{x} \in D$.

A rigorous mathematical derivation of the homogenized problem is given in Chapter 2. This theoretical approach does not provide a directly applicable method to compute the homogenized coefficients, except in the simple periodic case. Numerical homogenization methods aim to solve a given corrector problem to approximate a^0, but they all suffer from the so-called resonance error, described in Section 2.4, which often dominates the discretization errors. Several corrector problems have been studied in the past to improve the rate of decay of the resonance error, but they either do not reach arbitrary rates or require to solve very costly models (see Section 2.4.1).

In Chapter 3, a link between a parabolic corrector problem and the standard, elliptic corrector problem is described. No structural assumptions on the coefficients $a^\varepsilon(x)$ is needed at this point. Thanks to the link between parabolic and elliptic partial differential equations, we derive two novel corrector problems: a parabolic and a modified elliptic ones, which can be used to estimate a^0. In Chapter 4, the convergence rate of the parabolic approach is proved to be of arbitrarily high order. For the proof we assume the periodicity of a^ε and sufficient regularity, but numerical tests show that the results hold true also for less regular or stochastic coefficients. A similar analysis, for a novel modified elliptic problem is given in Chapter 5. The *a priori* error bounds of this chapter are established under the same assumptions as in Chapter 4. Numerical tests for non-periodic, non-smooth cases are performed to demonstrate the robustness of the approach. For both cases, we analyse the computational cost-accuracy ratio and compare it to the one of the standard method. The parabolic and the modified elliptic approach are new in the research field of numerical homogenization and their main advantage is that they achieve an exponential decay of the resonance error.

The homogenization of stationary random coefficients has been studied since the early '80s, but the derivation of upper bounds with *explicit* rates of convergence for the homogenization error $\|u^\varepsilon - u^0\|$ remains mainly unsolved. In Chapter 6 we describe the framework of homogenization for random coefficients and provide some of the most recent results in quantitative stochastic homogenization. Moreover, a review of some of the numerical approaches proposed to tackle stochastic homogenization problems is outlined.

The last contribution of the thesis is developed in Chapter 7 and it regards the study and the analysis of the modified elliptic corrector equation for random coefficients. The resonance error is decomposed in several terms: the boundary, the systematic and the statistical error. *A*

priori bounds on the systematic error are proved by exploiting the time decay of the solution of parabolic equations, without any regularity assumption on the coefficients. Numerical tests are provided both to support the theoretical results on the systematic error and to verify experimentally the convergence of the boundary and statistical errors.

2 Homogenization theory, multiscale methods and the resonance error

As we anticipated in the previous chapter, multiscale problems have to be addressed by properly designed multiscale numerical methods. Although they may be based on several different design principles, the numerical analysis of such methods cannot leave homogenization theory out of consideration, as it provides the theoretical foundations for the analysis of differential problems with fast oscillating coefficients. *A priori* analysis of multiscale methods shows that they all suffer from the so-called *resonance* error, due to the inexact coupling between the micro and the macro scale. Reducing the impact of such an error is thus crucial to improve the accuracy of multiscale algorithms.

Outline

In Section 2.1, the main results on the homogenization theory for second order elliptic operators with periodic coefficients are discussed. A non-exhaustive overview of existing multiscale numerical methods is given in Section 2.2. One of these methods, the Finite Element Heterogeneous Multiscale Method (FE-HMM), is described in more details in Section 2.3. Thanks to the results of homogenization theory, it is possible to understand the origin of the resonance error and to derive an *a priori* bound in the FE-HMM context, see Section 2.3.2. In the last decades, many works addressed the question of mitigating the effect of the resonance error and we review some of them in Section 2.4.

2.1 Main results in homogenization theory

Homogenization theory concerns the study of the solutions of PDEs in the regime for $\varepsilon \to 0$, where ε is the length scale of the oscillations of the coefficients [31, 45, 101]. We begin by considering a sequence of ε-indexed second order elliptic equations on a bounded domain $D \subset \mathbb{R}^d$, $d \leq 3$:

$$\begin{cases} -\nabla \cdot (a^\varepsilon(x) \nabla u^\varepsilon) = f & \text{in } D, \\ u^\varepsilon = 0 & \text{on } \partial D, \end{cases} \tag{2.1}$$

with $f \in H^{-1}(D)$. The differential problem (2.1) is well-posed, provided that the tensor $a^\varepsilon \in (L^\infty(D))^{d\times d}$ is symmetric, uniformly elliptic and bounded, i.e. there exist $\alpha, \beta > 0$ such that for any $\varepsilon > 0$,

$$\begin{aligned} &a^\varepsilon(x) = \left[a^\varepsilon(x)\right]^T \text{ and} \\ &\alpha|\xi|^2 \le \left|\xi \cdot a^\varepsilon(x)\xi\right|, \quad \left|a^\varepsilon(x)\xi\right| \le \beta|\xi|, \quad \forall \xi \in \mathbb{R}^d, \text{ a.e. } x \in \mathbb{R}^d. \end{aligned} \tag{2.2}$$

Coefficients satisfying the conditions (2.2) belong to the class $\mathcal{M}(\alpha, \beta, D)$. When $D = \mathbb{R}^d$ we will use the shorthand $\mathcal{M}(\alpha, \beta)$. For the moment, no further structural assumptions are taken on $a^\varepsilon(x)$, even though we one can think of ε as the oscillation length of the coefficients. Structural assumption on the coefficients will be given later on.

The convergence of the sequence $\left\{u^\varepsilon\right\}_\varepsilon \subset H_0^1(D)$ solving (2.1) in the limit for $\varepsilon \to 0$ can be studied through the concept of G-convergence, introduced by [129]. The notion of G-convergence was subsequently extended by [120] to H-convergence for the case of non-symmetric matrices.

Definition 2.1. *The sequence* $\left\{a^\varepsilon\right\}_\varepsilon \subset \mathcal{M}(\alpha, \beta, D)$ *is* G-convergent *to the matrix* $a^0 \in \mathcal{M}(\alpha, \beta, D)$ *if and only if, for every function* $f \in H^{-1}(D)$*, the function* $u^\varepsilon \in H_0^1(D)$ *that solves eq.* (2.1) *converges*

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

where u^0 *is the unique solution of the* homogenised *problem*

$$\begin{cases} -\nabla \cdot \left(a^0(x)\nabla u^0\right) = f & \text{in } D, \\ u^0 = 0 & \text{on } \partial D. \end{cases} \tag{2.3}$$

We will use the following notation to express the G-convergence to a^0 of a sequence $\left\{a^\varepsilon\right\}_\varepsilon$:

$$a^\varepsilon \xrightarrow{G} a^0.$$

In Theorem 2.2 some of the properties satisfied by G-converging sequences are listed.

Theorem 2.2 ([45])**.** *The following properties hold true:*

i) (uniqueness) The G-limit of a G-converging sequence $\left\{a^\varepsilon\right\}_\varepsilon \subset \mathcal{M}(\alpha, \beta, D)$ *is unique.*

ii) (L^2 *convergence) If* $\left\{a^\varepsilon\right\}_\varepsilon \subset \mathcal{M}(\alpha, \beta, D)$ *and* $a^\varepsilon \to a^0$ *strongly in* $\left(L^2(D)\right)^{d\times d}$*, then* $a^\varepsilon \xrightarrow{G} a^0$.

iii) (compactness) Let $\left\{a^\varepsilon\right\}_\varepsilon \subset \mathcal{M}(\alpha, \beta, D)$*. Then, there exist a G-converging subsequence* $\left\{a^{\varepsilon'}\right\}_{\varepsilon'}$.

iv) A sequence $\left\{a^\varepsilon\right\}_\varepsilon \subset \mathcal{M}(\alpha, \beta, D)$ *G-converges if and only if all its G-converging subsequences have the same limit.*

v) Let $\{a^\varepsilon\}_\varepsilon \subset \mathscr{M}(\alpha,\beta,D)$ *be a G-converging sequence to* $a^0 \in \mathscr{M}(\alpha,\beta,D)$*. Then*

$$a^\varepsilon \nabla u^\varepsilon \rightharpoonup a^0 \nabla u^0 \text{ in } \left(L^2(D)\right)^d. \tag{2.4}$$

Remark 2.3. *It is quite interesting to interpret the G-convergence in terms of convergence of inverse of elliptic operators. Let*

$$A^\varepsilon, A^0 : H_0^1(D) \mapsto H^{-1}(D)$$

be the elliptic operators defined by:

$$A^\varepsilon v = -\nabla\cdot\left(a^\varepsilon(x)\nabla v\right), \quad \text{and} \quad A^0 v = -\nabla\cdot\left(a^0(x)\nabla v\right),$$

for any $v \in H_0^1(D)$*. Then, the G-convergence of the sequence* $\{a^\varepsilon\}_\varepsilon$ *is equivalent to the convergence of* $(A^\varepsilon)^{-1} f$ *in the* $H_0^1(D)$*-weak topology, for any* $f \in H^{-1}(D)$*:*

$$\left(A^\varepsilon\right)^{-1} f \rightharpoonup \left(A^0\right)^{-1} f \quad \text{weakly in } H_0^1(D).$$

The compactness results of (2.2) guarantees the existence of a homogenized equation for some subsequence $\left\{a^{\varepsilon'}\right\}_{\varepsilon'}$. Without further assumptions, the G-limit may fail to be unique for different subsequences and the homogenized equation is not uniquely defined. Even when the full sequence $\{a^\varepsilon\}_\varepsilon$ G-converges, a closed form for the homogenized tensor is not available, in general.

Taking further structural assumptions on the tensor $a^\varepsilon(\cdot)$ allows to prove the uniqueness of the homogenized equation and provides a explicit form for a^0. The simplest non-trivial example is when $a^\varepsilon(x) = a(x/\varepsilon)$ is Y-periodic over some parallelepiped $Y \subset \mathbb{R}^d$. Under these assumptions, the homogenized limit is a constant matrix. The same results hold true when $a(\cdot)$ is quasi-periodic or a stationary ergodic random tensor field. These are simplification of coefficients of practical interest. For example one may be interested in the asymptotic behaviour of locally periodic coefficients with multiple scales:$a^\varepsilon(x) = a\left(x, \frac{x}{\varepsilon_1(\varepsilon)}, \ldots, \frac{x}{\varepsilon_N(\varepsilon)}\right)$, with $\lim_{\varepsilon\to 0}\varepsilon_i = 0$ and $\lim_{\varepsilon\to 0}\frac{\varepsilon_{i+1}}{\varepsilon_i} = 0$

In this thesis, we make the assumption of single scale oscillations of the original tensor: There exist $a \in \mathscr{M}\left(\alpha,\beta\right)$ such that

$$a^\varepsilon(x) = a\left(\frac{x}{\varepsilon}\right). \tag{2.5}$$

Two classes of tensors are considered: periodic and stationary ergodic. The results derived for periodic tensors are described in Chapters 4 and 5, while those for stationary ergodic random fields are reported in Chapters 6 and 7.

2.1.1 Periodic homogenization

Here, we briefly recall the convergence result for periodic homogenization. Let us assume that the multiscale tensor can be written as

$$a^\varepsilon(x) = a(x/\varepsilon),$$

where $a \in \mathscr{M}(\alpha,\beta)$ is Y-periodic over the parallelepiped $Y = \prod_{i_1}^{d}[0,l_i)$. Then, the *corrector problem*:

$$\begin{cases} -\nabla\cdot\left(a(y)\cdot\left(\nabla\chi^i+\mathbf{e}_i\right)\right)=0 & \text{on } Y,\\ \chi^i \text{is } Y\text{-periodic} & \int_Y \chi^i = 0 \end{cases} \tag{2.6}$$

has a unique solution $\chi^i \in W^1_{per}(Y)$ and it satisfies

$$\left\|\chi^i\right\|_{W^1_{per}(Y)} \leq \frac{1}{\alpha}\left\|a\mathbf{e}_i\right\|_{L^2(Y)} \tag{2.7}$$

The homogenized coefficient is computed as:

$$a^0_{ij} := \fint_Y \mathbf{e}_i\cdot a(y)\left(\nabla\chi^j+\mathbf{e}_j\right)dy = \fint_Y \left(\nabla\chi^i+\mathbf{e}_i\right)\cdot a(y)\left(\nabla\chi^j+\mathbf{e}_j\right)dy, \tag{2.8}$$

where the second inequality follows from the weak formulation of (2.6). By definition, the homogenized tensor is constant in space. Then, it is possible to prove that

$$u^\varepsilon \rightharpoonup u^0, \text{ weakly in } H^1_0(D), \quad \text{and} \quad a^\varepsilon(x)\nabla u^\varepsilon \rightharpoonup a^0\nabla u^0, \text{ weakly in } \left(L^2(D)\right)^d,$$

where u^0 is the weak solution of the homogenized equation (2.3), see [31, 45, 101].

2.1.2 Stochastic homogenization

The homogenization of stochastic fields concerns the study of (2.1) when a^ε is the realization of a random field. This research question was first addressed by [123, 104]. Standard assumptions in this context are the statistical stationarity and ergodicity of the random field a^ε. Besides providing a setting for proving the existence of the corrector function, stationarity implies that the homogenized coefficients are constant. The assumption of ergodicity, instead, implies the homogenized coefficients to be deterministic. In this setting, the corrector problem is similar the one of the periodic case eq. (2.6), but it is posed on the whole space:

$$-\nabla\cdot\left(a(y,\omega)\cdot\left(\nabla\chi^i+\mathbf{e}_i\right)\right)=0 \quad \text{on } \mathbb{R}^d, \tag{2.9}$$

where we included the variable $\omega \in \Omega$, the probability space. It is possible to prove that the corrector functions χ^i are uniquely defined up to an additive constant and that $\nabla\chi^i$ is a

stationary random field. The homogenized coefficients are computed as

$$a^0_{ij} := \mathbb{E}\left[\mathbf{e}_i \cdot a(\cdot)\left(\nabla\chi^j + \mathbf{e}_j\right)\right] = \fint_{\mathbb{R}^d} \mathbf{e}_i \cdot a(y,\omega)\left(\nabla\chi^j + \mathbf{e}_j\right) dy, \tag{2.10}$$

where the second equation follows from assumption of ergodicity. If the ergodicity condition is dropped, the homogenized coefficients are measurable with respect to the sub-σ-algebra of invariant sets:

$$a^0_{ij} = \mathbb{E}\left[\mathbf{e}_i \cdot a(\cdot)\left(\nabla\chi^j + \mathbf{e}_j\right)\middle|\mathscr{F}_{inv}\right].$$

A more complete description of stochastic homogenization is postponed to Chapter 6.

2.2 Overview of multiscale numerical methods

In this section we give a general introduction to multiscale numerical methods and briefly describe some of them. A more detailed description of the Finite Elements Heterogeneous Multiscale Method (FE-HMM) is given in Section 2.3 with the aim of introducing the main results of FE-HMM in Section 2.3.1 and a description of the resonance error in Section 2.3.2.

In Chapter 1, we anticipated that the aim of multiscale numerical methods is to solve multiscale problem with a number of degrees of freedom that is not constrained by the size of the fine scale heterogeneities. Multiscale problem can be written in general mathematical terms as

$$\mathscr{L}^\varepsilon u^\varepsilon = f, \tag{2.11}$$

where $\mathscr{L}^\varepsilon$ is the multiscale operator, u^ε the solution and f represents the forcing term. The above representation can be thought as a generalization of the model problem (2.1). The superscript ε is a small parameter denoting the size of small scale heterogeneities.

Standard numerical methods may not be applicable to solve this kind of problems. For example, the Finite Element Method (FEM) is not robust when used to discretize problems with rough coefficients, [28], as the numerical error between the true and the numerical solutions does converge to 0 only if the meshsize is sufficiently small, i.e. if $H \leq \varepsilon$. For small values of ε, the condition $H \leq \varepsilon$ can only be satisfied with a huge number of degrees of freedom, that eventually blows up as $\varepsilon \to 0$, thus making the computation unfeasible. On the other hand, most of the degrees of freedom are used to resolve the small scale features of the problem, while we are mostly interested in understanding the macroscopic behaviour, which would require far less degrees of freedom for problems with homogeneous coefficients.

Multiscale methods aim to propose a strategy to find a solution to eq. (2.11) such that the convergence results still hold true, without the constraint on the meshsize. So, the goal is to find a discrete operator $\mathscr{L}_H$ which is able to capture the large scale behaviour, i.e. such that the solution u_H of

$$\mathscr{L}_H u_H = f_H,$$

converges to the leading order approximation of the multiscale solution u^ε.

The operator $\mathscr{L}^\varepsilon$ can be any operator with heterogeneous coefficients. In order to fix ideas, we consider the second order elliptic operator

$$\mathscr{L}^\varepsilon v = -\nabla\left(a^\varepsilon(x)\nabla v\right).$$

In this case, the multiscale method will aim to approximate the homogenized operator $\mathscr{L}^0$ defined by

$$\mathscr{L}^0 v = -\nabla\left(a^0(x)\nabla v\right),$$

where a^0 is the H-limit of the sequence $\left\{a^\varepsilon\right\}_\varepsilon$.

A common ingredient of multiscale numerical methods is the need for solving a set of local microscale problems in order to extract the microscopic information and pass it to the macroscale model. As we will see, this implies using artificial conditions at the boundaries of the microscopic domains. These boundary conditions eventually worsen the overall accuracy of multiscale methods and improved methodologies with reduced boundary errors are needed in computations.

Generalised Finite Elements Method (GFEM)

The Generalised Finite Element Method (GFEM) addresses multiscale problems by assembling the stiffness matrix in a way that takes into account the small scale information. The method was first proposed in [28] for one dimensional problems, and further elaborated in [23, 24, 25, 26, 27, 130]. The strategy of the method is to partition the computational domain D into sub-domains D_j's and to introduce local approximation spaces V_j in which local solutions are looked for. The global solution is then sought in the global space S, obtained by "pasting together" the local spaces V_j's by means of a partition of unity $\left\{\phi_j\right\}_j$ (in fact, this method is referred to also as *partition of unity method*) over the sub-domains D_j. This approach offers a generalisation of FEM (hence the name) and it is possible to choose the sub-domains and the local spaces in a way to recover the standard global Finite Elements (FE) space.

The method was originally proposed to solve problems with perforated materials or crack dynamics, but it was also successfully applied to other multiscale problems.

The Variational Multiscale Method (VMM)

The Variational Multiscale Method (VMM) was proposed in [99, 100] and it is based on a decomposition of the solution into two terms, $u = u_H + u'$, the former can be treated numerically while the latter accounts for all the sub-grid effects and must be modelled. In the same spirit, the trial space V is decomposed into a finite dimensional space V_H and a residual space V' such that

$$V = V_H \oplus V'.$$

One can seek the solution by solving the variational problems: Find $u_H \in V_H$ and $u' \in V'$ such that

$$B\left(u_H + u', v_H\right) = \left(f, v_H\right), \quad \forall v_H \in V_H, \tag{2.12}$$

$$B\left(u_H + u', v'\right) = \left(f, v'\right), \quad \forall v' \in V', \tag{2.13}$$

where the bilinear from $B : V \times V$ is defined as

$$B\left(u, v\right) := \int_D \nabla v \cdot a^\varepsilon(x) \nabla u \, dx.$$

Re-writing eq. (2.13) as

$$B(u', v') = (f, v') - B(u_H, v') = (f - \mathscr{L}^\varepsilon u_H, v'),$$

one can write formally $u' = M(f - \mathscr{L}^\varepsilon u_H)$, where M is a bounded linear operator on V' obtained by restricting $f - \mathscr{L}^\varepsilon u_H$ to V', to obtain a variational problem in V_H:

$$B(u_H, v_H) + B(M(f - \mathscr{L}^\varepsilon u_H), v_H) = (f, v_H) \quad \forall v_H \in V_H.$$

For an actual numerical solution, the operator M has to be approximated and localized. The VMM is equivalent to the residual-free bubble method, as proved in [37].

The Equation–Free Method (EFM)

The equation–free method (EFM) was proposed in [103] as a mean to solve multiscale evolution problems with scale separation and unknown macroscopic model. This method was conceived in the context of computational chemistry, in which models describing the evolution of the system for short times are available, but running such models for long time horizons and large domains (which is in the interest of researchers) is not possible due to the large computational cost. EFM by-passes the derivation of an explicit macroscopic model by exploiting the scale separation, common to many multiscale problems. It is based on a *lift-evolve-restrict* procedure, which is used in the HMM as well. In the *lifting* process, a microscale initial condition over small scale *patches* is reconstructed from the macroscale distribution. The evolution at the microscale is very fast compared to the evolution at the macroscale and it is denoted as *patch dynamics*. The microscale system is evolved according to the known microscale equation until a time horizon much smaller than the time step size at the macroscale. As a third step, the evolved microscale solution is used to estimate the time derivative of the macroscopic variables (*restriction*). Finally, the macroscopic solution is interpolated in space and extrapolated in time (by means of the Projective Forward Euler method) to reconstruct it between the patches and advance of a macro time step. The lift-evolve-restrict procedure and the extrapolation method in time build up a coarse time stepper.

The equation-free method has some critical issues, as it is pointed out in [56]. Besides being applicable only to scale-separated systems, it fails to model problems with macroscale stochastic nature. In fact, in this case the equation-free method is either unable to capture the right effect or it is as expensive as the standard discretization scheme.

The Multiscale Finite Elements Method (MsFEM)

The Multiscale Finite Elements Method (MsFEM), described in [58], couples the small scale features with the macroscopic equation by using a finite element basis of multiscale nature, V_H^{MsFEM}, in place of the standard Finite Element space. Each basis element of the multiscale FE space is constructed by solving local boundary value problems posed over each macro-mesh element, which allows to take into account the effect of microscopic oscillations. For a given coarse mesh $\mathcal{T}_H$, and for each basis element Φ_i, $i = 1, \dots, n_H$, of the n_H-dimensional standard FE space V_H, the i-th multiscale FE basis element Φ_i^{MsFEM} is defined as the solution of the variational problem: Find $\Phi_i^{MsFEM} \in V_h$ such that $\Phi_i^{MsFEM}|_{\partial T} = \Phi^i$ and

$$\int_T a(x) \nabla \Phi_i^{MsFEM} \cdot \nabla \varphi_h \, dx = 0, \quad \forall \varphi_h \in V_h,$$

where V_h is a full resolution (high dimensional) finite element space. The values of the standard finite element basis function on the coarse mesh are used as boundary conditions for the local problem, which are discretized by a fine scale mesh of size $h \leq \varepsilon$. Then, the solution of the problem eq. (2.11) is sought into the space V_H^{MsFEM}, defined by linear combinations of the multiscale FE basis functions: Find $u_H \in V_H^{MsFEM}$ such that

$$B(u_H, v_H) = \big(f, v_H\big), \quad \forall v_H \in V_H^{MsFEM}.$$

This method was first proposed in [96] and later analysed in [97, 133]. The MsFEM allows to solve multiscale problem with and without scale separation at a cost that is only weakly dependent on the ratio $|\Omega| / h$, since the local computations are totally decoupled and can be solved in parallel, while the macro-problem has a cost independent of h.

As reported in [96], the main source of error in MsFEM is due to the *resonant effect* whose magnitude scales as ε / h. The resonance error can thus be quite large, especially for problems without scale separation, for which there will always be a value ε matching the mesh size h. This error represents the greatest challenge for MsFEM and numerical upscaling methods. In order to mitigate this error several techniques based on non-conforming formulations have been proposed, such as the use of oversampling [60] or a Petrov-Galerkin formulation of the finite dimensional problem [98].

The Localized Orthogonal Decomposition (LOD)

Localization of elliptic partial differential equations was first proposed in [114], and further developed in [88], as an approach to reduce the resonance error in multiscale problems without scale separation.

The method was first proposed for the solving multiscale linear second order PDEs with symmetric coefficients. The solution is sought in the full resolution (i.e., very high-dimensional) finite element space $V_h \subset H^1(D)$, which is decomposed into the direct sum:

$$V_h = V_H \oplus W_h \,,$$

where V_H is a coarse finite element space (with which the macroscopic behaviour is captured), while W_h is the kernel of a Clément-type quasi-interpolation operator $I_H : V_h \mapsto V_H$. The subspace W_h contains the fine scale features of V_h which cannot be captured by the coarse space V_H. However, the fact that W_h is the kernel of an interpolation operator suggests that the features of the (high dimensional) space W_h could be neglected. Consequently, we can look for a splitting

$$V_h = V_H^{ms} \oplus W_h \,,$$

where $\dim\left(V_H^{ms}\right) = \dim(V_H)$ and accurate approximations (in the H^1-norm) of u^ϵ can be found in V_H^{ms}. Hence, we look for the orthogonal complement of W_h in V_h with respect to the scalar product $B(\cdot,\cdot)$ associated to the elliptic problem. Let P_h be the $B(\cdot,\cdot)$-orthogonal projection on W_h, then one can define

$$V_H^{ms} = (I - P_h)(V_H). \tag{2.14}$$

In relation with the construction of the multiscale finite element space V_H^{MsFEM} we see that in the LOD method we take a substantially different approximation space. Indeed, the space V_H^{ms} can be equally defined as the set of solutions Φ_H^{ms} of

$$B\left(\Phi_H^{ms}, \varphi_h\right) = 0 \quad \forall \varphi_h \in W_h.$$

The support of the functions $\Phi_H^{ms} \in V_H^{ms}$ is not local, i.e. it may cover the whole computational domain D, in contrast to the functions Φ_i^{MsFEM}, which are local by construction. This represents a problem from the computational point of view, as, in order to compute a basis for V_H^{ms} it would be necessary to solve a full scale problem. A localization (whence, the name) technique is then used to mitigate this issue: k-th order patches around vertices x are introduced and denoted by $\omega_{x,k}$. The vector space of localized functions $W_h(\omega_{x,k})$ is defined for any patch and the local problems are solved on each patch to compute the basis elements of $V_{H,k}^{ms}$. Thanks to the exponential decay in space of the difference between the global and the local solutions, the localization error is negligible, provided that the patches are sufficiently large. Consequently, the discretization error can be bounded by the macro-mesh size H independently of the small scale oscillations.

A discussion of the LOD method in comparison to MsFEM can be found in [90] and an extension to the semi-linear case in [89]. In [66] the authors describe in details how to practically implement the LOD method. Like the MsFEM, the LOD method is more suitable for problems without scale separation. The reason is that each local problem is solved over a domain of size H with a mesh size $h \leq \varepsilon$. Therefore, there are $\mathcal{O}\left((H/h)^d\right) > \mathcal{O}\left((H/\varepsilon)^d\right)$ degrees of freedom, per local problem. In the regime for $\varepsilon \to 0$ the number of local degrees of freedom blows up and the computation becomes unfeasible.

2.3 The Finite Elements Heterogeneous Multiscale Method (FE-HMM)

The Heterogeneous Multiscale Method (HMM) is a general framework for designing multiscale algorithms for both ODEs with multiple time scales and PDEs with rough coefficients [54]. In this thesis, we will focus on applying the HMM to the solution of elliptic PDEs for which both the micro- and the macro-models are discretized by the FEM. This case is analysed by the Finite Elements Heterogeneous Multiscale Method (FE-HMM) [2, 14, 54, 55] which was developed for solving general multiscale problems, for which the G-limit of the fast oscillating tensor is not known *a priori*. Extension to the finite difference method [11], elasticity problems [3], parabolic problems [116, 15] and wave propagation [12, 62, 63] exist in literature. The algorithm may be used to solve problems with scale separation where the macroscopic model and the microscopic may be of different nature. In this section we describe the main ingredients of the methods in order to fix the notation and provide the mathematical framework to study the resonance error.

As the equation-free method, HMM is based on a "reconstruction-compression" paradigm (called "lifting-restriction" in the equation-free method) and it uses the microscale model throughout the computation, while other methods such as the VMM use it only to derive the macroscopic model. Contrarily to other approaches such as the sequential coupling, FE-HMM allows to derive *a priori* and *a posteriori* error estimate of the overall discretization, that also depend on the accuracy of the reconstruction of the effective operator $\mathscr{L}^0$.

In this description we will consider as model problem the weak form of the elliptic equation eq. (2.1) which is: Find $u^\varepsilon \in H_0^1(D)$ such that

$$B(u^\varepsilon, v) := \int_D \nabla v \cdot a^\varepsilon(x) \nabla u^\varepsilon \, dx = \langle f, v \rangle_{H^{-1}(D), H_0^1(D)}, \quad \forall v \in H_0^1(D).$$

The FE-HMM is composed of two main ingredients: a macro and a micro finite elements spaces, which are linked through the HMM bilinear form $B_H(\cdot,\cdot)$ of eq. (2.18).

Macro Finite Element Space

Let us consider a macroscopic partition $\mathcal{T}_H$ of the computational domain D and a finite element space $V^p(D,\mathcal{T}_H)$ on $\mathcal{T}_H$:

$$V^p(D,\mathcal{T}_H) = \left\{u^H \in H_0^1(D) : u^H|_E \in \mathscr{R}^p(E), \forall E \in \mathcal{T}_H\right\}, \tag{2.15}$$

where $\mathscr{R}^p(E)$ is the space of polynomials of total (resp. maximum) degree p over simplicial (resp. rectangular) elements of $\mathcal{T}_H$. The size H of the macroscopic partition $\mathcal{T}_H$ can be chosen arbitrarily and it is not constrained by the scale of heterogeneities. For any element $E \in \mathcal{T}_H$, we consider an index set J, a set of point $\left\{x_j\right\}_{j\in J} \subset E$ and weight $\left\{\omega_j\right\}_{j\in J} \subset (0,+\infty)$ constituting a quadrature formula (QF) $\left\{x_j,\omega_j\right\}_{j\in J}$ that exactly integrates polynomials in $\mathscr{R}^\sigma(E)$ and satisfies the following conditions:

- *Coercivity condition.* Using a QF to compute the bilinear form eq. (2.18) does not guarantee that the coercivity condition is satisfied. Thus, we require that

$$\sqrt{\sum_{j\in J} \omega_j \left|\nabla p(x_j)\right|} \quad \text{is a norm on the finite dimensional space } \mathscr{R}^\sigma(E)/\mathscr{R}^0(E). \tag{2.16}$$

  This property holds if the nodes $\left\{x_j\right\}_{j\in J}$ contain an unisolvent set for the derivatives of the considered polynomial set [44].

- *Approximation condition.* Let $u_{h,QF} \in V^p(D,\mathcal{T}_h)$ be the FEM solution of a variational elliptic problem where the integrals are computed by the numerical QF. We will assume that the QF is chosen such that the standard error estimates for FEM hold. Assuming enough regularity of the solution u, this reads

$$\left\|u_{h,QF} - u\right\|_{H^1} \leq Ch^p, \quad \left\|u_{h,QF} - u\right\|_{L^2} \leq Ch^{p+1}. \tag{2.17}$$

  For $p > 1$, the condition eq. (2.17) holds true if the QF is exact up to order $2p-2$ (resp. $2p-1$) for simplicial (resp. rectangular) elements, while for $p = 1$ we require the QF to be exact up to order 1 (resp. 2) for simplicial (resp. rectangular) elements [44].

Next, we consider *sampling domains* centred at each quadrature point

$$K_\delta(x_j) = x_j + [-\delta/2,\delta/2]^d,$$

where d is the dimension and $\delta \geq \varepsilon$. The macroscopic bilinear form is defined:

$$B_H\left(v^H, w^H\right) := \sum_{E\in\mathcal{T}_H} \sum_j \omega_j \fint_{K_\delta(x_j)} \nabla w_j^h \cdot a^\varepsilon(x) \nabla v_j^h \, dx, \tag{2.18}$$

where the micro function w_j^h, v_j^h are defined by the microscale problem eq. (2.23). The FE-HMM solution of problem eq. (2.1) is $u^H \in V^p(D,\mathscr{T}_H)$ such that

$$B_H\left(u^H, v^H\right) = F\left(v^H\right), \quad \forall v^H \in V^p(D,\mathscr{T}_H). \tag{2.19}$$

The bilinear form $B_H(\cdot,\cdot)$ is bounded and coercive on $V^p(D,\mathscr{T}_H)$ and the functional $F(\cdot)$ is linear and bounded. Hence, the assumption of the Lax-Milgram theorem as satisfied and the solution u^H is unique and bounded.

Micro Finite Element Space

Let us define, on each sampling domain, a microscopic partition $\mathscr{T}_h$ and a microscopic finite element space:

$$S^q(K_\delta(x_j),\mathscr{T}_h) = \left\{u^h \in W(K_\delta(x_j)) : u^h|_e \in \mathscr{R}^q(e), \forall e \in \mathscr{T}_h\right\}, \tag{2.20}$$

where $W(K_\delta(x_j))$ is a Sobolev space whose choice sets the boundary conditions for the corrector problems and thus determines the type of coupling between the micro and the macro problems. We consider the cases:

$$W(K_\delta(x_j)) = W^1_{per}\left(K_\delta(x_j)\right), \text{ for periodic boundary conditions,} \tag{2.21}$$

$$W(K_\delta(x_j)) = H^1_0\left(K_\delta(x_j)\right), \text{ for Dirichlet boundary conditions.} \tag{2.22}$$

Other choices for the boundary conditions are possible, but we will focus on these two. The micro functions are set to satisfy the microscale problem: Find v_j^h such that $v_j^h - v^H_{lin,j} \in S^q(K_\delta(x_j),\mathscr{T}_h)$ and

$$\int_{K_\delta(x_j)} a^\varepsilon(x)\nabla v_j^h \cdot \nabla z^h \, dx = 0, \quad \forall z^h \in S^q(K_\delta(x_j),\mathscr{T}_h), \tag{2.23}$$

where $v^H_{lin,j}$ is the linear approximation of v^H in a neighbourhood of x_j:

$$v^H_{lin,j} = v^H(x_j) + (x - x_j)\cdot\nabla v^H(x_j).$$

A schematic representation of the method is depicted in Figure 2.1, where the red square dots represent the quadrature points in the macro mesh $\mathscr{T}_H$. Around each quadrature point x_j a sampling domain $K_\delta(x_j)$ is defined and discretized by the micro mesh $\mathscr{T}_h$.

2.3.1 A few results on the *a priori* error analysis of FE-HMM

Here we recall, without proofs, a few results of the *a priori* error analysis for FE-HMM. The numerical error is composed of three terms, respectively accounting for the discretization error at the macroscale (e_{MAC}), the coupling conditions between the two scales (e_{MOD}) and

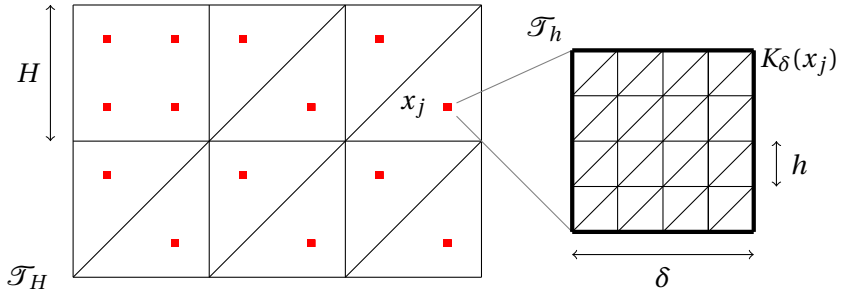


Figure 2.1 – Schematic representation of FE-HMM.

the discretization error at the microscale (e_{MIC}):

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

where $\|\cdot\|$ denotes the H^1 and L^2 norms. The second error, e_{MOD} is the resonance error and it is connected to the reconstruction of the homogenized coefficients[1]. In order to prove the error decomposition above, it is sufficient to prove the coercivity of the modified bilinear form eq. (2.18). Explicit estimates for e_{MOD} and e_{MIC} are only possible under further assumptions on the coefficients a^ε, i.e. local periodicity with macroscopic collocation, [4]. The focus of this thesis is on the local reconstruction of the homogenized coefficients, so we decided to ignore the macroscopic variations of the coefficients and only consider the periodic case:

$$a^\varepsilon(x) = a(x/\varepsilon),\ a \in \mathscr{M}(\alpha,\beta) \text{ is } K\text{-periodic}\,, \tag{2.24}$$

where $K := [-1/2, 1/2]^d$: However, the upscaling strategies proposed in this thesis can be exploited in the homogenization of more general coefficients, not only periodic.

We define two "intermediate" solutions $u^{0,H}$ and $\bar{u}^H$ to split the global error into the three error terms above. The two intermediate solutions are, respectively, the numerical solution of the *exact* homogenized problem eq. (2.3) and the approximate solution obtained by solving the corrector problem exactly in $W(K_\delta(x_j))$. Let us define the bilinear form on $V^p(D,\mathcal{T}_H) \times V^p(D,\mathcal{T}_H)$ for the homogenised problem eq. (2.3):

$$B_{0,H}\left(v^H, w^H\right) := \sum_{E\in\mathcal{T}_H} \sum_{j\in J} \omega_j\, a^0 \nabla v^H(x_j) \cdot \nabla w^H(x_j). \tag{2.25}$$

Ellipticity and boundedness of the G-limit a^0 and the coercivity condition of the QF imply coercivity and boundedness of the bilinear form eq. (2.25). Here, any link with a microscale problem is absent, as we are treating the numerical approximation of the eq. (2.3). The homogenised tensor is constant, thanks to the assumption in eq. (2.24). Under the sufficient regularity conditions and accuracy of the quadrature formula we have the classical FEM

[1] The resonance error is denoted as e_{MOD} in agreement to the notation used in several papers on FE-HMM, e.g. [4], where this term is called *modelling* error.

convergence result for the homogenized problem:

Proposition 2.4 ([4]). *Suppose that $u^0 \in H^{p+1}(D)$ and let $u^{0,H} \in V^p(D,\mathcal{T}_H)$ be the solution of the variational problem: Find $u^{0,H} \in V^p(D,\mathcal{T}_H)$ such that*

$$B_{0,H}\left(u^{0,H}, v^H\right) = F(v^H), \quad \forall v^H \in V^p(D,\mathcal{T}_H). \tag{2.26}$$

Then,

$$e_{MAC,H^1} := \left\| u^0 - u^{0,H} \right\|_{H^1(D)} \leq CH^p, \quad e_{MAC,L^2} := \left\| u^0 - u^{0,H} \right\|_{L^2(D)} \leq CH^{p+1}.$$

The microscopic error accounts for the propagation, at the macroscale, of the corrector problems' discretization error, [2]. Let us consider the bilinear form on $V^p(D,\mathcal{T}_H) \times V^p(D,\mathcal{T}_H)$ with exact micro solutions:

$$\bar{B}_H\left(v^H, w^H\right) := \sum_{E\in\mathcal{T}_H} \sum_{j\in J} \omega_j \fint_{K_\delta(x_j)} a^\varepsilon(x) \nabla v_j \cdot \nabla w_j \, dx, \tag{2.27}$$

where v_j, w_j are the *exact* solution of eq. (2.23) in the infinite dimensional space $W(K_\delta(x_j))$. Coercivity and boundedness of $\bar{B}_H(\cdot,\cdot)$ can be proved as for $B_H(\cdot,\cdot)$, [4]. Hence, the variational problem: Find $\bar{u}^H \in V^p(D,\mathcal{T}_H)$ such that

$$\bar{B}_H\left(\bar{u}^H, v^H\right) = F(v^H), \quad \forall v^H \in V^p(D,\mathcal{T}_H), \tag{2.28}$$

has a unique solution $\bar{u}^H$, which is the discrete macro solution with exactly solved micro scale problem. The micro-error measures the discrepancy between u^H and $\bar{u}^H$, which uniquely depends on the discretization error at the microscale.

Proposition 2.5 ([4]). *Let $u^H, \bar{u}^H \in V^p(D,\mathcal{T}_H)$ be the solutions, respectively, of eqs.* (2.19) *and* (2.28) *and based on corrector problems that satisfy the same boundary condition. Additionally, suppose that the corrector functions defined in eq.* (2.32) *satisfy*

$$\varphi^i_{\delta,j} \in W^{q+1,\infty}(K_\delta(x_j)), \quad \textit{and} \quad \left\| D^\alpha \varphi^i_{\delta,j} \right\|_{L^\infty(K_\delta(x_j))} \leq C\varepsilon^{-|\alpha|}.$$

Then

$$e_{MIC,H^1} := \left\| \bar{u}^H - u^H \right\|_{H^1(D)} \leq C\left(\frac{h}{\varepsilon}\right)^{2q}. \tag{2.29}$$

The last error to be estimated is the resonance error, which is defined and bounded in the following Proposition 2.6. A proof of this result is given in Section 2.3.2 for the periodic case, and an extension to the locally periodic case is given in Proposition 2.9.

Proposition 2.6 ([4]). *Let $a \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^{d\times d})$ be K-periodic. Let $u^{0,H}$ and $\bar{u}^H$ be the solutions*

of eqs. (2.26) *and* (2.28). *Then*

$$e_{MOD} := \left\| u^{0,H} - \bar{u}^H \right\|_{H^1(D)} \leq \begin{cases} 0 & \text{if } \delta/\varepsilon \in \mathbb{N} \text{ and } W(K_\delta(x_j)) = W^1_{per}\left(K_\delta(x_j)\right), \\ C\frac{\varepsilon}{\delta} & \text{if } \delta/\varepsilon \notin \mathbb{N} \text{ and } W(K_\delta(x_j)) = H^1_0(K_\delta(x_j)). \end{cases} \tag{2.30}$$

2.3.2 Proof of the resonance error bound for periodic tensors

This section is devoted to the proof of the resonance error bound for e_{MOD}. Such an error arises when the boundary conditions of the micro-problems do not match the values of the exact solution on the boundary of the sampling domains. The error $\left\| \bar{a}^0(x_j) - a^0 \right\|_F$ can only be estimated in the cases where an exact form of a^0 is available, for example in the periodic case which is analysed in this section. Under the periodicity assumption, the resonance error arises if the boundary conditions of the corrector problems do not fit the periodic settings, e.g. when Dirichlet boundary conditions are used in problem eq. (2.32). By contrast, if periodic boundary conditions are considered in eq. (2.32), and δ is an integer multiple of ε, the homogenized tensor is reconstructed exactly. In the ε-periodic case, we denote the corrector functions as $\chi^i_{\varepsilon,j}(x)$, which solve the problem

$$\int_{K_{n\varepsilon}(x_j)} a^\varepsilon(x)\left(\nabla\chi^i_{\varepsilon,j} + \mathbf{e}_i\right)\cdot\nabla z\,dx = 0, \quad \forall z \in W_{per}\left(K_\varepsilon(x_j)\right). \tag{2.31}$$

We will compare the periodic solution with $\varphi_{\delta,j} \in H^1_0\left(K_\delta(x_j)\right)$ obtained by using homogeneous Dirichlet boundary conditions.

Before estimating the resonance error, we show that the bilinear form $\bar{B}_H$ can be re-interpreted as a bilinear form at the macroscale (so, it ignores the microscale effects), with modified coefficients. By linearity, the microscale exact solution v_j used in eq. (2.27) can be written as

$$v_j = v^H_{j,lin} + \sum_{i=1}^d \frac{\partial v^H_{j,lin}}{\partial x_i}\varphi^i_{\delta,j},$$

where the *local corrector* $\varphi^i_{\delta,j} \in W\left(K_\delta\left(x_j\right)\right)$ solves

$$\int_{K_\delta(x_j)} a^\varepsilon(x)\left(\nabla\varphi^i_{\delta,j} + \mathbf{e}_i\right)\cdot\nabla z\,dx = 0, \quad \forall z \in W\left(K_\delta\left(x_j\right)\right). \tag{2.32}$$

If $W\left(K_\delta\left(x_j\right)\right) = W^1_{per}\left(K_\delta\left(x_j\right)\right)$ and $\delta = n\varepsilon$, with $n \in \mathbb{N}$, then $\varphi^i_{\delta,j}(x) = \chi^i_{\varepsilon,j}(x)$. Thus we have:

$$\begin{aligned}\bar{B}_H\left(v^H, w^H\right) &:= \sum_{E\in\mathcal{T}_H}\sum_{j\in J}\omega_j \fint_{K_\delta(x_j)} a^\varepsilon(x)\nabla v_j\cdot\nabla w_j\,dx \\ &= \sum_{E\in\mathcal{T}_H}\sum_{j\in J}\omega_j \fint_{K_\delta(x_j)} a^\varepsilon(x)\nabla\left(v^H_{j,lin} + \sum_{i=1}^d \frac{\partial v^H_{j,lin}}{\partial x_i}\varphi^i_{\delta,j}\right)\cdot\nabla\left(w^H_{j,lin} + \sum_{l=1}^d \frac{\partial w^H_{j,lin}}{\partial x_l}\varphi^l_{\delta,j}\right)dx\end{aligned}$$

$$= \sum_{E\in\mathcal{T}_H}\sum_{j\in J}\omega_j \sum_{i,l=1}^{d} \frac{\partial v^H_{j,lin}}{\partial x_i}\frac{\partial w^H_{j,lin}}{\partial x_l} \fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\mathbf{e}_i+\nabla\varphi^i_{\delta,j}\right)\cdot\left(\mathbf{e}_l+\nabla\varphi^l_{\delta,j}\right)dx$$

By the identity $\nabla v^H_{j,lin} = \nabla v^H(x_j)$ and by defining the approximate homogenised matrix at x_j, $\bar{a}^0(x_j)$, with components

$$\bar{a}^0_{li}(x_j) := \fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\mathbf{e}_i+\nabla\varphi^i_{\delta,j}\right)\cdot\left(\mathbf{e}_l+\nabla\varphi^l_{\delta,j}\right)dx, \tag{2.33}$$

and the bilinear form $\bar{B}_H$ can be rewritten as

$$\bar{B}_H\left(v^H,w^H\right) = \sum_{E\in\mathcal{T}_H}\sum_{j\in J}\omega_j\bar{a}^0(x_j)\nabla v^H(x_j)\cdot\nabla w^H(x_j). \tag{2.34}$$

The difference between $\bar{a}^0$ and the true G-limit a^0 provides a bound for the resonance error through the following Proposition 2.7.

Proposition 2.7 ([5])**.** *Let $u^{0,H}$ and $\bar{u}^H$ be, respectively, the solutions of eqs.* (2.26) *and* (2.28)*. Then*

$$\left\|u^{0,H}-\bar{u}^H\right\|_{H^1(D)} \le \sup_{\substack{E\in\mathcal{T}_H\\ x_j\in E}}\left\|\bar{a}^0(x_j)-a^0\right\|_F \tag{2.35}$$

where $\|\cdot\|_F$ is the Frobenius norm.

Proof. Let $w^H := u^{0,H}-\bar{u}^H$. From the coercivity of $\bar{B}_H$ it follows that there exist $c>0$ independent of H and ε such that

$$\begin{aligned} c\left\|u^{0,H}-\bar{u}^H\right\|^2_{H^1(D)} &\le \bar{B}_H\left(u^{0,H}-\bar{u}^H,w^H\right)\\ &= \bar{B}_H\left(u^{0,H},w^H\right)-F\left(w^H\right)\\ &= \bar{B}_H\left(u^{0,H},w^H\right)-\bar{B}_{0,H}\left(u^{0,H},w^H\right). \end{aligned}$$

Hence, we conclude that there exists $C>0$, independent of H, such that

$$\left\|u^{0,H}-\bar{u}^H\right\|_{H^1(D)} \le C \sup_{w^H\in V^p(D,\mathcal{T}_H)} \frac{\left|\bar{B}_H\left(u^{0,H},w^H\right)-\bar{B}_{0,H}\left(u^{0,H},w^H\right)\right|}{\left\|w^H\right\|_{H^1(D)}}. \tag{2.36}$$

We conclude by bounding the right-hand side through

$$\begin{aligned} \left|\bar{B}_H\left(u^{0,H},w^H\right)-\bar{B}_{0,H}\left(u^{0,H},w^H\right)\right| &= \left|\sum_{E\in\mathcal{T}_H}\sum_{j\in J}\omega_j\left(\bar{a}^0(x_j)-a^0\right)\nabla u^{0,H}(x_j)\cdot\nabla w^H(x_j)\right|\\ &\le \sup_{\substack{K\in\mathcal{T}_H\\ x_j\in K}}\left\|\bar{a}^0(x_j)-a^0\right\|_F\left\|u^{0,H}\right\|_{H^1(D)}\left\|w^H\right\|_{H^1(D)}. \end{aligned}$$

In conclusion, eq. (2.35) follows from the uniform bound on $\left\|u^{0,H}\right\|_{H^1(D)} \le C\left\|f\right\|_{H^{-1}}$ given by the variational problem eq. (2.26). □

From Proposition 2.7, we see that the resonance error comes from the inaccurate reconstruction of the homogenized tensor and it is independent of the numerical error at the microscale, as the analysis is done assuming that the corrector problems are solved exactly. The task is now to bound the right-hand side of eq. (2.35). This is done by the following statement, already proved in [55, 5].

Proposition 2.8 ([55, 5]). *Let $a \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^{d\times d})$ be $[0,1)^d$-periodic, let a^0 be the G-limit of $a(x/\varepsilon)$, for $\varepsilon \to 0$ and let $\bar{a}(x_j)$ be given by eq.* (2.33) *for $\varphi_{\delta,j} \in H_0^1\left(K_\delta\left(x_j\right)\right)$ and $\delta > \varepsilon$. There exists a constant $C > 0$ independent of ε, δ such that*

$$\sup_{\substack{E\in\mathcal{T}_H \\ x_j\in E}} \left\| \bar{a}^0(x_j) - a^0 \right\|_F \leq C\frac{\varepsilon}{\delta}. \tag{2.37}$$

Proof. Let us define $n \in \mathbb{N}$ as

$$n = \begin{cases} \left\lfloor \frac{\delta}{\varepsilon} \right\rfloor & \text{if } \delta/\varepsilon \notin \mathbb{N}, \\ \frac{\delta}{\varepsilon} - 1 & \text{if } \delta/\varepsilon \in \mathbb{N}. \end{cases}$$

By following a standard procedure to estimate $\left\| u^\varepsilon - \left(u^0 + \varepsilon \sum_i \frac{\partial u^0}{\partial x_i}\chi^i\right)\right\|_{H^1}$, we define the boundary layer $\Delta := K_\delta \setminus K_{n\varepsilon}$ and a cut-off function $\rho_\varepsilon \in C_0^\infty(K_\delta)$ such that

$$\rho_\varepsilon(x) = 1 \quad \text{in } K_{n\varepsilon}, \quad \text{and} \quad \left\|\nabla\rho_\varepsilon\right\|_{L^\infty(K_\delta)} \leq \frac{C}{\varepsilon},$$

for some $C > 0$ independent of ε. Then, we can define the function $\theta_0^i := \varphi_{\delta,j} - \rho_\varepsilon \chi_{\varepsilon,j}^i$ solving the variational problem

$$\int_{K_\delta(x_j)} a^\varepsilon(x)\nabla\theta_0^i \cdot \nabla z\, dx = \int_{K_\delta(x_j)} a^\varepsilon(x)\nabla\left(\chi_{\varepsilon,j}^i(1-\rho_\varepsilon)\right)\cdot\nabla z\, dx \quad \forall z \in H_0^1\left(K_\delta\left(x_j\right)\right). \tag{2.38}$$

The Lax-Milgram theorem provides the bound

$$\left\|\nabla\theta_0^i\right\|_{L^2(K_\delta(x_j))} \leq \frac{\beta}{\alpha}\left\|\nabla\left(\chi_{\varepsilon,j}^i(1-\rho_\varepsilon)\right)\right\|_{L^2(K_\delta(x_j))},$$

whose right-hand side is bounded by using the assumptions on ρ_ε and $\chi_{\varepsilon,j}^i \in W^{1,\infty}\left(K_\delta\left(x_j\right)\right)$:

$$\left\|\nabla\left(\chi_{\varepsilon,j}^i(1-\rho_\varepsilon)\right)\right\|_{L^2(K_\delta(x_j))} \leq \left\|\nabla\chi_{\varepsilon,j}^i\right\|_{L^2(\Delta)} + \frac{C}{\varepsilon}\left\|\chi_{\varepsilon,j}^i\right\|_{L^2(\Delta)} \leq C|\Delta|^{1/2}\left\|\chi^i\right\|_{W^{1,\infty}(K)}.$$

From the estimates above and the bound of the volume of the boundary layer $|\Delta| \leq d\frac{\varepsilon}{\delta}|K_\delta|$, we deduce that

$$\left\|\nabla\theta_0^i\right\|_{L^2(K_\delta(x_j))} \leq C|K_\delta|^{\frac{1}{2}}\left(\frac{\varepsilon}{\delta}\right)^{\frac{1}{2}}. \tag{2.39}$$

The difference $\bar{a}_{ik}^0(x_j) - a_{ik}^0$ is decomposed into two terms, I_{ik}^1 and I_{ik}^2. The former stands for the difference between $\varphi_{\delta,j}$ and χ_ε^i and the latter for the mismatch between the sampling

domain and the periodic cell:

$$\bar{a}^0_{ik}(x_j) - a^0_{ik} = I^1_{ik} + I^2_{ik},$$

with

$$I^1_{ik} := \fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\mathbf{e}_k + \nabla\varphi^k_{\delta,j}\right)\cdot\left(\mathbf{e}_i + \nabla\varphi^i_{\delta,j}\right)dx - \fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\mathbf{e}_k + \nabla\left(\rho_\varepsilon\chi^k_\varepsilon\right)\right)\cdot\left(\mathbf{e}_i + \nabla\left(\rho_\varepsilon\chi^i_\varepsilon\right)\right)dx,$$
$$I^2_{ik} := \fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\mathbf{e}_k + \nabla\left(\rho_\varepsilon\chi^k_\varepsilon\right)\right)\cdot\left(\mathbf{e}_i + \nabla\left(\rho_\varepsilon\chi^i_\varepsilon\right)\right)dx - \fint_{K_{n\varepsilon}(x_j)} a^\varepsilon(x)\left(\mathbf{e}_k + \nabla\chi^k_\varepsilon\right)\cdot\left(\mathbf{e}_i + \nabla\chi^i_\varepsilon\right)dx.$$

By symmetry of a^ε and the weak problem eq. (2.38), the term I^1_{ik} equals

$$I^1_{ik} = -\fint_{K_\delta(x_j)} a^\varepsilon(x)\nabla\theta^k_0\cdot\nabla\theta^i_0\,dx + \underbrace{\fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\nabla\varphi_k + \mathbf{e}_k\right)\cdot\nabla\theta^i_0\,dx}_{=0} + \underbrace{\fint_{K_\delta(x_j)} a^\varepsilon(x)\left(\nabla\varphi_i + \mathbf{e}_i\right)\cdot\nabla\theta^k_0\,dx}_{=0},$$

and it is estimated by the Cauchy-Schwarz inequality and the bound eq. (2.39):

$$\left|I^1_{ik}\right| \le \beta\frac{1}{\left|K_\delta\left(x_j\right)\right|}\left\|\nabla\theta^i_0\right\|^2_{L^2(K_\delta(x_j))} \le C\frac{\varepsilon}{\delta}. \tag{2.40}$$

The second term, I^2_{ik} is rewritten as

$$I^2_{ik} = \frac{1}{\left|K_\delta\left(x_j\right)\right|}\int_\Delta a^\varepsilon(x)\left(\mathbf{e}_k + \nabla\left(\rho_\varepsilon\chi^k_\varepsilon\right)\right)\cdot\left(\mathbf{e}_i + \nabla\left(\rho_\varepsilon\chi^i_\varepsilon\right)\right)dx + \left(\frac{1}{\left|K_\delta\left(x_j\right)\right|} - \frac{1}{\left|K_{n\varepsilon}(x_j)\right|}\right)\int_{K_{n\varepsilon}(x_j)} a^\varepsilon(x)\left(\mathbf{e}_k + \nabla\chi^k_\varepsilon\right)\cdot\left(\mathbf{e}_i + \nabla\chi^i_\varepsilon\right)dx,$$

and it is bounded by using the Cauchy-Schwarz inequality and the fact that $\chi^i \in W^{1,\infty}(K)$:

$$\begin{aligned}\left|I^2_{ik}\right| &\le \frac{\beta}{\left|K_\delta\left(x_j\right)\right|}\left\|\mathbf{e}_k + \nabla\left(\rho_\varepsilon\chi^k_\varepsilon\right)\right\|_{L^2(\Delta)}\left\|\mathbf{e}_i + \nabla\left(\rho_\varepsilon\chi^i_\varepsilon\right)\right\|_{L^2(\Delta)} + \frac{|\Delta|}{\left|K_\delta\left(x_j\right)\right|}a^0_{ik}\\ &\le \frac{|\Delta|}{\left|K_\delta\left(x_j\right)\right|}\left[\beta\sup_{1\le i\le d}\left(1 + C\left\|\chi^i\right\|_{L^\infty(K)} + \left\|\nabla\chi^i\right\|_{L^\infty(K)}\right)^2 + a^0_{ik}\right].\end{aligned}$$

Since the term inside the parenthesis can be bounded by a constant independent by δ and ε, we conclude that

$$\left|I^2_{ik}\right| \le C\frac{\varepsilon}{\delta}, \tag{2.41}$$

and eq. (2.37) follows from eqs. (2.40) and (2.41). □

So far, we have seen that using Dirichlet BCs in the corrector ε-problems for periodic multiscale coefficients causes a resonance error which decays as ε/δ. As a matter of fact, the same error decay holds true if other BCs are used or if the size of the sampling domains does not match with the period of the coefficients. This result can be extended to the case of locally periodic coefficients, $a^\varepsilon(x) = a(x, x/\varepsilon)$ where $a(\cdot,\cdot)$ is Lipschitz continuous in its first argument and periodic in the second:

Proposition 2.9. *Let $a \in W^{1,\infty}\left(D, L^\infty\left(K, \mathbb{R}^{d\times d}\right)\right)$ be $[0,1)^d$-periodic in the second variable. Let $u^{0,H}$ and $\bar{u}^H$ be the solutions of eqs.* (2.26) *and* (2.28)*. Then*

$$e_{MOD} := \left\| u^{0,H} - \bar{u}^H \right\|_{H^1(D)} \leq \begin{cases} C\delta & \text{if } \delta/\varepsilon \in \mathbb{N} \text{ and } W(K_\delta(x_j)) = W^1_{per}\left(K_\delta(x_j)\right), \\ C\left(\frac{\varepsilon}{\delta} + \delta\right) & \text{if } \delta/\varepsilon \notin \mathbb{N} \text{ and } W(K_\delta(x_j)) = H^1_0(K_\delta(x_j)). \end{cases}$$

By comparing the error estimates in Propositions 2.6 and 2.9, it can be noticed that approximating the effective coefficient over a sampling domain of size δ has the effect of introducing an additional error term, which scales as δ. However, this term is often negligible with respect to the ε/δ term, thus making the reduction of the latter more crucial.

2.3.3 Reduced basis method for locally periodic coefficients

In connection to the discussion above, we briefly describe here a recently developed strategy for applying the FE-HMM to problems with locally periodic coefficients. In this case, the micro-corrector problem should be solved for each quadrature point x_j, but this is computationally very expensive. Such an increase of the computational cost can be avoided by exploiting the smoothness of $a(x,y)$ with respect to the x-variable and by using the Reduced Basis method to approximate $\bar{a}^0(x_j)$, defined in eq. (2.33), in any arbitrary point x_j. Let us assume that the locally periodic tensor satisfies the affine representation

$$a(x,y) = \sum_{q=1}^{Q} \Theta_q(x) a_q(y), \tag{2.42}$$

where the functions $\Theta_q(x)$ are known and of cheap evaluation and $a_q \in L^\infty(\mathbb{R}^{d\times d})$ are $[0,1)^d$-periodic. Then, it is possible to compute the $\bar{a}^0(x_j)$ on few locations (the "snapshots" of the reduced basis method) and to use them to approximate $\bar{a}^0(x)$ on arbitrary locations x. When the affine decomposition eq. (2.42) is not available, it is possible to approximate it by the Empirical Interpolation Method (EIM, [30, 113]). The application of the reduced basis method to the FE-HMM is known as RB-FE-HMM and was first proposed in [10].

2.4 The resonance error in FE-HMM and alternative corrector problems

In Section 2.3 the FE-HMM was discussed and *a priori* bounds on the approximation errors were provided. Besides the macro- and micro-errors ,which are due to numerical approximations, the FE-HMM suffers from the resonance error, which comes from the coupling conditions between the micro- and the macroscales. In particular, the resonance error has two sources: the mismatching BCs used in the corrector problems and the wrong size of the sampling domains. We have discussed this error in the FE-HMM framework, but other multiscale numerical methods show a similar error due to mismatching coupling conditions between the micro and the macro scales. In the context of the MsFEM, *a priori* error estimates are derived in [96, 97].

Under the assumption of periodicity of the coefficients, the resonance error vanishes only if the corrector problems are solved with periodic boundary conditions on the periodic cell. When the period of the microstructure is not known exactly and the sampling domains of size δ are used to reconstruct it, the resonance error has a magnitude proportional to ε/δ. If the periodicity assumption is relaxed, e.g. if a is random stationary ergodic or quasi-periodic tensor, or in the case of non-linear models, it is not possible to reconstruct exactly the homogenized coefficients a^0, due to the lack of computable microscale models. For example, if a is a stationary ergodic random field, a^0 exists but one should solve and average an auxiliary model over the whole $\mathbb{R}^d$ in order to evaluate a^0. Therefore, the resonance error for non periodic micro structures is not avoidable. Moreover, there are cases for which the resonant effect is even worse than the estimate provided in Proposition 2.6. For example, for stationary ergodic random coefficient, the resonance error scales as $(\varepsilon/\delta)^r$, where r depends on the dimension, but is in general $r < 1$, [55].

From a computational point of view, the first order decay rate of the error is the efficiency and accuracy bottleneck of numerical upscaling schemes. Indeed, while the micro and macro discretization errors can be reduced by *simultaneous* refinement of the micro and macro meshes, [2], or by increment of the FE order at both the macro and the micro scales, the only available strategy to reduce the resonance error is to increase the size of the sampling domains K_δ. This approach is computationally inefficient, as the numerical cost increases as δ^d, while the error decays as δ^{-r}, with $r \leq 1 \leq d$, for possibly every quadrature point in the macro mesh. Hence, in order to reduce the resonance error down to practically reasonable accuracies, one needs to solve the corrector problem eq. (2.23) over "large" ($\delta/\varepsilon \gg 1$) sampling domains K_δ. This inefficiency triggered the birth and development of a number of numerical methodologies aiming at improving the resonance error convergence rate, in order to reduce it under acceptable accuracies without substantial enlargement of the sampling domains K_δ.

2.4.1 Existing approaches for reducing the resonance error

Over the last two decades, several interesting approaches have been proposed to reduce the resonance error. It is worth mentioning that the improvement of the resonance error decay is an active subject of research in the MsFEM framework, too. Several strategies have been proposed in this context , such as oversampling [97, 90], Petrov–Galerkin approach [98] and the Localized Orthogonal Decomposition (LOD) [88, 89, 114, 66]. They are mostly based on the choice of the FE spaces at the micro scale.

Most of the techniques proposed in the FE-HMM framework rely on modified microscale models to achieve better accuracies in the estimation of the homogenized coefficients. Some of these approaches are described below.

From now on, we will use the notation of the *rescaled* corrector problems, consistently to what is done in the next chapters and several of the works cited below. Hence, we will not consider the corrector problem

$$-\nabla\cdot\left(a^{\varepsilon}(x)\left(\nabla\varphi_{\delta}^{i}+\mathbf{e}_i\right)\right)=0 \quad \text{on } K_{\delta},$$

but its ε-rescaled version:

$$-\nabla\cdot\left(a(y)\left(\nabla\chi_{R}^{i}+\mathbf{e}_i\right)\right)=0 \quad \text{on } K_{R}, \tag{2.43}$$

where we have used the rescaling

$$y=\frac{x}{\varepsilon}, \quad \chi_{R}^{i}(y)=\frac{1}{\varepsilon}\varphi_{\delta}^{i}(x), \quad \text{and } R=\frac{\delta}{\varepsilon}.$$

Oversampling and filtering

The first attempt to improve the convergence rate was by modifying the averaging formula eq. (2.33). In oversampling, the microscopic corrector problem eq. (2.23) is solved over the cell K_R, while the computation of the homogenized coefficient takes place in an interior domain $K_L \subset K_R$:

$$\begin{cases} -\nabla\cdot\left(a(y)\left(\nabla\chi_{R}^{i}+\mathbf{e}_i\right)\right)=0 & \text{on } K_R \\ \chi_{R}^{i}=0 & \text{on } \partial K_R, \end{cases} \tag{2.44}$$

$$a_{ij}^{0,R,L}:=\int_{K_L}\mathbf{e}_i\cdot a(y)\left(\nabla\chi_{R}^{j}+\mathbf{e}_j\right)\mu_L(y)\,dy, \tag{2.45}$$

where μ_L is a smooth function of mass one and supported on K_L (the filter). Another attempt is based on exploring the combined effect of oversampling and imposing different BCs (Dirichlet, Neumann and periodic) in eq. (2.23), see [134]. It has been found that periodic BCs perform better than the other two. Moreover, the Dirichlet BCs tend to overestimate the effective coefficients, while Neumann BCs underestimate them. The use of these strategies becomes questionable if one is interested in practically relevant error tolerances, since there

is still a need for substantially enlarging the computational domain K_R to reach a satisfactory accuracy. Additionally, the weighted averaging of the fluxes is compared with the standard geometric average of eq. (2.33): thought the weighted average (filtering), does not improve the convergence rate, it reduces the prefactor in the *a priori* bound for the resonance error.

Filtered corrector problems

As it was remarked above, the resonance error is due to two factors: mismatch in the boundary conditions of the corrector problems and the true solution, and the averaging over a bounded domain. In order to address these issues, a *filtered corrector problem* to approximate the homogenized limit of (non-)periodic coefficients was proposed, [33]:

$$\begin{cases} -\nabla\cdot\left[\left(a(y)\left(\nabla\chi_R^i+\mathbf{e}_i\right)+\lambda\right)\mu_R(y)\right]=0 \quad \text{on } K_R \\ \int_{K_R}\nabla\chi_R^i\mu=0, \end{cases} \tag{2.46}$$

where $p,\lambda\in\mathbb{R}^d$ and μ is a *filter*, a class of functions defined later. The Lagrange multiplier λ allows us to solve eq. (2.46) without imposing any boundary condition. Indeed, it is not difficult to prove that eq. (2.46) is well posed in the quotient space (following the notation of [33]) $H^1_\mu(K_R)/\mathbb{R}$, where

$$L^2_\mu(K_R)=\left\{u:K_R\mapsto\mathbb{R},\ \text{measurable},\ \int_{K_R}u^2\mu<\infty\right\},\ \text{and}$$
$$H^1_\mu(K_R)=\left\{u\in L^2_\mu(K_R),\nabla u\in\left(L^2_\mu(K_R)\right)^d\right\}.$$

By-passing the use of boundary conditions to solve the corrector problems allows us to get rid of the first source of the resonance error.

The homogenized coefficients are then approximated by

$$\begin{aligned} a_{ij}^{0,R}&:=\int_{K_R}\left(\nabla\chi_R^i+\mathbf{e}_i\right)\cdot a(y)\left(\nabla\chi_R^j+\mathbf{e}_j\right)\mu_R(y)\,dy \\ &=\int_{K_R}\mathbf{e}_i\cdot a(y)\left(\nabla\chi_R^j+\mathbf{e}_j\right)\mu_R(y)\,dy. \end{aligned} \tag{2.47}$$

The filter μ_R is used here in order to address the second source of error, i.e. the averaging over a "wrong" domain. Filters are a class of compact support functions with unit mass, q-th order regularity up to the boundary and vanishing derivatives up to the order $q-1$. This class of function is a powerful tool to average periodic and quasi-periodic functions, as it is shown in [39, 40] and in Definition 3.7. Indeed, they allow to approximate the average of (quasi-)periodic functions with accuracy R^{-q}: Let us consider a quasi-periodic function $f:\mathbb{R}^d\mapsto\mathbb{R}$ and let us define

$$\langle f\rangle:=\lim_{R\to\infty}\fint_{K_R}f\left(y\right)dy,$$

then, there exists $C > 0$ independent of δ such that

$$\left| \int_{K_R} f(y)\, \mu_R\, dy - \langle f \rangle \right| \leq C R^{-q}. \tag{2.48}$$

This approach is analysed in [33]. The resonance error $\left| a^{0,R} - a^0 \right|$ decays as R^{-q} for one dimensional, periodic coefficients. By using the multiscale expansion ansatz, a second order convergence rate R^{-2} is shown in higher dimension, independently of the order of the filter. Numerical simulations demonstrate the optimality of the second order rate in dimension $d > 1$.

An elliptic model with zero-th order term

Another promising strategy is to use an elliptic corrector problem with a small (i.e., converging to zero) zero-th order regularization term, [78]:

$$\frac{1}{T}\chi^i_{T,R} - \nabla \cdot \left(a(y) \left(\nabla \chi^i_{T,R} + \mathbf{e}_i \right) \right) = 0 \text{ on } K_R, \tag{2.49}$$

with suitable boundary conditions to ensure well posedness. This problem is widely used in stochastic homogenization to prove existence of the corrector functions. In the non-stochastic case, the cell problem eq. (2.49) allows to reduce the effect that mismatching boundary conditions have on the values of the solution *inside* the domain. Indeed, the exponential decay of the Green's function for eq. (2.49) entails exponentially fast decay of the boundary error. The homogenized coefficients are then computed by using the filters of eq. (2.47), rescaled in order to be supported on the smaller domain $K_L \subset K_R$:

$$a^{0,R,L,T}_{ij} := \int_{K_R} \left(\nabla \chi^i_{T,R} + \mathbf{e}_i \right) \cdot a(y) \left(\nabla \chi^j_{T,R} + \mathbf{e}_j \right) \mu_L(y)\, dy \tag{2.50}$$

This method suffers from a bias (or systematic error) due to added regularization term, which limits the convergence rate to fourth order. The error bound for this approach is

$$\left| a^{0,R,L,T}_{ij} - a^0_{ij} \right| \leq C \left(L^{-q-1} + T^{-2} + T^{1/4} \exp\left(-c \frac{|R-L|}{\sqrt{T}} \right) \right). \tag{2.51}$$

If an high order filter ($q > 3$) is chosen and the condition $L \gg \sqrt{T}$ holds, then, by choosing

$$R = 3L/2, \quad T = L^2 \log(L)^{-4}$$

we have $\left| a^{0,R,L,T}_{ij} - a^0_{ij} \right| \leq C R^{-4} \left(\log R \right)^8$. Numerical simulations in [78] show that the method requires very large values of R to achieve the optimal fourth order asymptotic rate.

In [79], Richardson extrapolation is used to increase the convergence rate to higher orders at the expense of solving the corrector problem several times with different regularization terms.

The Richardson iterates are defined by

$$\chi_{T,R,k+1} = \frac{1}{2^k - 1}\left(2^k \chi_{2T,R,k} - \chi_{T,R,K}\right).$$

If the k-th iterate is used in eq. (2.51) to compute the approximate homogenized tensor, the error bound improves to

$$\left|a_{ij}^{0,R,L,T} - a_{ij}^0\right| \leq C\left(L^{-q-1} + T^{-2k} + T^{1/4}\exp\left(-c\frac{|R-L|}{\sqrt{2^{k-1}T}}\right)\right).$$

This last error bound allows to reach arbitrary high convergence rate in the asymptotic regime $R \to \infty$.

The wave equation approach

An interesting idea to cancel the error due to mismatching boundary condition in the cell problem is proposed in [16, 18]. Here, a second order hyperbolic equation on $K_R \times (-T/2, T/2)$ is solved:

$$\begin{cases} \partial_{tt}\chi_R - \nabla\cdot\left(a(y)\left(\nabla\chi_R^i + \mathbf{e}_i\right)\right) = 0 & \text{in } K_R \times (-T/2, T/2), \\ \chi_R^i = 0, & \text{in } K_R \times 0, \\ \partial_t \chi_R^i = 0, & \text{in } K_R \times 0, \end{cases} \tag{2.52}$$

with suitable boundary conditions. The rationale behind this approach is that the boundary conditions do not affect the solution in any interior subdomain sufficiently far from the boundary, because of the finite speed of wave propagation. The approximation of a^0 is computed by averaging over a subdomain $K_L \times (-T, T)$:

$$a_{ij}^{0,R} = \int_{-T/2}^{T/2}\int_{K_L} \mathbf{e}_i \cdot a^\varepsilon(y)\left(\nabla\chi_R^j + \mathbf{e}_j\right)\mu_L(y)\,dy\,\mu_T(t)\,dt, \tag{2.53}$$

where the filtering function μ_L and μ_T belong to the same class of filters used in the sections above. The choices $T = L$ and $L \leq R - T\sqrt{\|a\|_{L^\infty}}$ ensure that the boundary conditions do not pollute the solution in the averaging domain K_L. The only error present in the upscaling algorithm is thus the one connected to the averaging which, by the periodicity of a^ε, decays as L^{-q-1}. Hence, this method provides an arbitrary rate of accuracy in approximating the G-limit of periodic media. However, there are a few computational challenges with this method:

- the spatial domain size increases linearly with the wave speed;
- the solution of the wave equation depends on time, and therefore additional degrees of freedom are needed to approximate the cell-solution;
- practically, accurate approximations of solutions of the wave equation require high resolutions per-wavelength, which makes the method less efficient (when compared to solving an elliptic cell-problem).

The embedded method

An embedded corrector problem is proposed in [41, 42]. The original tensor $a(\cdot)$ is replaced in the (infinite) sampling domains by $a_R(\cdot)$ defined as

$$a_R(y) = \begin{cases} a(y) & y \in K_R, \\ \bar{a} & y \in \mathbb{R}^d \setminus K_R, \end{cases}$$

where $\bar{a} \in \mathbb{R}^{d\times d}$ is an *a priori unknown* constant matrix approximating the homogenized coefficients. The *embedded corrector problem* is

$$-\nabla \cdot \left(a_R(y)\left(\chi_R + \xi\right)\right) = 0 \quad \text{in } \mathscr{D}'\left(\mathbb{R}^d\right). \tag{2.54}$$

The solution χ_R to the corrector problem eq. (2.54) can be used to define consistent approximations of a^0. Let us define the vector space

$$V_0 := \left\{ v \in L^2_{loc}\left(\mathbb{R}^d\right), \nabla v \in \left(L^2\left(\mathbb{R}^d\right)\right)^d, \int_{K_R} v = 0 \right\}.$$

The differential equation eq. (2.54) is the Euler–Lagrange equation for the minimization problem: Find $\mathscr{J}_p(\bar{a}) \in \mathbb{R}$ such that

$$\mathscr{J}_p(\bar{a}) := \min_{v\in V_0} \frac{1}{2|K_R|}\int_{K_R} (\nabla v + \xi)\cdot a^\varepsilon(y)\,(\nabla v + \xi)\,dy \\ + \frac{1}{2|K_R|}\int_{\mathbb{R}^d\setminus K_R} \nabla v \cdot \bar{a}\nabla v\,dy - \frac{1}{|K_R|}\int_{\partial K_R} \mathbf{n}\cdot\bar{a}\xi\, v\,d\sigma(y). \tag{2.55}$$

Next, by linearity of the mapping $\xi \mapsto \chi_R^\xi$ we know that there exists a matrix $G(\bar{a}) \in \mathbb{R}^{d\times d}$ such that

$$\frac{1}{2}\xi\cdot G(\bar{a})\xi = \mathscr{J}_p(\bar{a}).$$

The following three approximations of a^0 are proposed:

$$a_1^{0,R} = \operatorname*{argmax}_{\bar{a}\in\mathscr{M}(\alpha,\beta)} \operatorname{Tr}\left(G(\bar{a})\right), \quad a_2^{0,R} = G(a_1^{0,R}), \quad a_3^{0,R} = G(a_3^{0,R}).$$

Then, it can be proved that $a_1^{0,R}, a_2^{0,R} \to a^0$ as $R\to\infty$ and that there exist a subsequence $\left\{a_3^{0,R_k}\right\}_k$ such that $a_3^{0,R_k} \to a^0$. The embedded method shows a first order convergence rate for $a_{1,}^{0,R}$ and a second order of convergence rate for $a_2^{0,R}$ and $a_3^{0,R}$, with respect to the modelling parameter R.

An iterative method

A new method for efficient computation of the homogenized coefficients was developed in [119]. The method was originally proposed for the homogenization of stochastic coefficients

on random networks and later adapted to the continuous case in [87]. The main goal is not to provide better convergence rates of the resonance error but, rather, to propose an iterative scheme to reconstruct the homogenized coefficients and to derive estimates for the error introduced by the finite number of iterations. Nevertheless, we think that it is worth to mention this method in this section, as the corrector problems proposed in Chapter 4 are based on it.

This iterative approach can be seen as the backward Euler discretization in time of the parabolic operator $\partial_t - \nabla\cdot(a\nabla)$, where the spatial differential operators are to be meant in a discrete settings. The first element of the iterative scheme is $\chi_{-1} = \nabla\cdot\big(a(y)\xi\big)$, and the k-th iteration, for $k \le n$, is defined by solving

$$2^{-k}\chi_k - \nabla\cdot\big(a(y)\nabla\chi_k\big) = 2^{-k}\chi_{k-1} \quad \text{in } \mathbb{Z}^d. \tag{2.56}$$

By using the properties of the Green's function, it is possible to prove that the correctors solving eq. (2.56) can be approximated within an exponentially decaying error by the bounded domain solutions

$$\begin{cases} 2^{-k}\chi_{R,k} - \nabla\cdot\big(a(y)\nabla\chi_{R,k}\big) = 2^{-k}\chi_{R,k-1} & \text{in } K_{R_k}, \\ \chi_{R,k} = 0 & \text{on } \partial K_{R_k}, \end{cases} \tag{2.57}$$

so that the boundary error is negligible compared to the systematic and statistical errors. These two errors are due to:

- the use of a model with the addition of a zero-th order term, and
- the computation of the homogenized tensor by averaging over bounded domains:

$$\xi\cdot a^{0,R,L,n}\xi := \fint_{K_{L_0}} \xi\cdot a(y)\xi\, dy + \sum_{k=0}^{n} 2^k \fint_{K_{L_k}} \chi_{k-1}\chi_k + \chi_k^2\, dy. \tag{2.58}$$

In eqs. (2.57) and (2.58) we have defined

$$L_k = 2^{n-\left(\frac{1}{2}-\varepsilon\right)k}, \quad \text{and} \quad R_k = L_k + C(1+n)2^{\frac{k}{2}}.$$

In this case, the mean square resonance error is bounded by

$$\mathbb{E}\left[\left|\xi\cdot\big(a^0 - a^{0,R,L,n}\big)\xi\right|^2\right]^{\frac{1}{2}} \le CL_n^{-\frac{d}{2}}. \tag{2.59}$$

2.4.2 Conclusion

As discussed above, many approaches to reconstruct the homogenized coefficients with reduced resonance error have been proposed in the past years. Some of these methods could only achieve a reduction of the prefactor in the error bounds, while others improved the speed

of convergence with respect to the size of the sampling domain. High convergence rates for the resonance error will allow to estimate the homogenized coefficients with better accuracy and reduced computational cost, thus being valuable for simulating multiscale materials. However, some of the proposed strategies have limited orders of convergence or the error decays at the expected rate only in the asymptotic regime, i.e. for sufficiently large values of R. This motivated us to carry out the research exposed in this thesis.

3 New parabolic and modified elliptic corrector problems

Let us consider the second order elliptic equation:

$$\begin{cases} -\nabla\cdot(a^\varepsilon(x)\nabla u^\varepsilon) = f & \text{in } D\subset\mathbb{R}^d, \\ u^\varepsilon = 0 & \text{on } \partial D, \end{cases} \tag{3.1}$$

where the tensor a^ε oscillates at the small scale, denoted by $\varepsilon \ll 1$. The macroscopic behaviour of this equation is approximated by solving the homogenized problem:

$$\begin{cases} -\nabla\cdot\left(a^0(x)\nabla u^0\right) = f & \text{in } D\subset\mathbb{R}^d, \\ u^0 = 0 & \text{on } \partial D. \end{cases}$$

In general, the homogenized matrix a^0 cannot be computed in closed form, thus multiscale numerical methods are used to approximate it by solving auxiliary equations at the microscale. However, such methods are affected by the resonance error which strongly limits the accuracy of multiscale simulations. In the FE-HMM context, the resonance error depends on the size of the sampling domains and the imposed boundary conditions (BCs) in the corrector problems. Letting R denote the size of the sampling domain, the resonance error for periodic coefficients decays as R^{-1}, as proved in Section 2.3.2. Thus, the objective of many alternative corrector problems, as those presented in Section 2.4.1, is to approximate the correct homogenized coefficients within an improved decay of the resonance error. In this chapter, we present two novel methods for computing the effective parameters of linear second order elliptic PDEs with fast oscillating coefficients. The main advantage of the two methods is that their resonance error decays with arbitrary rates of convergence.

Outline

First of all, we prove an equivalence result between the periodic elliptic correctors and the solutions of parabolic PDEs with periodic boundary conditions. This is a fundamental result that will be used to derive the two new schemes, as well as in the *a priori* error analysis. A

proof is given in Section 3.1.

The first upscaling method is described in Section 3.2. It relies on the solution of parabolic problems at the microscale, which was already proved to be successful in the context of stochastic homogenization over discrete networks, [119]. The advantage of using parabolic corrector problems is that the correctors are minimally affected by mismatching boundary conditions, thanks to the exponential decay of the Green's function. This allows to prove that the convergence rate of the resonance error for this model is arbitrarily high, as it will be proven in Chapter 4.

As a second step, we derive a modified elliptic corrector problem as the time integral of the parabolic one in Section 3.3. The integration in time of the parabolic correctors brings a new term in the right-hand side of the elliptic corrector problem, so we will denote it as the "modified elliptic" approach. The connection with the parabolic problem allows to transfer many of its properties to the elliptic problem, for example the exponential decay of the error due to mismatching boundary conditions. Hence, we can achieve arbitrarily high convergence rates, as for the parabolic case. The proof of the convergence rate for this method is postponed to Chapter 5.

The content of this chapter is based on [6].

3.1 Equivalence between the parabolic and the standard elliptic problems in the periodic setting

In this section, we discuss an approach to compute a^0 based on parabolic auxiliary problems and prove its equivalence to the use of periodic correctors. This will allow to derive two corrector problems suitable for the computation of effective coefficients under more general assumptions. Let us assume that the multiscale tensor a^ε satisfies:

i) $a^\varepsilon(x) = a(x/\varepsilon)$, for $a \in \mathscr{M}(\alpha, \beta)$;

ii) $a(\cdot)$ is K-periodic, with $K := [-1/2, 1/2]^d$.

Under assumptions i) and ii), the exact homogenized tensor can be computed as

$$a^0_{ij} := \fint_K \mathbf{e}_i \cdot a(x)\left(\nabla \chi^j + \mathbf{e}_j\right) dx, \tag{3.2}$$

where the corrector functions χ^i solve the periodic auxiliary problems of Section 2.1.1:

$$\begin{cases} -\nabla \cdot \left(a(x)\left(\nabla \chi^i + \mathbf{e}_i\right)\right) = 0 & \text{in } K \\ \chi^i \text{ is } K\text{-periodic}, & \int_K \chi^i = 0. \end{cases} \tag{3.3}$$

By symmetry of a and the weak form of eq. (3.3), a^0 is equivalently computed as

$$a^0_{ij} = \fint_K a_{ij}(x)\,dx - \fint_K \nabla\chi^i \cdot a(x)\nabla\chi^j\,dx, \tag{3.4}$$

where second term of eq. (3.4) is denoted as *correction term*.

Let us introduce the following parabolic problem with periodic boundary conditions

$$\begin{cases} \dfrac{\partial v^i}{\partial t} - \nabla\cdot(a(x)\nabla v^i) = 0 & \text{in } K\times(0,+\infty) \\ v^i(\cdot,t) \quad K\text{-periodic}, \forall\, t\geq 0 & \\ v^i(x,0) = \nabla\cdot(a(x)\mathbf{e}_i) & \text{in } K. \end{cases} \tag{3.5}$$

The solution of eq. (3.5) is well-posed in the space $L^2\left([0,+\infty), W^1_{per}(K)\right)\cap C\left([0,+\infty), L^2_0(K)\right)$.

Proposition 3.1. *Let $a\in\mathscr{M}(\alpha,\beta)$ be K-periodic and $\nabla\cdot(a(x)\mathbf{e}_i)\in L^2_0(K)$. Then, eq. (3.5) has a unique weak solution v^i such that*

$$v^i \in L^2\left([0,+\infty), W^1_{per}(K)\right), \partial_t v^i \in L^2\left([0,+\infty), W^1_{per}(K)'\right).$$

It follows that $v^i\in C\left([0,+\infty), L^2_0(K)\right)$, and there exist constants $C>0$ such that the following bounds hold true:

$$\left\|v^i\right\|_{L^\infty\left([0,+\infty),L^2(K)\right)} + \left\|v^i\right\|_{L^2\left([0,+\infty),W^1_{per}(K)\right)} \leq C\left\|\nabla\cdot(a(x)\mathbf{e}_i)\right\|_{L^2(K)}. \tag{3.6}$$

Moreover, v^i is Hölder continuous in $K\times(0,+\infty)$.

Here, the space $W^1_{per}(K)'$ is the dual space of $W^1_{per}(K)$ (a characterization of this space can be found in [45]). Next, we prove that a^0, defined as in eq. (3.4), can be equivalently computed through the solutions v_i's.

Theorem 3.2. *Let $a(\cdot)\in\mathscr{M}(\alpha,\beta)$ be K-periodic, $v^i\in C\left([0,+\infty), L^2_0(K)\right)$ be the weak solution of eq. (3.5) and $\chi^i\in W^1_{per}(K)$ be the weak solution of eq. (3.3). Then, for $1\leq i,j\leq d$, the following identities hold*

$$\chi^i = \int_0^{+\infty} v^i(\cdot,t)\,dt \quad \text{in } W^1_{per}(K), \tag{3.7}$$

$$\frac{1}{2}\int_K \nabla\chi^i(x)\cdot a(x)\nabla\chi^j(x)dx = \int_0^{+\infty}\int_K v^i(x,t)\,v^j(x,t)\,dx\,dt. \tag{3.8}$$

Proof. We reformulate problem (3.11) as the abstract Cauchy problem in $L^2_0(K)$

$$\begin{cases} \dfrac{dv^i}{dt} + Av^i = 0 \\ v^i(0) = g^i, \quad g^i(x) = \nabla\cdot(a(x)\mathbf{e}_i) \text{ in } L^2_0(K). \end{cases}$$

Here, the operator $A : W^1_{per}(K) \to W^1_{per}(K)'$ is defined as $Av := -\nabla \cdot (a\nabla v)$. Then, $v^i(t) = e^{-tA} g^i$. We know that $\sigma(A)$, the spectrum of A, is contained in an open sectorial domain $\alpha + S_\omega$, where $\alpha \in \mathbb{R}$, $\alpha > 0$ and

$$S_\omega = \left\{ z \in \mathbb{C} : \left| \arg z \right| < \omega, 0 < \omega < \frac{\pi}{2} \right\}.$$

Then, the Dunford integral representation

$$e^{-tA} = \frac{1}{2\pi i} \int_\Gamma e^{-tz} (zI - A)^{-1} \, dz$$

holds, where Γ is an infinite curve lying in $\rho(A) := \mathbb{C} \setminus \sigma(A)$ and surrounding $\sigma(A)$ counterclockwise. Then, integrating in time we obtain

$$\begin{aligned} \int_0^{+\infty} v^i(t) \, dt &= \int_0^{+\infty} \frac{1}{2\pi i} \int_\Gamma e^{-tz} (zI - A)^{-1} g^i \, dz \, dt \\ &= \frac{1}{2\pi i} \int_\Gamma \int_0^{+\infty} e^{-tz} \, dt \, (zI - A)^{-1} g^i \, dz \\ &= \frac{1}{2\pi i} \int_\Gamma \frac{1}{z} (zI - A)^{-1} g^i \, dz = A^{-1} g^i. \end{aligned}$$

The first equality is given by the Dunford integral formula. The second equality is obtained by Fubini's theorem. The third equality is true because the double integral is bounded, $\lim_{t \to +\infty} e^{-tz} = 0$ since $Re(z) > 0$ on Γ. The last equality follows from the fact that the function $f(z) = 1/z$ is holomorphic in the interior of $\alpha + S_\omega$. Since A is an isomorphism and χ^i is the weak solution of $A\chi^i = g^i$, we have that $A^{-1} g^i = \chi^i$ and (3.7) is proved.

To prove (3.8), we write the weak formulation of (3.3) and choose $\chi^j = \int_0^{+\infty} v^j \, dt$ as test function:

$$\begin{aligned} \int_{K_R} \nabla \chi^j \cdot a(x) \nabla \chi^i \, dy &= \left(\nabla \cdot (a \mathbf{e}_i), \chi^j \right)_{L^2_0(K)} \\ &= \int_0^{+\infty} \left(\nabla \cdot (a \mathbf{e}_i), v^j \right)_{L^2_0(K)} \, dt. \end{aligned}$$

Using the semigroup property of e^{-tA} and the self-adjointness of A we obtain

$$\begin{aligned} \int_K \nabla \chi^i(x) \cdot a(x) \nabla \chi^j(x) \, dy &= \int_0^{+\infty} \left(v^i(\cdot, 0), e^{-tA} v^i(\cdot, 0) \right)_{L^2_0(K)} \, dt \\ &= \int_0^{+\infty} \left(v^i(\cdot, 0), e^{-\frac{t}{2}A} e^{-\frac{t}{2}A} v^i(\cdot, 0) \right)_{L^2_0(K)} \, dt \\ &= \int_0^{+\infty} \left(e^{-\frac{t}{2}A} v^i(\cdot, 0), e^{-\frac{t}{2}A} v^i(\cdot, 0) \right)_{L^2_0(K)} \, dt \end{aligned}$$

$$= \int_0^{+\infty} \left(v^i(\cdot, t/2), v^j(\cdot, t/2) \right)_{L^2(K_R)} dt,$$

and conclude the proof by the change of variable $t/2 \mapsto t$. □

Theorem 3.2 implies that a^0 can equivalently be computed by using the parabolic correctors v^i's. Indeed, we can either plug eq. (3.8) into eq. (3.4) or eq. (3.7) into eq. (3.2) to get the equivalent formulations eq. (3.9) and eq. (3.10).

Corollary 3.3. *Let $a(\cdot) \in \mathcal{M}(\alpha, \beta)$ be K-periodic. Let $v^i \in C\left([0,+\infty), L^2_0(K)\right)$ solve eq.* (3.5)*. Then*

$$a^0_{ij} = \fint_K a_{ij}(x)\, dx - 2 \int_0^{+\infty} \fint_K v^i(x,t) v^j(x,t)\, dx\, dt, \tag{3.9}$$

$$a^0_{ij} = \fint_K \mathbf{e}_i \cdot a(x) \left(\int_0^{+\infty} \nabla v^j(x,t)\, dt + \mathbf{e}_j \right) dx. \tag{3.10}$$

The two formulations, eqs. (3.9) and (3.10), will be used in Chapters 4 and 5 to prove *a priori* estimates for the resonance error.

3.2 The parabolic corrector problems

In this section we propose a parabolic corrector problem to approximate a^0. This approach does not rely on the periodicity assumption for $a(\cdot)$, thus it is suitable for a very general classes of tensors. The parabolic correctors are defined as the solutions u^i_R, $i = 1, \dots, d$ of the following problems:

$$\begin{cases} \dfrac{\partial u^i_R}{\partial t} - \nabla \cdot (a(x) \nabla u^i_R) = 0 & \text{in } K_R \times (0,+\infty) \\ u^i_R = 0 & \text{on } \partial K_R \times (0,+\infty) \\ u^i_R(x,0) = \nabla \cdot (a(x)\mathbf{e}_i) & \text{in } K_R, \end{cases} \tag{3.11}$$

The well-posedness of eq. (3.11) is well-known (see, e.g., [111]), and is summarized below.

Proposition 3.4. *Let $a \in \mathcal{M}(\alpha, \beta, K_R)$ and $a\mathbf{e}_i \in H_{div}(K_R)$. Then, eq.* (3.11) *has a unique weak solution u^i_R such that*

$$u^i_R \in L^2\left([0,+\infty), H^1_0(K_R)\right), \partial_t u^i_R \in L^2\left([0,+\infty), H^{-1}(K_R)\right).$$

It follows that $u^i_R \in C\left([0,+\infty), L^2(K_R)\right)$, and there exists a constant $C > 0$ such that the following bound holds true:

$$\left\| u^i_R \right\|_{L^\infty\left([0,+\infty), L^2(K_R)\right)} + \left\| u^i_R \right\|_{L^2\left([0,+\infty), H^1_0(K_R)\right)} \leq C \left\| \nabla \cdot (a(\cdot)\mathbf{e}_i) \right\|_{L^2(K_R)}.$$

Moreover, u^i_R is Hölder continuous in $K_R \times (0,+\infty)$.

We now describe how the parabolic correctors can be used to approximate the homogenized tensor. As a first guess, we could replace K with K_R and $v^i(x,t)$ with $u_R^i(x,t)$ in eq. (3.9) and define

$$a^{0,R,R,+\infty} := \fint_{K_R} a_{ij}(x)\,dx - 2\int_0^{+\infty} \fint_{K_R} u_R^i(x,t)\,u_R^j(x,t)\,dx\,dt \tag{3.12}$$

as an approximant of a^0, but this strategy does not bring any advantage. The reason is that the use of the parabolic method is equivalent to the standard approach of eq. (2.43), as the following Proposition 3.5 and Corollary 3.6 show. The proof is essentially identical to the one of Theorem 3.2 and is thus omitted.

Proposition 3.5. *Let $a(\cdot) \in \mathscr{M}(\alpha,\beta)$, $u_R^i \in C\left([0,+\infty), L^2(K_R)\right)$ be the weak solution of eq.* (3.11) *and $\chi_R^i \in H_0^1(K_R)$ be the weak solution of eq.* (2.43)*. Then, for $1 \le i,j \le d$, the following identities hold*

$$\chi_R^i = \int_0^{+\infty} u_R^i(\cdot,t)\,dt \quad \text{in } H_0^1(K_R),$$
$$\frac{1}{2}\int_{K_R} \nabla\chi_R^i(x)\cdot a(x)\nabla\chi_R^j(x)dx = \int_0^{+\infty}\int_{K_R} u_R^i(x,t)\,u_R^j(x,t)\,dx\,dt.$$

Corollary 3.6. *Let $a(\cdot) \in \mathscr{M}(\alpha,\beta)$ be K-periodic, $a^{0,R}$ be defined by* (2.45) *and $a^{0,R,R,+\infty}$ be defined by* (3.12)*. Then*

$$a^{0,R,R,+\infty} = a^{0,R}.$$

Hence, using the classical result stated in Section 2.3.2, there exist a constant $C>0$ independent of R such that

$$\left\| a^{0,R,R,+\infty} - a^0 \right\|_F \le \frac{C}{R}.$$

This result implies that the formula (3.12) must be modified in order to achieve higher convergence orders. This is achieved by truncating the time integral at a finite time T and employing the kernel averages of Definition 3.7.

Smooth averaging functions are commonly used to accurately approximate the mean of K-periodic function f over a larger domain K_R. Here, we define a class of averaging kernels (also known as *filters*) that can be used to approximate a^0. Filters have the property of approximating the average of periodic functions with arbitrary rate of accuracy, as stated in Lemma 4.3.

Definition 3.7 (Definition 3.1 in [78])**.** *We say that a function $\mu : [-1/2,1/2] \mapsto \mathbb{R}^+$ belongs to the space $\mathbb{K}^q$, $q \ge 1$, if:*

i) $\mu \in C^q([-1/2,1/2]) \cap W^{q+1,\infty}((-1/2,1/2))$;

ii) $\mu^{(k)}(-1/2) = \mu^{(k)}(1/2) = 0,\ \forall k \in \{0,\dots,q-1\}$;

iii) $\displaystyle\int_{-\frac{1}{2}}^{\frac{1}{2}} \mu(y)\,dy = 1.$

For $q = 0$, we define $\mu \in \mathbb{K}^0$ as $\mu(y) = \mathbb{1}_{[-1/2,1/2]}$, where $\mathbb{1}_I$ is the characteristic function on the interval I.

In dimension d, a q-th order filter $\mu_L : K_L \subset \mathbb{R}^d \to \mathbb{R}^+$ with $L > 0$ is defined by

$$\mu_L(x) := L^{-d} \prod_{i=1}^{d} \mu\Big(\frac{x_i}{L}\Big),$$

where μ is a one dimensional q-th order filter and $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$. In this case, we will say that $\mu_L \in \mathbb{K}^q(K_L)$. Note that filters μ_L are considered extended to 0 outside of K_L.

Hence, we propose to approximate the homogenized matrix by modifying the correction part of eq. (3.12):

$$a_{ij}^{0,R,L,T} = \int_{K_L} a_{ij}(x)\mu_L(x)\,dx - 2\int_0^T \int_{K_L} u_R^i(x,t)u_R^j(x,t)\mu_L(x)\,dx\,dt, \tag{3.13}$$

where $L \le R$, $0 < T < +\infty$ and the function μ_L belongs to a class of filtering functions given in Definition 3.7. The goal of Chapter 4 is to prove that the resonance error for the parabolic reconstruction of the homogenized coefficients decays with an arbitrary rate of convergence, upon choosing $L, T = \mathcal{O}(R)$:

$$e_{MOD} := \left\| a^{0,R,L,T} - a^0 \right\|_F \le C\left[R^{-(q+1)} + e^{-\sqrt{2\lambda_0 c}(1-k_o)R}\right]. \tag{3.14}$$

The constants $q \in \mathbb{N}$, $\lambda_0, c, C \in \mathbb{R}_+$, $0 < k_o < 1$ will be specified later.

3.3 The modified elliptic corrector problems

In Section 3.2 we proposed a parabolic corrector problem to compute the homogenized coefficients. A key-point for achieving better convergence rates in the approximation of a^0 is to use a filtering function and integrate over a *finite* interval $[0, T]$ in time. Based on this last consideration, and in view of eq. (3.10) for computing a^0, we will now derive a modified elliptic corrector problem. The statement of the following Theorem 3.8 is proved by the eigenvalue decomposition of the operator $A : H_0^1(K_R) \mapsto H^{-1}(K_R)$ defined by

$$A := -\nabla \cdot (a\nabla),$$

but it could be equally proved by exploiting the Dunford integral representation, as in Theorem 3.2.

Theorem 3.8. *Let $a \in \mathcal{M}(\alpha, \beta, K_R)$ and $a\mathbf{e}_i \in H_{div}(K_R)$. Let u_R^i be the weak solution of eq. (3.11) and $g^i(x) := \nabla \cdot (a(x)\mathbf{e}_i)$. Then, the function $\chi_{T,R}^i \in H_0^1(K_R)$ defined as the time integral*

$$\chi_{T,R}^i(x) := \int_0^T u_R^i(x,t)\,dt$$

is the unique solution of

$$\begin{cases} -\nabla\cdot\Big(a(x)\nabla\chi^i_{T,R}(x)\Big) = g^i(x) - [e^{-AT}g^i](x) & \text{in } K_R, \\ \chi^i_{T,R}(x) = 0 & \text{on } \partial K_R, \end{cases} \tag{3.15}$$

and it satisfies

$$\|\chi^i_{T,R}\|_{H^1_0(K_R)} \leq C\|a\mathbf{e}_i\|_{H_{div}(K_R)},$$

where $C^2 = \frac{1+C_p^2}{\alpha^2}$, C_p *is the Poincaré constant and* α *is the coercivity constant.*

Proof. Let $\{\varphi_k\}_{k=0}^{\infty} \subset H^1_0(K_R)$ be the eigenfunctions of the operator A. They form an orthonormal basis in $L^2(K_R)$ and the expansion of the solution u^i_R on such a basis is

$$u^i_R(t,x) = \sum_{k=0}^{\infty} u^i_k(t)\varphi_k(x), \quad u^i_R(0,x) = \sum_{k=0}^{\infty} g^i_k\varphi_k(x),$$

where $g^i_k := \langle g^i, \varphi_k\rangle_{L^2(K_R)}$ and $u^i_k := \langle u^i_R, \varphi_k\rangle_{L^2(K_R)}$. Plugging this expansion into the equation eq. (3.11), we obtain

$$\sum_{k=0}^{\infty}\left(\frac{d}{dt}u^i_k(t)\varphi_k(x) + u^i_k(t)\lambda_k\varphi_k(x)\right) = 0, \quad \text{for } i = 1,\dots,d,$$

where $\{\lambda_k\}_{k=0}^{\infty}$ are the positive eigenvalues of A. From the orthogonality of the the eigenfunctions basis, we conclude that

$$u^i_k(t) = e^{-\lambda_k t}u^i_k(0) = e^{-\lambda_k t}g^i_k.$$

So the semigroup $e^{-AT} : L^2(K_R) \to L^2(K_R)$ can be defined as

$$e^{-AT}g^i := \sum_{k=0}^{\infty} e^{-\lambda_k T}g^i_k, \tag{3.16}$$

from which it follows that

$$\|e^{-AT}g^i\|^2_{L^2(K_R)} = \sum_{k=0}^{\infty} e^{-2\lambda_k T}|g^i_k|^2 \leq \sum_{k=0}^{\infty}|g^i_k|^2 = \|g^i\|^2_{L^2(K_R)}, \tag{3.17}$$

i.e., the semigroup e^{-AT} is a contraction. By integrating in time, we obtain

$$\chi^i_{T,R}(x) := \int_0^T u^i_R(t,x)\,dt = \sum_{k=0}^{\infty} g^i_k\varphi_i(x)\int_0^T e^{-\lambda_k t}\,dt = \sum_{k=0}^{\infty}\frac{1}{\lambda_k}\left(1-e^{-\lambda_k T}\right)g^i_k\varphi_k(x).$$

Moreover, evaluating $A\chi_{T,R}(x)$ we obtain

$$A\chi^i_{T,R}(x) = \sum_{k=0}^{\infty} g^i_k\varphi_k(x) - \sum_{k=0}^{\infty} e^{-\lambda_k T}g^i_k\varphi^i_k(x) = g^i(x) - e^{-AT}g^i(x).$$

The Lax-Milgram theorem guarantees the existence and uniqueness of $\chi_{T,R}$ in the space $H_0^1(K_R)$. By uniform ellipticity of the coefficients $a(x)$ and Hölder inequality we derive that

$$\alpha\|\nabla\chi^i_{T,R}\|^2_{L^2(K_R)} \leq \|a\mathbf{e}_i\|_{L^2(K_R)}\|\nabla\chi^i_{T,R}\|_{L^2(K_R)} + \|e^{-AT}g^i\|_{L^2(K_R)}\|\chi^i_{T,R}\|_{L^2(K_R)},$$

that, by application of eq. (3.17) and the Poincaré inequality for $\chi_{T,R}$ and Young inequality, leads to the final bound

$$\|\chi^i_{T,R}\|_{H_0^1(K_R)} \leq \sqrt{\frac{1+C_p^2}{\alpha^2}}\|a\mathbf{e}_i\|_{H^1_{div}(K_R)},$$

where α is the ellipticity constant and C_p is the Poincaré constant. □

Remark 3.9. *Note that periodicity of a is not necessary for the well-posedness of $\chi^i_{T,R}$.*

The modified corrector $\chi^i_{T,R}$ plays the same role as the standard corrector χ^i_R, but with a reduced influence from the boundary conditions, thanks to the additional term $e^{-AT}g^i$ at the right-hand side. By recalling eq. (3.10), we approximate the homogenized coefficient by

$$a^{0,R,L,T}_{ij} = \int_{K_L} \mathbf{e}_i \cdot a(x)\left(\mathbf{e}_j + \nabla\chi^j_{T,R}(x)\right)\mu_L(x)\,dx,$$

where μ_L is a filtering function as in Definition 3.7. The goal of Chapter 5 is to prove that the resonance error for the modified elliptic approach decays with an arbitrary rate of convergence, upon choosing $L, T = \mathcal{O}(R)$:

$$e_{MOD} := \left\|a^{0,R,L,T} - a^0\right\|_F \leq C\left[R^{-(q+\frac{1}{2})} + e^{-\sqrt{2\lambda_0}c(1-k_o)R}\right]. \tag{3.18}$$

The constants $q \in \mathbb{N}$, $\lambda_0, c, C \in \mathbb{R}_+$, $0 < k_o < 1$ will be specified later.

3.4 Conclusion

In this chapter we presented two novel corrector models to approximate the macroscopic coefficients for multiscale second order elliptic problems. The former method relies on the solution of parabolic equations to compute the effective tensor, even if the original problem is time-independent, but has the advantage of allowing to achieve arbitrarily high orders of convergence, as proven in Chapter 4. The modified elliptic model is directly derived from the parabolic one and it shows similar convergence properties, see Chapter 5 for the proofs. On the other hand, this approach has the advantage of not requiring to use the additional time dimension.

4 Reduction of the resonance error via parabolic corrector problems

In Chapter 3 we proposed the use of a parabolic model to approximate the homogenized limit of the coefficients $a^\varepsilon(x) = a(x/\varepsilon)$ for the linear second order elliptic PDE:

$$\begin{cases} -\nabla\cdot(a^\varepsilon(x)\nabla u^\varepsilon) = f & \text{in } D\subset\mathbb{R}^d, \\ u^\varepsilon = 0 & \text{on } \partial D, \end{cases}$$

The approximation of a^0 is defined as

$$a_{ij}^{0,R,L,T} = \int_{K_L} a_{ij}(x)\mu_L(x)\,dx - 2\int_0^T\int_{K_L} u_R^i(x,t)\,u_R^j(x,t)\mu_L(x)\,dx\,dt, \tag{4.1}$$

where the parabolic correctors solve

$$\begin{cases} \dfrac{\partial u_R^i}{\partial t} - \nabla\cdot(a(x)\nabla u_R^i) = 0 & \text{in } K_R\times(0,+\infty), \\ u_R^i = 0 & \text{on } \partial K_R\times(0,+\infty), \\ u_R^i(x,0) = \nabla\cdot(a(x)\mathbf{e}_i) & \text{in } K_R. \end{cases} \tag{4.2}$$

In this chapter we carry out the convergence analysis for the presented parabolic approach and we provide *a priori* error bounds on the approximation of the homogenized coefficients given by eq. (4.1). The following assumptions on the multiscale tensor a^ε are taken:

i) $a^\varepsilon(x) = a(x/\varepsilon)$, for $a\in\mathscr{M}(\alpha,\beta)$;

ii) $a(\cdot)$ is K-periodic, with $K := [-1/2,1/2]^d$;

iii) $a(\cdot)\mathbf{e}_i \in H_{div}(K)$, for $i = 1,\dots,d$;

iv) $a(\cdot)\in\left[C^{1,\gamma}(K_R)\right]^{d\times d}$.

By these assumptions,we prove that the resonance error decays with arbitrary convergence rates, thus significantly improving the accuracy of numerical homogenization methods.

Outline

The main contribution of this chapter is the proof of *a priori* error bounds on the resonance error for the parabolic method. The proof is carried out in Section 4.1 and it is based on:

- the use of filters to approximate the average of periodic functions with arbitrary rates of convergence;
- the exponential decay in time of L^2-norm of the parabolic correctors;
- the exponential decay in space of the parabolic Green's function.

As it will be more clear later on, integrating over a bounded time interval is crucial to get arbitrary convergence rates of the resonance error. Some results for too large final integration times T are also described. The theoretical convergence analysis is supported by several numerical experiments, described in Section 4.2. The tests are run also for coefficients that do not comply with the conditions above. Nevertheless, the same convergence trends are found, suggesting that the convergence rates are valid under more general assumptions. As a last step, we discuss the numerical complexity of this approach and compared it against the standard approach in Section 4.3, and we show that the computational cost for the parabolic case grows more slowly than the one for the standard approach, thus making the parabolic method favourable for larger cells.

The content of this chapter is based on [9].

4.1 *A priori* analysis of the resonance error

In this section we show how one can reach high orders of convergence for the resonance error by using the parabolic corrector problems eq. (4.2) and prove the upper bound of eq. (3.18). The main result of the present chapter is the following Theorem 4.1.

Theorem 4.1. *Let $K_R \subset \mathbb{R}^d$, with $d \leq 3$ and $R \geq 1$. Let the coefficient matrix $a(\cdot)$ satisfy:*

i) $a(\cdot) \in \mathscr{M}(\alpha, \beta)$,

ii) $a(\cdot)$ *is K-periodic,*

iii) $a(\cdot)\mathbf{e}_i \in H_{div}(K_R)$, $i = 1, \dots, d$,

iv) $a(\cdot) \in \left[C^{1,\gamma}(K_R)\right]^{d\times d}$ *for some* $0 < \gamma \leq 1$.

Let $a^{0,R,L,T}$ and a^0 be defined, respectively, as in eq. (3.13) *and eq.* (3.4), *with u_R^i satisfying eq.* (3.11) *for any $i = 1, \dots, d$. Let $\mu_L \in \mathbb{K}^q(K_L)$, with $0 < L < R - 3/2$ and $T \leq \frac{2c}{d+1}|R-L|^2$, with*

$c = 1/(4\beta)$. Then, there exists constants $\lambda_0(\alpha, d)$ and $C > 0$ independent of R, L or T (but it may depend on d, $a(\cdot)$ and $\mu_L(\cdot)$) such that

$$\left\| a^{0,R,L,T} - a^0 \right\|_F \leq C \left[L^{-(q+1)} + e^{-2\lambda_0 T} + \frac{1}{T} \left(\frac{R}{\sqrt{T}} \right)^{d-1} e^{-c\frac{|R-L|^2}{T}} + \left(\frac{T}{|R-L|^2} \right)^{3-d} e^{-2c\frac{|R-L|^2}{T}} \right]. \quad (4.3)$$

Additionally, if $\nabla \cdot (a(\cdot)\mathbf{e}_i) \in W^1_{per}(K)$, then there exists a constant $C > 0$ independent of R, L or T (but it may depend on d, $a(\cdot)$ and $\mu_L(\cdot)$) such that

$$\left\| a^{0,R,L,T} - a^0 \right\|_F \leq C \left[L^{-(q+1)} + e^{-2\lambda_0 T} + \frac{1}{|R-L|} \left(\frac{R}{\sqrt{T}} + 1 \right)^{d-1} e^{-c\frac{|R-L|^2}{T}} + \frac{1}{|R-L|^2} \left(\frac{R^2}{T} \right) e^{-2c\frac{|R-L|^2}{T}} \right]. \quad (4.4)$$

The choice

$$L = k_o R, \quad T = k_T R,$$

with $0 < k_o < 1$ and $k_T = \sqrt{\frac{c}{2\lambda_0}}(1 - k_o)$ results in the following convergence rate in terms of R

$$\left\| a^{0,R,L,T} - a^0 \right\|_F \leq C \left[R^{-(q+1)} + e^{-\sqrt{2\lambda_0 c}(1-k_o)R} \right], \quad (4.5)$$

for a constant $C > 0$ independent of R, L or T.

Remark 4.2. *Note that the exponent in the exponential term $\sqrt{2\lambda_0 c} \approx \sqrt{\alpha/\beta}$ depends on the contrast ratio.*

The term $L^{-(q+1)}$ is the averaging error induced by using the filter function $\mu_L \in \mathbb{K}^q(K_L)$, and it can be made arbitrarily small by taking higher values for q. The term $e^{-2\lambda_0 T}$ originates from using a finite T for the parabolic corrector problem (3.11). The remaining terms are the errors due to the boundary conditions, which decay exponentially provided $T < |R-L|^2$. Moreover, the optimal scaling $L \approx R$, $T = \mathcal{O}(R)$ are found by equating the exponents of the truncation and boundary errors. Note that bounds eqs. (4.3) and (4.4) are similar to the one obtained in [78], except for the term T^{-2} that accounts for the effect of using a biased model equation.

4.1.1 Error decomposition

In this section we outline the steps to prove the bound stated in Theorem 4.1:

Step 1: Decomposition the error into four terms:

$$a_{ij}^{0,R,L,T} - a_{ij}^0 = \underbrace{\int_{K_L} a_{ij}(x)\mu_L(x)\,dx - \fint_K a_{ij}(x)\,dx}_{e_{AV}(a_{ij})}$$

$$\underbrace{+2\int_0^T\int_{K_L} v^i(x,t)v^j(x,t)\mu_L(x)\,dx\,dt - 2\int_0^T\int_{K_L} u_R^i(x,t)u_R^j(x,t)\mu_L(x)\,dx\,dt}_{e_{BC}}$$

$$\underbrace{+2\int_0^T \fint_K v^i(x,t)v^j(x,t)\,dx\,dt - 2\int_0^T\int_{K_L} v^i(x,t)v^j(x,t)\mu_L(x)\,dx\,dt}_{e_{AV}(v^iv^j)}$$

$$\underbrace{+2\int_0^{+\infty} \fint_K v^i(x,t)v^j(x,t)\,dx\,dt - 2\int_0^T \fint_K v^i(x,t)v^j(x,t)\,dx\,dt}_{e_{TR}}, \quad (4.6)$$

by exploiting the fact that the exact homogenized coefficient a^0 can be equally calculated by eq. (3.4) or eq. (3.9)

Step 2: Estimation of the *averaging* errors $e_{AV}(a_{ij})$ and $e_{AV}(v^iv^j)$ by means of Lemma 4.3.

Step 3: Estimation of the *truncation* error e_{TR} by means of the exponential decrease in time of $\|v^i(\cdot,t)\|_{L^2(K)}$.

Step 4: Estimation of the *boundary* error e_{BC} by means of upper bounds for the fundamental solution of the parabolic problem and integration over finite time intervals $[0,T]$.

The coming subsections will be devoted to the derivation of upper bounds for $e_{AV}(a_{ij})$, e_{BC}, $e_{AV}(v^iv^j)$ and e_{TR}.

4.1.2 Averaging errors bounds

The two error terms studied in this subsection originate from the fact that we are approximating the averages of periodic functions by a weighted average over a bounded domain. For such a reason, these errors will be referred to as *averaging* error for a and for v^i and are denoted by, respectively, $e_{AV}(a_{ij})$ and $e_{AV}(v^iv^j)$. In order to bound these terms we rely on the fact that filtering functions approximate the average of periodic functions with arbitrary rate of accuracy, as stated in the following lemma (see[78] for a proof).

Lemma 4.3 (Lemma 3.1 in [78])**.** *Let $\mu_L \in \mathbb{K}^q(K_L)$. Then, for any K-periodic function $f \in L^p(K)$ with $1 < p \le 2$, we have*

$$\left|\int_{K_L} f(x)\mu_L(x)\,dx - \fint_K f(x)\,dx\right| \le C\,\|f\|_{L^p(K)}\,L^{-(q+1)},$$

where C is a constant independent of L.

Remark 4.4. *The result of Lemma 4.3 was proved in [78] for K-periodic $f \in L^2(K)$ and, then, extended to the case $f \in L^p(K)$, $1 < p < 2$.*

Corollary 4.5 is a direct consequence of Lemma 4.3, and therefore the proof is omitted.

Corollary 4.5. *Let $a \in \mathcal{M}(\alpha,\beta)$ be K-periodic. Then, there exists $C_1 > 0$, independent of L, such that*

$$\left|e_{AV}(a_{ij})\right| \le C_1 L^{-(q+1)}, \quad i,j = 1,\dots,d.$$

Before providing a convergence result for $e_{AV}(v^i v^j)$ we recall the following property about product rule in Sobolev spaces (see [36] for a proof).

Lemma 4.6. *Let $\Omega \subset \mathbb{R}^d$ be a domain and $u, v \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$, with $1 \le p \le +\infty$. Then, $uv \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ and the product rule for derivation holds:*

$$\frac{\partial}{\partial x_i}(uv) = \frac{\partial u}{\partial x_i} v + u \frac{\partial v}{\partial x_i}, \quad i = 1,\dots,d.$$

Lemma 4.7. *Let $a(\cdot)$ satisfy conditions i), ii) and iii) of Theorem 4.1, let $v^i \in L^2([0,+\infty), W^1_{per}(K))$ be the K-periodic solution of eq.* (3.5) *and $\mu_L \in \mathbb{K}^q(K_L)$. Then, there exists $C_3 > 0$, independent of L, such that*

$$\left|e_{AV}(v^i v^j)\right| \le C_3 L^{-(q+1)}.$$

Proof. By applying Lemma 4.3 to the function $2v^i v^j$ we get:

$$\left|e_{AV}(v^i v^j)\right| \le \int_0^T C \left\| v^i(\cdot,t) v^j(\cdot,t) \right\|_{L^p(K)} L^{-(q+1)}\, dt, \tag{4.7}$$

with $1 < p \le 2$. Following the proof of Lemma 4.3 (see Appendix A, [78]), we deduce that, for any $q \ge 2$ one can also choose $p = 1$ in the inequality above. Therefore, by the use of Cauchy–Schwarz and Hölder inequalities, $e_{AV}(v^i v^j)$ can be estimated as

$$\begin{aligned}
\left|e_{AV}(v^i v^j)\right| &\le \int_0^T C \left\| v^i(\cdot,t) v^j(\cdot,t) \right\|_{L^1(K)} L^{-(q+1)}\, dt \\
&\le C L^{-(q+1)} \int_0^T \left\| v^i(\cdot,t) \right\|_{L^2(K)} \left\| v^j(\cdot,t) \right\|_{L^2(K)} dt \\
&\le C L^{-(q+1)} \left\| v^i \right\|_{L^2([0,+\infty),L^2(K))} \left\| v^j \right\|_{L^2([0,+\infty),L^2(K))}.
\end{aligned}$$

The result follows by choosing

$$C_3 := C \left\| v^i \right\|_{L^2([0,+\infty),L^2(K))} \left\| v^j \right\|_{L^2([0,+\infty),L^2(K))}.$$

In the case $q \in \{0,1\}$ we cannot utilize any more the L^1-norm of the product. In view of eq. (4.7),

with the choice $p = 3/2$, it follows that

$$\begin{aligned}\left|e_{AV}(v^i v^j)\right| &\leq \int_0^T C \left\| v^i(\cdot,t)\, v^j(\cdot,t)\right\|_{L^{3/2}(K)} L^{-(q+1)}\, dt \\ &\leq \int_0^T C \left\| v^i(\cdot,t)\, v^j(\cdot,t)\right\|_{W^{1,1}(K)} L^{-(q+1)}\, dt \\ &\leq \int_0^T C \left\| v^i(\cdot,t)\right\|_{W^1_{per}(K)} \left\| v^j(\cdot,t)\right\|_{W^1_{per}(K)} L^{-(q+1)}\, dt \\ &\leq C L^{-(q+1)} \left\| v^i\right\|_{L^2\left([0,+\infty),W^1_{per}(K)\right)} \left\| v^j\right\|_{L^2([0,+\infty),W^1_{per}(K))},\end{aligned}$$

where the first inequality is a direct application of Lemma 4.3, the second inequality follows from the continuous inclusion of $W^{1,1}(K)$ in $L^{3/2}(K)$, the third inequality comes from the embedding $W^1_{per}(K) \subset W^{1,1}(K)$ and the validity of Lemma 4.6 for functions v^i which implies:

$$\left\| v^i(\cdot,t)\, v^j(\cdot,t)\right\|_{W^{1,1}(K)} \leq C \left\| v^i(\cdot,t)\right\|_{W^1_{per}(K)} \left\| v^j(\cdot,t)\right\|_{W^1_{per}(K)}.$$

Finally, the last inequality is the Chauchy-Schwarz inequality. The result follows by choosing

$$C_3 := C \left\| v^i\right\|_{L^2\left([0,+\infty),W^1_{per}(K)\right)} \left\| v^j\right\|_{L^2\left([0,+\infty),W^1_{per}(K)\right)}.$$

□

4.1.3 Truncation error bound

In this subsection we derive an *a priori* estimate for the truncation error, which originates from the restriction of the time integral in eq. (3.13) on the finite interval $[0, T]$. As it will be more clear from the coming analysis, the time truncation is essential for improving the convergence rate of the resonance error, as large values of T result in a "pollution" of the correctors inside K_L. The spectral properties of the elliptic operator in the space of periodic functions are used to derive an estimate of the truncation error.

For K-periodic coefficients $a \in \mathcal{M}(\alpha,\beta)$, the bilinear form $B : W^1_{per}(K) \times W^1_{per}(K) \mapsto \mathbb{R}$ defined by

$$B(u,v) = \int_K \nabla u \cdot a(x) \nabla v\, dx. \tag{4.8}$$

is continuous and coercive and there exists a non-decreasing sequence of strictly positive eigenvalues $\{\lambda_j\}_{j=0}^{\infty}$ and a L^2-orthonormal set of eigenfunctions $\{\varphi_j\}_{j=0}^{\infty} \subset W^1_{per}(K)$ such that

$$B(\varphi_j, w) = \lambda_j \langle \varphi_j, w\rangle_{L^2(K)}, \quad \forall\, w \in W^1_{per}(K). \tag{4.9}$$

Based on this result, we can prove the following lemma on the exponential decay in time of $\left\| v^i(\cdot,t)\right\|_{L^2(K)}$.

Lemma 4.8. *Let $v^i \in C\left([0,\infty), L^2(K)\right)$ be the solution of eq.* (3.5) *and let $\lambda_0 > 0$ be the smallest eigenvalue of the bilinear form B introduced in eq.* (4.8)*. Then*

$$\left\| v^i(\cdot,t) \right\|_{L^2(K)} \leq e^{-\lambda_0 t} \left\| v^i(\cdot,0) \right\|_{L^2(K)}, \quad \text{a.e. } t \in [0,+\infty).$$

Proof. The variational formulation of eq. (3.5) reads: Find $v^i \in L^2\left([0,+\infty), W^1_{per}(K)\right)$ and $\partial_t v^i \in L^2\left([0,+\infty), W^1_{per}(K)'\right)$ such that

$$\begin{aligned} &\left(\partial_t v^i, w\right) + B\left(v^i, w\right) = 0, \qquad \forall\, w \in W^1_{per}(K), \\ &v^i(\cdot,0) = \nabla\cdot(a\mathbf{e}_i) \in L^2_0(K). \end{aligned}$$

By using $w = v^i(\cdot,t)$, the second line becomes

$$\frac{1}{2}\frac{d}{dt}\left\| v^i \right\|^2_{L^2(K)} = -B\left(v^i, v^i\right).$$

Let $\{\lambda_j\}_{j=0}$ and $\{\varphi_j\}_{j=0}$ be, respectively, the eigenvalues and eigenfunctions of B and let us denote $\hat{v}^i_j := \langle v^i, \varphi_j \rangle_{L^2(K)}$. By orthogonality of the eigenfunctions and Parseval's identity, it holds

$$B\left(v^i, v^i\right) = \sum_{j=0}^{\infty} \lambda_j \left|\hat{v}^i_j\right|^2 \geq \lambda_0 \sum_{j=0}^{\infty} \left|\hat{v}^i_j\right|^2 = \lambda_0 \left\| v^i \right\|^2_{L^2(K)}.$$

Then, by coercivity of the bilinear form B and use of the above inequality, we get

$$\left\| v^i \right\|_{L^2(K)} \frac{d}{dt}\left\| v^i \right\|_{L^2(K)} = \frac{1}{2}\frac{d}{dt}\left\| v^i \right\|^2_{L^2(K)} = -B\left(v^i, v^i\right) \leq -\lambda_0 \left\| v^i \right\|^2_{L^2(K)}.$$

So, the following differential inequality is derived:

$$\frac{d}{dt}\left\| v^i \right\|_{L^2(K)} \leq -\lambda_0 \left\| v^i \right\|_{L^2(K)}.$$

As proved in [67], $\left\| v^i(\cdot,t) \right\|_{L^2(K)}$ is absolutely continuous in time, and the result is obtained by Gronwall's inequality. □

Remark 4.9. *It is easy to prove that $\lambda_0 \geq \frac{\alpha}{C_P^2}$, where the Poincaré constant for a convex domain K is $C_P = \frac{diam(K)}{\pi}$, see [124].*

Lemma 4.10 (Truncation error)**.** *Let $v^i \in C([0,+\infty), L^2(K))$ solve eq.* (3.5)*,*

$$e_{TR} := 2\int_T^{+\infty} \fint_K v^i(x,t)\, v^j(x,t)\, dx\, dt$$

and λ_0 be the smallest eigenvalue of B. Then, there exist $C_4 > 0$, independent of T, such that

$$|e_{TR}| \leq C_4 e^{-2\lambda_0 T}. \tag{4.10}$$

Proof. We start by applying the Cauchy-Schwarz inequality on $L^2(K)$:

$$|e_{TR}| \leq \frac{2}{|K|} \int_T^\infty \left\| v^i(\cdot,t) \right\|_{L^2(K)} \left\| v^j(\cdot,t) \right\|_{L^2(K)} dt. \tag{4.11}$$

Then, we plug the result of lemma 4.8 into eq. (4.11):

$$\begin{aligned} |e_{TR}| &\leq \frac{2}{|K|} \int_T^\infty e^{-2\lambda_0 t} \left\| v^i(\cdot,0) \right\|_{L^2(K)} \left\| v^j(\cdot,0) \right\|_{L^2(K)} dt \\ &\leq \frac{1}{|K|} \left\| v^i(\cdot,0) \right\|_{L^2(K)} \left\| v^j(\cdot,0) \right\|_{L^2(K)} \frac{1}{\lambda_0} e^{-2\lambda_0 T}. \end{aligned}$$

The results follows by choosing

$$\begin{aligned} C_4 &= \frac{1}{\lambda_0 |K|} \left\| v^i(\cdot,0) \right\|_{L^2(K)} \left\| v^j(\cdot,0) \right\|_{L^2(K)} \\ &= \frac{1}{\lambda_0 |K|} \|\nabla \cdot (a(\cdot)\mathbf{e}_i)\|_{L^2(K)} \left\| \nabla \cdot \left(a(\cdot)\mathbf{e}_j\right) \right\|_{L^2(K)}. \end{aligned}$$

□

4.1.4 Boundary error bound

From the definition,

$$e_{BC} := \int_0^T \int_{K_L} \left(u_R^i u_R^j - v^i v^j \right) \mu_L \, dx \, dt, \tag{4.12}$$

one can notice that the source of the error e_{BC} is the mismatch between u_R^i and v^i on the boundary ∂K_R. Therefore, we refer to such an error as the *boundary* error. The boundary error converges to zero at an exponential rate, as stated in Lemma 4.11.

Lemma 4.11. *Let $a(\cdot)$ satisfy conditions i), ii), iii) and iv) of Theorem 4.1, $T \leq \frac{2c}{d+1}|R-L|^2$ and let e_{BC} be defined by eq.* (4.12)*. Then, there exist constants $C, c > 0$, independent of R, L and T such that*

$$|e_{BC}| \leq C \left[\frac{1}{T} \left(\frac{R}{\sqrt{T}} \right)^{d-1} e^{-c\frac{|R-L|^2}{T}} + \left(\frac{T}{|R-L|^2} \right)^{3-d} e^{-2c\frac{|R-L|^2}{T}} \right].$$

Additionally, if $\nabla \cdot (a(\cdot)\mathbf{e}_i) \in W^1_{per}(K)$, then there exist constants $C, c > 0$, independent of R, L and T such that

$$|e_{BC}| \leq C \left[\frac{1}{|R-L|} \left(\frac{R}{\sqrt{T}} + 1 \right)^{d-1} e^{-c\frac{|R-L|^2}{T}} + \frac{1}{|R-L|^2} \left(\frac{R^2}{T} \right) e^{-2c\frac{|R-L|^2}{T}} \right].$$

The proof of Lemma 4.11 directly follows from Propositions 4.16 and 4.18. We need Definitions 4.12 and 4.13 in order to define a *boundary error function* which will be used in the estimation of e_{BC}.

Definition 4.12 (Boundary layer)**.** *Let us define a sub-domain $K_{\tilde{R}} \subset K_R$, where $\tilde{R}$ is defined to*

be the largest integer such that $\tilde{R} \leq R - 1/2$. *The* boundary layer *is defined as the set* $\Delta := K_R \setminus K_{\tilde{R}}$. *We observe that* $|\Delta| = R^d - \tilde{R}^d \leq 2dR^{d-1}$.

The boundary layer and $K_{\tilde{R}}$ are depicted in Figure 4.1.

Definition 4.13 (Cut-off function)**.** *A* cut-off function *on* K_R *is a function* $\rho \in C^\infty(K_R, [0,1])$ *such that*

$$\rho(x) = \begin{cases} 1 \text{ in } K_{\tilde{R}} \\ 0 \text{ on } \partial K_R \end{cases} \quad \text{and} \quad \left|\nabla \rho\right| \leq C \text{ on } \Delta,$$

where the subdomain $K_{\tilde{R}}$ *and the boundary layer* Δ *are defined according to Definition 4.12.*

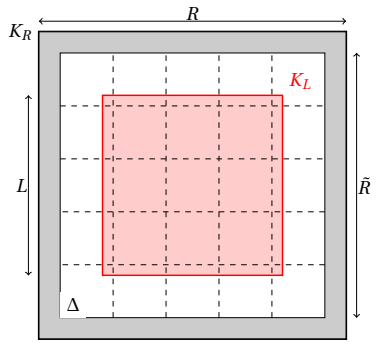


Figure 4.1 – Scheme of the sampling domain K_R and its subsets K_L, $K_{\tilde{R}}$ and Δ.

Let us define the boundary error function $\theta^i \in L^2\left([0,+\infty), H_0^1(K_R)\right)$ through the relation $\theta^i := u_R^i - \rho v^i$. For the analysis it is fundamental that $\rho = 1$ in $K_{\tilde{R}}$ and that $L < \tilde{R}$. By the definition of θ^i, we write

$$e_{BC} = \int_0^T \int_{K_L} \left[v^i v^j \left(\rho^2 - 1\right) + \theta^i v^j + v^i \theta^j + \theta^i \theta^j \right] \mu_L \, dx \, dt.$$

One readily notice that the first term in the integral vanishes on the integration domain, since $\rho^2(x) = 1$ for all $x \in K_{\tilde{R}} \supset K_L$. So, we have to study the integrals

$$e_{BC}^b := \int_0^T \int_{K_L} v^i \theta^j \mu_L \, dx \, dt, \text{ and } e_{BC}^c := \int_0^T \int_{K_L} \theta^i \theta^j \mu_L \, dx \, dt. \tag{4.13}$$

As both integrals depend on the values that the functions θ^i take over the averaging domain K_L, we need to provide pointwise estimates for $\theta^i(x,t)$ on $K_L \times [0,T]$. This is done in Section 4.1.4 by the use of the fundamental solution of eq. (3.11).

Estimates for θ^i

Here, we derive an upper bound for θ^i on $K_L \times [0,T]$. By definition and linearity of the correctors problem, the function θ^i satisfies the problem:

$$\frac{\partial \theta^i}{\partial t} - \nabla \cdot (a(x)\nabla\theta^i) = -\nabla(1-\rho(x)) \cdot a(x)\nabla v^i - \nabla \cdot \left[a(x)\nabla(1-\rho(x))\, v^i \right] \tag{4.14}$$

in $K_R \times (0,+\infty)$, with boundary and initial conditions

$$\begin{aligned} &\theta^i = 0 && \text{on } \partial K_R \times (0,+\infty), \\ &\theta^i(x,0) = v^i(x,0)(1-\rho(x)) && \text{in } K_R. \end{aligned} \tag{4.15}$$

As the integrals in eq. (4.13) are performed over a subset K_L of the domain K_R of eq. (4.14), we are not really interested in estimating the norm of θ^i over the whole K_R, but rather on K_L. Thus, thanks to the use of fundamental solution for problem eqs. (4.14) and (4.15), we will derive *a priori* pointwise estimates for $\theta^i(x,t)$, for $(x,t) \in K_L \times (0,T)$. The legitimacy of pointwise estimates for θ^i is guaranteed by the fact that u^i and v^i are Hölder continuous functions for $t>0$, and so is θ^i. Hence, for $t>0$, the pointwise values of $\theta^i(x,t)$ is meaningful. Moreover, since $\theta^i(x,0)=0$ in K_L, $\theta^i(x,t)$ is bounded in $K_L \times [0,+\infty)$.

Usually, the existence of a fundamental solution for equations like eqs. (4.14) and (4.15) and the derivation of its properties are done for parabolic problems in non-divergence form with Hölder continuous coefficients [71, 105]. In this setting it is possible to prove pointwise bounds (of the type of eq. (4.29)) on the spatial (up to second order) and time (up to first order) derivatives of the fundamental solution. The existence result can be extended to the case of equations in divergence form with discontinuous coefficients, under the only assumption of uniform ellipticity, see [20]. In this weaker setting it is possible to prove the well-known Nash-Aronson estimate on the fundamental solution, but there is no proof, to the best of my knowledge, of the existence of similar bound for the derivative. Therefore, we need to assume $C^{1,\gamma}$-regularity for $a(\cdot)$ in order to be able to write the equation in non-divergence form and use the results of [71, 105].

We will denote by $\Gamma(x,t;\xi,\tau) \in C^{0,\gamma}(K_R \times (\tau,+\infty))$ the fundamental solution of the parabolic operator with homogeneous Dirichlet boundary conditions

$$\begin{aligned} L_{(x,t)}: \quad L^2([\tau,+\infty), H_0^1(K_R)) &\mapsto L^2([\tau,+\infty), H^{-1}(K_R)) \\ u &\mapsto \partial_t u - \nabla_x \cdot (a(x)\nabla_x u), \end{aligned}$$

i.e. $\Gamma(x,t;\xi,\tau)$ satisfies

$$L_{(x,t)}\Gamma(x,t;\xi,\tau) = 0, \quad (x,\xi,t) \in K_R \times K_R \times (\tau,+\infty), \tag{4.16a}$$

$$g(x) = \lim_{t \to \tau^+} \int_{K_R} \Gamma(x,t;\xi,\tau) g(\xi)\, d\xi, \quad \forall g \in C(K_R). \tag{4.16b}$$

Subscript (x,t) in eq. (4.16a) is to indicate that the differentiation is operated with respect to the x- and t-variables. Equation eq. (4.16b) can be interpreted as the fact that the initial condition (given that the initial time instant is $t=\tau$) for the fundamental solution is $\Gamma(x,\tau;\xi,\tau)=\delta(x-\xi)$, the Dirac's delta function centred at ξ. In the same way, one can define the adjoint operator, given the symmetry of a, as

$$\begin{aligned} L^*_{(y,s)}: \quad L^2((-\infty,\tau],H^1_0(K_R)) &\mapsto L^2((-\infty,\tau],H^{-1}(K_R)) \\ u &\mapsto -\partial_s u-\nabla_y\cdot\big(a(y)\nabla_y u\big). \end{aligned}$$

The fundamental solution of $L^*_{(y,s)}$ is denoted by $\Gamma^*(y,s;x,t)$ and satisfies

$$\begin{aligned} L^*_{(y,s)}\Gamma^*(y,s;\xi,\tau)=0, \quad (y,\xi,s)\in K_R\times K_R\times(-\infty,\tau), \\ g(y)=\lim_{s\to\tau^-}\int_{K_R}\Gamma^*(y,s;\xi,\tau)g(\xi)\,d\xi, \quad \forall g\in C(K_R). \end{aligned}$$

A well-known result is that the differential problems

$$L_{(x,t)}u=f \quad \text{and} \quad L^*_{(y,s)}v=f$$

are well-posed only for $t>\tau$ and $s<\tau$, respectively, where τ is the time of the initial (resp. final) condition. Thus, we formally define

$$\Gamma(x,t;\xi,\tau)=0, \text{ for } t<\tau, \quad \Gamma^*(y,s;\xi,\tau)=0, \text{ for } s>\tau.$$

A central property of the two fundamental solutions is

$$\Gamma(x,t;y,s)=\Gamma^*(y,s;x,t), \text{ for } s<t. \tag{4.17}$$

The identity between two fundamental solution is proved in Theorem 17, §3.7 [71] for the case of Hölder continuous coefficients, but it can be extended to the discontinuous case by following the same proof, as done in [22]. Pointwise *a priori* estimates for Γ are derived in [21], following the results obtained in [121]. Such estimates can be extended to the derivatives of the fundamental solution under additional regularity assumptions, see, e.g., [71, 105]. The solution of eq. (4.14) can be written as

$$\begin{aligned} \theta^i(x,t)=&\int_{K_R}\Gamma(x,t;y,0)v^i(y,0)(1-\rho(y))\,dy \\ &-\int_{K_R}\int_0^t\Gamma(x,t;y,s)\nabla_y(1-\rho(y))\cdot a(y)\nabla_y v^i(y,s)\,ds\,dy \\ &+\int_{K_R}\int_0^t\nabla_y\Gamma(x,t;y,s)\cdot a(y)\nabla_y(1-\rho(y))v^i(y,s)\,ds\,dy, \end{aligned} \tag{4.18}$$

for any $t>0$. Now, we provide a lemma for rewriting eq. (4.18) in the form of boundary flux integral.

Lemma 4.14. *Let $a(\cdot)$ satisfy conditions i), ii) and iii) of Theorem 4.1, θ^i be the weak solution of eq.* (4.14) *and let v^i be Hölder continuous in $K_R \times (0,+\infty)$. Then, for any $(x,t) \in K_L \times (0,+\infty)$,*

$$\theta^i(x,t) = \int_{\partial K_R} \int_0^t \mathbf{n} \cdot a(y) \nabla_y \Gamma(x,t;y,s)\, v^i(y,s)\, ds\, d\sigma_y, \tag{4.19}$$

where $\mathbf{n}$ *denotes the unit vector orthogonal to ∂K_R pointing outward.*

Proof. First of all, we derive an integral equality for Γ^*. Multiplying $L^*_{(y,s)}\Gamma^* = 0$ by $v^i(1-\rho)$, integrating over $K_R \times (0,t)$ and using integration by parts, one gets:

$$\begin{aligned}
\int_0^t \int_{K_R} & -\partial_s \Gamma^*(y,s;x,t)\, v^i(y,s)(1-\rho(y)) \\
& + \nabla_y \Big(v^i(y,s)(1-\rho(y)) \Big) \cdot a(y) \nabla_y \Gamma^*(y,s;x,t)\, dy\, ds \\
& = \int_0^t \int_{\partial K_R} \mathbf{n} \cdot a(y) \nabla_y \Gamma(x,t;y,s)\, v^i(y,s)(1-\rho(y))\, d\sigma_y\, ds,
\end{aligned} \tag{4.20}$$

since $\nabla_y \Gamma(x,t;y,s) = \nabla_y \Gamma^*(y,s;x,t)$ for any $s < t$. Then the second and third integrals in eq. (4.18) are rewritten as

$$\begin{aligned}
\int_{K_R} \int_0^t & -\Gamma(x,t;y,s) \nabla_y (1-\rho(y)) \cdot a(y) \nabla_y v^i(y,s)\, ds\, dy \\
& + \int_{K_R} \int_0^t \nabla_y \Gamma(x,t;y,s) \cdot a(y) \nabla_y (1-\rho(y))\, v^i(y,s)\, ds\, dy \\
& = \int_{K_R} \int_0^t -\nabla_y \left[\Gamma(x,t;y,s)(1-\rho(y)) \right] \cdot a(y) \nabla_y v^i(y,s)\, ds\, dy \\
& + \int_{K_R} \int_0^t \nabla_y \Gamma(x,t;y,s) \cdot a(y) \nabla_y \Big[(1-\rho(y))\, v^i(y,s) \Big]\, ds\, dy \\
& = \int_{K_R} \int_0^t \Gamma(x,t;y,s)(1-\rho(y)) \partial_s v^i(y,s)\, ds\, dy \\
& + \int_{K_R} \int_0^t \nabla_y \Gamma(x,t;y,s) \cdot a(y) \nabla_y \Big[(1-\rho(y))\, v^i(y,s) \Big]\, ds\, dy,
\end{aligned} \tag{4.21}$$

where the last equality follows from the weak form of eq. (3.5). Then, we integrate the former of the two last integrals by parts, thus obtaining

$$\begin{aligned}
\int_{K_R} \int_0^t & \Gamma(x,t;y,s)(1-\rho(y)) \partial_s v^i(y,s)\, ds\, dy \\
& = \lim_{\epsilon \to 0^+} \int_{K_R} \Gamma(x,t;y,t-\epsilon) v^i(y,t-\epsilon)(1-\rho(y))\, dy \\
& \quad - \int_{K_R} \Gamma(x,t;y,0) v^i(y,0)(1-\rho(y))\, dy
\end{aligned}$$

$$-\int_{K_R}\int_0^t \partial_s\Gamma(x,t;y,s)(1-\rho(y))\,v^i(y,s)\,ds\,dy. \quad (4.22)$$

From the fact that $\rho(x)=1$ for all $x\in K_L$ and from the continuity of v^i we deduce

$$\lim_{\epsilon\to 0^+}\int_{K_R}\Gamma(x,t;y,t-\epsilon)\,v^i(y,t-\epsilon)(1-\rho(y))\,dy = v^i(x,t)(1-\rho(x))=0,$$

for any $x\in K_L$. By putting eqs. (4.18), (4.21) and (4.22) together we get

$$\theta^i(x,t)=\int_{K_R}\int_0^t -\partial_s\Gamma(x,t;y,s)\,v^i(y,s)(1-\rho(y))\,ds\,dy + \int_{K_R}\int_0^t \nabla_y\Gamma(x,t;y,s)\cdot a(y)\nabla_y\left[v^i(y,s)(1-\rho(y))\right]\,ds\,dy.$$

Finally, from eqs. (4.17) and (4.20) we conclude that

$$\theta^i(x,t)=\int_{\partial K_R}\int_0^t \mathbf{n}\cdot a(y)\nabla_y\Gamma(x,t;y,s)\,v^i(y,s)\,ds\,d\sigma_y.$$

□

From now on we will distinguish two cases in the derivation of the estimates, based on the regularity of the initial condition $v^i(\cdot,0)=\nabla\cdot(a(\cdot)\mathbf{e}_i)$, i.e. on the regularity of the tensor $a(\cdot)$.

Lemma 4.15. *Let $a(\cdot)$ satisfy conditions i), ii), iii) and iv) of Theorem 4.1*[1]*, let $\theta^i\in C([0,+\infty),L^2(K_R))$ be the solution of eq.* (4.14)*, and let $v^i\in L^2((0,+\infty),W^1_{per}(K))$ be the solution of eq.* (3.5)*. Then, there exist a constant $\tilde{C}>0$, independent of R and L such that*

$$\sup_{x\in K_L}\left|\theta^i(x,t)\right|\le \tilde{C}\frac{R^{d-1}}{|R-L|}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}\left[\frac{1}{t}+\frac{1}{2c\,|R-L|^2}\right]^{\frac{d-1}{2}} e^{-c\frac{|R-L|^2}{t}}, \quad (4.23)$$

for $c=1/4\beta$.

Otherwise, if $v^i\in C\left([0,+\infty)W^1_{per}(K)\right)$, then

$$\sup_{x\in K_L}\left|\theta^i(x,t)\right|\le \tilde{C}R^{d-1}\left\|v^i(\cdot,0)\right\|_{W^1_{per}(K)} e^{-\lambda_0 t}\int_0^t \frac{1}{s^{(d+1)/2}}e^{-c\frac{|R-L|^2}{s}}e^{\lambda_0 s}\,ds, \quad (4.24)$$

where $\lambda_0>0$ is the smallest eigenvalue of the bilinear form B.

Proof. From eq. (4.19) we can write

$$\left|\theta^i(x,t)\right|\le\int_0^t\int_{\partial K_R}\left|\mathbf{n}\cdot a(y)\nabla_y\Gamma(x,t;y,s)\right|\left|v^i(y,s)\right|\,d\sigma_y\,ds.$$

[1] The assumption of Hölder continuity of $\partial_k a_{ij}(x)$ is to ensure the correctness of (4.29).

By applying the Hölder inequality we get

$$\left|\theta^i(x,t)\right| \leq |\partial K_R|^{1/2} \int_0^t \sup_{y\in\partial K_R} \left|\mathbf{n}\cdot a(y)\nabla_y\Gamma(x,t;y,s)\right| \left\|v^i(\cdot,s)\right\|_{L^2(\partial K_R)} ds. \tag{4.25}$$

The value of $\left\|v^i(\cdot,s)\right\|_{L^2(\partial K_R)}$ is well defined for any time $s>0$ (unless we have a more regular initial condition, e.g. $v^i(\cdot,0)\in W^1_{per}(K)$, in that case the trace is defined also for $s=0$) and we can estimate it by the following inequality

$$\left\|v^i(\cdot,s)\right\|_{L^2(\partial K_R)} = \left\|v^i(\cdot,s)(1-\rho)\right\|_{L^2(\partial K_R)} \leq C_{tr}\left\|v^i(\cdot,s)(1-\rho)\right\|_{H^1(\Delta)},$$

where C_{tr} is fixed, thanks to the fact that the distance between $K_{\tilde{R}}$ and ∂K_R is larger or equal to $1/2$. As $\rho\in C^1(K_R)$ and $\partial_{x_k} v^i(\cdot,s)\in L^2(K_R)$ the product rule holds and we can write

$$\left\|\nabla(v^i(1-\rho))\right\|_{L^2(\Delta)} \leq \left\|\nabla v^i\right\|_{L^2(\Delta)} + \left\|\nabla\rho\right\|_{L^\infty(\Delta)}\left\|v^i\right\|_{L^2(\Delta)}.$$

Let us now consider a covering of Δ, defined as $\Delta_K := \bigcup_{y\in\partial K_{\frac{R+\tilde{R}}{2}}} K+y$. Then, $|\Delta_K| = c(d)\left|\partial K_{\frac{R+\tilde{R}}{2}}\right|\operatorname{diam}(K) \leq CR^{d-1}$. By exploiting the periodic structure of v^i we have that

$$\begin{aligned}
\left\|v^i\right\|_{L^2(\Delta)} &\leq \left\|v^i\right\|_{L^2(\Delta_K)} \leq \left(\frac{|\Delta_K|}{|K|}\right)^{1/2}\left\|v^i\right\|_{L^2(K)},\\
\left\|\nabla v^i\right\|_{L^2(\Delta)} &\leq \left\|\nabla v^i\right\|_{L^2(\Delta_K)} \leq \left(\frac{|\Delta_K|}{|K|}\right)^{1/2}\left\|\nabla v^i\right\|_{L^2(K)}.
\end{aligned}$$

Finally, we recall that in the space $W^1_{per}(K)$ the Poincaré-Wirtinger inequality holds:

$$\left\|v^i\right\|_{L^2(K)} \leq C_P\left\|\nabla v^i\right\|_{L^2(K)} \tag{4.26}$$

so that

$$\begin{aligned}
\left\|v^i(\cdot,s)\right\|_{L^2(\partial K_R)} &\leq C_{tr}C_\rho C_P\left(\frac{|\Delta|}{|K|}\right)^{1/2}\left\|\nabla v^i(\cdot,s)\right\|_{L^2(K)}\\
&\leq CR^{\frac{d-1}{2}}\left\|\nabla v^i(\cdot,s)\right\|_{L^2(K)}.
\end{aligned} \tag{4.27}$$

Now, we go back to the estimation of θ^i: putting together eqs. (4.25) and (4.27) (and recalling that $|\partial K_R| = 2dR^{d-1}$) we get

$$\left|\theta^i(x,t)\right| \leq CR^{d-1}\int_0^t \sup_{y\in\partial K_R}\left|\mathbf{n}\cdot a(y)\nabla_y\Gamma(x,t;y,s)\right|\left\|\nabla v^i(\cdot,s)\right\|_{L^2(K)} ds. \tag{4.28}$$

Now, we will derive different *a priori* estimates for different regularity assumption on the initial condition. Both of them rely on the Nash-Aronson type estimate

$$\nabla_y\Gamma(x,t;y,s) \leq \frac{C}{(t-s)^{\frac{d+1}{2}}}e^{-c\frac{|x-y|^2}{t-s}}, \tag{4.29}$$

with $C = (4\pi\alpha)^{-d/2}$ and $c = (4\beta)^{-1}$. The bound eq. (4.29) is proved in [71, 105] for parabolic equations in non-divergence form with Hölder continuous coefficients. In [68] the authors claim that eq. (4.29) is valid also for parabolic equation in divergence form with Hölder continuous coefficients, but the statement remains unproved.
Case $v^i(\cdot,0) \in L^2(K)$: We apply the Hölder inequality in time and the estimates on $\nabla_y\Gamma$ for Hölder coefficients to get:

$$\begin{aligned}\left|\theta^i(x,t)\right| &\le CR^{d-1}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}\left(\int_0^t \sup_{y\in\partial K_R}\left|\mathbf{n}\cdot a(y)\nabla_y\Gamma(x,t;y,s)\right|^2 ds\right)^{1/2}\\ &\le CR^{d-1}\|a\|_{L^\infty(K)}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}\left(\int_0^t \frac{C^2}{(t-s)^{(d+1)}}e^{-2c\frac{|x-\bar y(x)|^2}{t-s}} ds\right)^{1/2}, \quad (4.30)\end{aligned}$$

where $\bar y(x) = \arg\min_{y\in\partial K_R}\left|x-y\right|$. By the change of variables $\sigma = 2c\frac{|x-\bar y(x)|^2}{t-s}$ and the fact that the primitive function of $t^N e^{-t}$ (with $N\in\mathbb{N}$) is $-\sum_{k=0}^N \frac{N!}{k!}t^k e^{-t}$, the inequality eq. (4.30) becomes

$$\begin{aligned}\left|\theta^i(x,t)\right| &\le C\frac{\|a\|_{L^\infty(K)}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}R^{d-1}}{(2c\left|x-\bar y(x)\right|^2)^{d/2}}\\ &\quad\left[\sum_{k=0}^{d-1}\frac{(d-1)!}{k!}\left(2c\frac{\left|x-\bar y(x)\right|^2}{t}\right)^k\right]^{\frac12} e^{-c\frac{|x-\bar y(x)|^2}{t}}\\ &\le \frac{C}{\sqrt{2c}}\|a\|_{L^\infty(K)}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}\frac{R^{d-1}}{\left|x-\bar y(x)\right|}\\ &\quad\left[(d-1)!\sum_{k=0}^{d-1}\binom{d-1}{k}\frac{1}{t^k}\left(\frac{1}{2c\left|x-\bar y(x)\right|^2}\right)^{d-1-k}\right]^{\frac12} e^{-c\frac{|x-\bar y(x)|^2}{t}}\\ &\le \frac{C(d-1)!}{\sqrt{2c}}\|a\|_{L^\infty(K)}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}\frac{R^{d-1}}{\left|x-\bar y(x)\right|}\\ &\quad\left[\frac1t+\frac{1}{2c\left|x-\bar y(x)\right|^2}\right]^{\frac{d-1}{2}} e^{-c\frac{|x-\bar y(x)|^2}{t}}.\end{aligned}$$

Including all the terms that do not depend on R, L nor t in a single constant $\tilde C$ and by the lower bound $\inf_{x\in K_L}\left|x-\bar y(x)\right| \ge |R-L|$ we deduce

$$\left|\theta^i(x,t)\right| \le \tilde C\frac{R^{d-1}}{|R-L|}\left\|\nabla v^i\right\|_{L^2((0,t),L^2(K))}\left[\frac1t+\frac{1}{2c|R-L|^2}\right]^{\frac{d-1}{2}} e^{-c\frac{|R-L|^2}{t}}.$$

Case $v^i(\cdot,0) \in W^1_{per}(K)$: Again, we use the eigenvalues $\{\lambda_j\}_{j=0}$ and eigenvectors $\{\varphi_j\}_{j=0}$ of B. Let us denote $\hat v^i_j(t) := \langle v^i(\cdot,t),\varphi_j\rangle_{L^2(K)}$. Then,

$$\hat v^i_j(t) = e^{-\lambda_j t}\langle v^i(\cdot,0),\varphi_j\rangle_{L^2(K)}.$$

From the above characterization of the components $\hat{v}^i_j(t)$ and the coercivity of B we have

$$\alpha \left\| \nabla v^i(\cdot,t) \right\|^2_{L^2(K)} \leq B[v^i(\cdot,t), v^i(\cdot,t)] = \sum_{j=0}^{+\infty} e^{-2\lambda_j t} \lambda_j \left| \langle v^i(\cdot,0), \varphi_j \rangle_{L^2(K)} \right|^2,$$

for any $t \geq 0$. The Parseval's identity also holds for $t = 0$, since $v^i(\cdot,0) \in W^1_{per}(K)$, by assumption. So,

$$\begin{aligned} \alpha \left\| \nabla v^i(\cdot,t) \right\|^2_{L^2(K)} &\leq e^{-2\lambda_0 t} \sum_{j=0}^{+\infty} \lambda_j \left| \langle v^i(\cdot,0), \varphi_j \rangle_{L^2(K)} \right|^2 \\ &= e^{-2\lambda_0 t} B[v^i(\cdot,0), v^i(\cdot,0)] \\ &\leq \beta e^{-2\lambda_0 t} \left\| \nabla v^i(\cdot,0) \right\|^2_{L^2(K)}. \end{aligned}$$

Thus,

$$\left\| \nabla v^i(\cdot,t) \right\|_{L^2(K)} \leq e^{-\lambda_0 t} \left(\frac{\beta}{\alpha} \right)^{1/2} \left\| \nabla v^i(\cdot,0) \right\|_{L^2(K)}. \tag{4.31}$$

Then, we apply again the known inequality for $\nabla_y \Gamma$ and the estimate in eq. (4.28) becomes

$$\begin{aligned} \left| \theta^i(x,t) \right| &\leq R^{d-1} \frac{\beta^{3/2}}{\alpha^{1/2}} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} \int_0^t \frac{C}{(t-s)^{(d+1)/2}} e^{-c\frac{|x-\bar{y}(x)|^2}{t-s}} e^{-\lambda_0 s} \, ds \\ &= R^{d-1} \frac{\beta^{3/2}}{\alpha^{1/2}} e^{-\lambda_0 t} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} \int_0^t \frac{C}{s^{(d+1)/2}} e^{-c\frac{|x-\bar{y}(x)|^2}{s}} e^{\lambda_0 s} \, ds, \end{aligned}$$

and we get eq. (4.24) by posing $\tilde{C} = \frac{C\beta^{3/2}}{\alpha^{1/2}}$ and re-using the lower bound

$$\inf_{x \in K_L} \left| x - \bar{y}(x) \right| \geq |R - L|.$$

□

Term e^b_{BC}

Proposition 4.16. *Let the hypotheses of Lemma 4.15 be satisfied. Moreover, let $v^i \in C([0,+\infty), L^2(K))$, $\theta^i \in L^\infty(K_L \times [0,+\infty))$, let e^b_{BC} be defined as in eq.* (4.13) *and let L/R be constant. Then, there exist constants $C_{2,b}, C'_{2,b}, c > 0$ independent of R, L, T such that*

$$\left| e^b_{BC} \right| \leq \frac{C_{2,b}}{|R-L|} \left\| v^i \right\|_{L^2([0,+\infty), W^1_{per}(K))} \left(\frac{R}{\sqrt{T}} + C'_{2,b} \right)^{d-1} e^{-c\frac{|R-L|^2}{T}}, \tag{4.32}$$

Otherwise, if $v^i(\cdot,0) \in W^1_{per}(K)$ and $T \leq \frac{2c}{d+1} |R-L|^2$ then there exist constants $C_{2,b}, c > 0$ independent of R, L, T such that

$$\left| e^b_{BC} \right| \leq \frac{C_{2,b}}{T} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} \left(\frac{R}{\sqrt{T}} \right)^{d-1} e^{-c\frac{|R-L|^2}{T}}. \tag{4.33}$$

Proof. Applying Hölder inequality on the space integral, we obtain:

$$\left|\int_0^T \int_{K_L} v^i(x,t)\theta^j(x,t)\mu_L(x)\,dx\,dt\right| \leq \int_0^T \int_{K_L} \left|v^i(x,t)\theta^j(x,t)\mu_L(x)\right| dx\,dt$$
$$\leq \int_0^T \left\|v^i(\cdot,t)\right\|_{L^2(K_L)} \left\|\theta^j(\cdot,t)\right\|_{L^\infty(K_L)} \left\|\mu_L\right\|_{L^2(K_L)} dt.$$

By assumption, $\mu_L \in L^\infty(K_L) \subset L^2(K_L)$ with continuous inclusion, and

$$\left\|\mu_L\right\|_{L^2(K_L)} \leq |K_L|^{1/2} \left\|\mu_L\right\|_{L^\infty(K_L)} \leq C_\mu L^{-d/2}.$$

Next, we estimate $\left\|v^i(\cdot,t)\right\|_{L^2(K_L)}$. Since v^i, we have for integer L

$$\left\|v^i(\cdot,t)\right\|_{L^2(K_L)} = L^{d/2} \left\|v^i(\cdot,t)\right\|_{L^2(K)},$$

while, for non-integer L

$$\left\|v^i(\cdot,t)\right\|_{L^2(K_L)} \leq \lceil L \rceil^{d/2} \left\|v^i(\cdot,t)\right\|_{L^2(K)}.$$

Finally, we recall the exponential decay of $\left\|v^i(\cdot,t)\right\|_{L^2(K)}$ and we derive the estimate:

$$\left|\int_0^T \int_{K_L} v^i(x,t)\theta^j(x,t)\mu_L(x)\,dx\,dt\right| \leq$$
$$\leq L^{d/2} \left\|v^i(\cdot,0)\right\|_{L^2(K)} \int_0^T e^{-\lambda_0 t} \left\|\theta^j(\cdot,t)\right\|_{L^\infty(K_L)} dt\, C_\mu L^{-d/2}$$
$$\leq C_\mu \left\|v^i(\cdot,0)\right\|_{L^2(K)} \int_0^T e^{-\lambda_0 t} \left\|\theta^j(\cdot,t)\right\|_{L^\infty(K_L)} dt. \quad (4.34)$$

Case $v^i(\cdot,0) \in L^2(K)$: We use eq. (4.23) in lemma 4.15 to bound the last integral in eq. (4.34):

$$\int_0^T e^{-\lambda_0 t} \left\|\theta^j(\cdot,t)\right\|_{L^\infty(K_L)} dt \leq$$
$$\leq \tilde{C} \frac{R^{d-1}}{|R-L|} \left\|v^i\right\|_{L^2([0,+\infty),W^1_{per}(K))} \int_0^T e^{-\lambda_0 t} \left[\frac{1}{t} + \frac{1}{2c|R-L|^2}\right]^{\frac{d-1}{2}} e^{-c\frac{|R-L|^2}{t}} dt$$
$$\leq \frac{\tilde{C}}{\lambda_0} \frac{R^{d-1}}{|R-L|} \left\|v^i\right\|_{L^2([0,+\infty),W^1_{per}(K))} \left[\frac{1}{T} + \frac{1}{2c|R-L|^2}\right]^{\frac{d-1}{2}} e^{-c\frac{|R-L|^2}{T}}$$
$$= \frac{\tilde{C}}{\lambda_0} \left\|v^i\right\|_{L^2([0,+\infty),W^1_{per}(K))} \frac{1}{|R-L|} \left[\frac{R^2}{T} + \frac{R^2}{2c|R-L|^2}\right]^{\frac{d-1}{2}} e^{-c\frac{|R-L|^2}{T}},$$

where we bounded the integral by the $L^1 - L^\infty$ Hölder inequality. Then, by posing

$$C_{2,b} = \frac{C_\mu \tilde{C}}{\lambda_0} \left\|v^i(\cdot,0)\right\|_{L^2(K)}, \text{ and } C'_{2,b} = \frac{1}{\sqrt{2c}(1-L/R)}, \text{ with } 0 < L/R < 1,$$

we get eq. (4.32).

Case $v^i(\cdot,0) \in W^1_{per}(K)$: We can use the estimate eq. (4.24) to bound the last integral in eq. (4.34):

$$\begin{aligned}
&\int_0^T e^{-\lambda_0 t} \left\| \theta^j(\cdot,t) \right\|_{L^\infty(K_L)} dt \\
&\quad \leq \tilde{C} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} R^{d-1} \int_0^T e^{-2\lambda_0 t} \int_0^t s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{\lambda_0 s} \, ds \, dt \\
&\quad = \tilde{C} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} R^{d-1} \int_0^T \int_s^T e^{-2\lambda_0 t} \, dt \, s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{\lambda_0 s} \, ds,
\end{aligned} \tag{4.35}$$

by Fubini's theorem. We bound the double integral in time as

$$\begin{aligned}
\int_0^T \int_s^T e^{-2\lambda_0 t} \, dt \, s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{\lambda_0 s} \, ds &\leq \frac{1}{2\lambda_0} \int_0^T s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{-\lambda_0 s} \, ds \\
&\leq \frac{1}{2\lambda_0} \left(\max_{s\in[0,T]} s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} \right) \int_0^T e^{-\lambda_0 s} \, ds \\
&\leq \frac{1}{2\lambda_0^2} T^{-(d+1)/2} e^{-c\frac{|R-L|^2}{T}},
\end{aligned}$$

under the assumption that $T \leq \frac{2c}{d+1} |R-L|^2$. Thus we get the final bound

$$\int_0^T e^{-\lambda_0 t} \left\| \theta^j(\cdot,t) \right\|_{L^\infty(K_L)} dt \leq \frac{\tilde{C}}{2\lambda_0^2} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} \left(\frac{R}{\sqrt{T}} \right)^{d-1} \frac{1}{T} e^{-c\frac{|R-L|^2}{T}},$$

and the proof is complete by taking

$$C_{2,b} = \frac{C_\mu \tilde{C}}{2\lambda_0^2} \left\| v^i(\cdot,0) \right\|_{L^2(K)}.$$

□

Remark 4.17. *The estimates provided in Proposition 4.16 for regular initial condition are subjected to the final time constraint $T \leq \frac{2c}{d+1} |R-L|^2$. If such a condition is not satisfied, then the convergence rate of the resonance error is deteriorated as the solution is polluted by the boundary error for longer times. The analysis for this case is postponed to Section 4.1.6.*

Term e^c_{BC}

Here, we provide estimates for the term e^c_{BC} of eq. (4.13) under two regularity conditions. This term decays faster than e^b_{BC} and can be neglected.

Proposition 4.18. *Let the hypotheses of Lemma 4.15 be satisfied. Moreover, let $v^i \in C([0,+\infty), L^2(K))$, $\theta^i \in L^\infty(K_L \times [0,+\infty))$, let e^c_{BC} be defined as in eq.* (4.13) *and let L/R be constant. Then, there*

exist a constants $C_{2,c}, c > 0$ *independent of* R, L, T *such that*

$$\left|e_{BC}^{c}\right| \leq \frac{C_{2,c}}{|R-L|^2} \left\| v^i \right\|^2_{L^2([0,+\infty),W^1_{per}(K))} \left(\frac{R^2}{T}\right)^{d-1} e^{-\frac{2c|R-L|^2}{T}}. \tag{4.36}$$

Otherwise, if $v^i \in C([0,+\infty), W^1_{per}(K))$*, then, there exist constants* $C_{2,c}, c > 0$ *independent of* R, L, T *such that*

$$\left|e_{BC}^{c}\right| \leq C_{2,c} \left\| v^i(\cdot,0) \right\|^2_{W^1_{per}(K)} \left(\frac{T}{c|R-L|^2}\right)^{3-d} e^{-2c\frac{|R-L|^2}{T}}. \tag{4.37}$$

Proof. From the positivity of μ_L and the fact that its integral is equal to one, we derive the inequality

$$\begin{aligned}\left|\int_0^T \int_{K_L} \theta^i(x,t)\theta^j(x,t)\mu_L(x)\,dx\,dt\right| &\leq \int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\theta^j(x,t)\right| dt \int_{K_L} \mu_L(x)\,dx \\ &\leq \max_i \int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt.\end{aligned}$$

Then, the task now is to estimate $\int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt$.

<u>*Case* $v^i(\cdot,0) \in L^2(K)$*:*</u> By eq. (4.23) we derive

$$\begin{aligned}\int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt \leq \tilde{C}^2 \frac{R^{2(d-1)}}{|R-L|^2} \left\| v^i \right\|^2_{L^2([0,+\infty),W^1_{per}(K))} \\ \int_0^T \left[\frac{1}{t} + \frac{1}{2c|R-L|^2}\right]^{d-1} e^{-2c\frac{|R-L|^2}{t}}\,dt.\end{aligned} \tag{4.38}$$

By the change of variable $\sigma = 2c\frac{|R-L|^2}{t}$ we bound the integral

$$\begin{aligned}&\int_0^T \left[\frac{1}{t} + \frac{1}{2c|R-L|^2}\right]^{d-1} e^{-2c\frac{|R-L|^2}{t}}\,dt \\ &\qquad = \left(\frac{1}{2c|R-L|^2}\right)^{d-2} \int_{\frac{2c|R-L|^2}{T}}^{+\infty} \frac{(\sigma+1)^{d-1}}{\sigma^2} e^{-\sigma}\,d\sigma \\ &\qquad \leq \left(\frac{1}{2c|R-L|^2}\right)^{d-2} \left(\frac{2c|R-L|^2}{T}+1\right)^{d-1} \left(\frac{2c|R-L|^2}{T}\right)^{-2} \int_{\frac{2c|R-L|^2}{T}}^{+\infty} e^{-\sigma}\,d\sigma \\ &\qquad = \left(\frac{1}{2c|R-L|^2}\right)^{d-2} \left(1+\frac{T}{2c|R-L|^2}\right)^{d-1} \left(\frac{T}{2c|R-L|^2}\right)^{3-d} e^{-\frac{2c|R-L|^2}{T}} \\ &\qquad \leq \frac{C}{T^{d-1}} e^{-\frac{2c|R-L|^2}{T}},\end{aligned} \tag{4.39}$$

since $\left(1+\frac{T}{2c|R-L|^2}\right)$ and $\frac{T^2}{2c|R-L|^2}$ can be bounded from above by a constant, due to $T \leq C|R-L|$. By plugging eq. (4.39) into eq. (4.38) we get:

$$\int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt$$
$$\leq \tilde{C}^2 \left\| v^i \right\|^2_{L^2([0,+\infty),W^1_{per}(K))} \left(\frac{R^2}{2c|R-L|^2}\right)^{d-1} \frac{1}{T^{d-1}} e^{-\frac{2c|R-L|^2}{T}}$$
$$\leq \tilde{C}^2 \left\| v^i \right\|^2_{L^2([0,+\infty),W^1_{per}(K))} \left(\frac{R^2}{T}\right)^{d-1} \frac{1}{2c|R-L|^2} e^{-\frac{2c|R-L|^2}{T}}.$$

We get eq. (4.36) with $C_{2,c} = \frac{\tilde{C}^2}{2c}$.
Case $v^i(\cdot,0) \in W^1_{per}(K)$: We recall eq. (4.24) and apply Minkowski integral inequality:

$$\int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt \leq \tilde{C}^2 R^{2(d-1)} \left\| v^i(\cdot,0) \right\|^2_{W^1_{per}(K)}$$
$$\int_0^T \left(e^{-\lambda_0 t} \int_0^t s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{\lambda_0 s}\, ds \right)^2 dt$$
$$\leq \tilde{C}^2 R^{2(d-1)} \left\| v^i(\cdot,0) \right\|^2_{W^1_{per}(K)}$$
$$\left\{ \int_0^T \left(\int_s^T e^{-2\lambda_0 t} s^{-(d+1)} e^{-2c\frac{|R-L|^2}{s}} e^{2\lambda_0 s}\, dt \right)^{1/2} ds \right\}^2$$
$$\leq \tilde{C}^2 R^{2(d-1)} \left\| v^i(\cdot,0) \right\|^2_{W^1_{per}(K)}$$
$$\left\{ \int_0^T \frac{1}{\sqrt{2\lambda_0}} \left(e^{-2\lambda_0 s} - e^{-2\lambda_0 T} \right)^{1/2} s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{\lambda_0 s}\, ds \right\}^2$$
$$\leq \frac{\tilde{C}^2}{2\lambda_0} R^{2(d-1)} \left\| v^i(\cdot,0) \right\|^2_{W^1_{per}(K)}$$
$$\leq \frac{\tilde{C}^2}{2\lambda_0} R^{2(d-1)} \left\| v^i(\cdot,0) \right\|^2_{W^1_{per}(K)} \left\{ \int_0^T s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}}\, ds \right\}^2,$$

by the fact that $\left(1 - e^{-2\lambda_0(T-s)}\right) \leq 1$. We estimate the integral by the change of variables $\sigma = c\frac{|R-L|^2}{s}$:

$$\int_0^T s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}}\, ds = \left(\frac{1}{c|R-L|^2}\right)^{\frac{d-1}{2}} \int_{c\frac{|R-L|^2}{T}}^{+\infty} \sigma^{(d-3)/2} e^{-\sigma}\, d\sigma$$
$$\leq \left(\frac{1}{c|R-L|^2}\right)^{\frac{d-1}{2}} \left(\sup_{\sigma \geq c\frac{|R-L|^2}{T}} \sigma^{(d-3)/2} \right) \int_{c\frac{|R-L|^2}{T}}^{+\infty} e^{-\sigma}\, d\sigma$$
$$\leq \frac{1}{c|R-L|^2} T^{(3-d)/2} e^{-c\frac{|R-L|^2}{T}}.$$

And by plugging the bound for the integral into the bound for $\int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt$ we get

$$\int_0^T \sup_{x\in K_L} \left|\theta^i(x,t)\right|^2 dt \le \frac{\tilde{C}^2}{2\lambda_0} R^{2(d-1)} \left\|v^i(\cdot,0)\right\|^2_{W^1_{per}(K)} \frac{1}{c^2|R-L|^4} T^{3-d} e^{-2c\frac{|R-L|^2}{T}}$$
$$\le \frac{\tilde{C}^2}{2\lambda_0} \left\|v^i(\cdot,0)\right\|^2_{W^1_{per}(K)} \left(\frac{T}{c|R-L|^2}\right)^{3-d} \left(\frac{1}{c(1-L/R)}\right)^{2(d-1)} e^{-2c\frac{|R-L|^2}{T}}.$$

since $\frac{1}{c(1-L/R)}$ is constant, we get eq. (4.37) with $C_{2,c} = \frac{\tilde{C}^2}{2\lambda_0}\left(\frac{1}{c(1-L/R)}\right)^{2(d-1)}$. □

4.1.5 *A priori* bound on the resonance error for the parabolic approach

We can now prove Theorem 4.1, by recalling the results of Sections 4.1.2 to 4.1.4.

Proof of Theorem 4.1. The decomposition eq. (4.6) implies

$$\left\|a^{0,R,L,T} - a^0\right\|_F \le d^2 \max_{i,j}\left(\left|I^1_{ij}\right| + \left|I^2_{ij}\right| + \left|I^3_{ij}\right| + \left|I^4_{ij}\right|\right).$$

By using the upper bounds in Corollary 4.5, Lemmas 4.7 and 4.10, and Propositions 4.16 and 4.18 in the above inequality we get

$$\left\|a^{0,R,L,T} - a^0\right\|_F \le C\left[L^{-(q+1)} + e^{-2\lambda_0 T} + \frac{1}{|R-L|}\left(\frac{R}{\sqrt{T}}+1\right)^{d-1} e^{-c\frac{|R-L|^2}{T}} \right.$$
$$\left. + \frac{1}{|R-L|^2}\left(\frac{R^2}{T}\right)^{d-1} e^{-2c\frac{|R-L|^2}{T}}\right], \quad (4.40)$$

for some constant C independent of R, L and T. Using the optimal values $L = k_o R$ and $T = k_T R$, with $0 < k_o < 1$ and $k_T = \sqrt{\frac{c}{2\lambda_0}}(1-k_o)$, we write eq. (4.40) as:

$$\left\|a^{0,R,L,T} - a^0\right\|_F \le C\left[R^{-(q+1)} + e^{-\sqrt{2\lambda_0 c}(1-k_o)R} + \frac{1}{R}\left(\sqrt{R}+1\right)^{d-1} e^{-\sqrt{2\lambda_0 c}(1-k_o)R} \right.$$
$$\left. + R^{d-3} e^{-2\sqrt{2\lambda_0 c}(1-k_o)R}\right],$$

The last term is of higher order than the third one, so it can be omitted. Finally, we get

$$\left\|a^{0,R,L,T} - a^0\right\|_F \le C\left[R^{-(q+1)} + \left(1 + \frac{\left(\sqrt{R}+1\right)^{d-1}}{R}\right) e^{-\sqrt{2\lambda_0 c}(1-k_o)R}\right]. \quad (4.41)$$

In the case of more regular initial conditions, $\nabla\cdot(a\mathbf{e}_i) \in W^1_{per}(K)$, we have:

$$\left\|a^{0,R,L,T} - a^0\right\|_F \le C\left[L^{-(q+1)} + e^{-2\lambda_0 T} + \left(\frac{R}{\sqrt{T}}\right)^{d-1} \frac{1}{T} e^{-c\frac{|R-L|^2}{T}} + \left(\frac{T}{|R-L|^2}\right)^{3-d} e^{-2c\frac{|R-L|^2}{T}}\right].$$

Also in this case, we use $L = k_o R$ and $T = k_T R$ and omit the last term to get

$$\left\| a^{0,R,L,T} - a^0 \right\|_F \leq C \left[R^{-(q+1)} + \left(1 + R^{\frac{d-3}{2}}\right) e^{-\sqrt{2\lambda_0 c}(1-k_o)R} \right]. \tag{4.42}$$

Finally, using the fact that $R \geq 1$, we can bound the prefactors in front of the exponential terms in eq. (4.41) and eq. (4.42) by a constant independent of R and get eq. (4.5). □

4.1.6 Effect of integration over long time

Since the estimates provided in Lemma 4.11 for regular initial condition are derived under the constraint of $T < \frac{2c}{d+1}|R-L|^2$, we provide here a short analysis for the case where such a condition is not satisfied.

Proposition 4.19. *Let $v^i \in C\left([0,+\infty), W^1_{per}(K)\right)$ be the weak solution of the parabolic model or its periodic extension on K_R let $\theta^i \in L^\infty([0,+\infty) \times K_R)$ be the solution of eq.* (4.14) *and let L/R be constant. Then, there exist constants $\tilde{C}_{2,b}, c > 0$ independent of R, L, T such that, for any $\frac{2c}{d+1}|R-L|^2 \leq T < +\infty$,*

$$\left| e^b_{BC} \right| \leq \frac{\tilde{C}_{2,b}}{|R-L|^2} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)}.$$

Proof. From the proof of Lemma 4.11 we know that

$$\begin{aligned}\left| \int_0^T \int_{K_L} v^i(x,t)\theta^j(x,t)\mu_L(x)\,dx\,dt \right| \leq & C_\mu \left\| v^i(\cdot,0) \right\|_{L^2(K)} \frac{\tilde{C}}{2\lambda_0} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)} R^{d-1} \\ & \int_0^T s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{-\lambda_0 s}\,ds.\end{aligned}$$

We bound the last integral by $L^1 - L^\infty$ Hölder inequality:

$$\begin{aligned}\int_0^T s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} e^{-\lambda_0 s}\,ds &\leq \sup_{s\in[0,+\infty)} s^{-(d+1)/2} e^{-c\frac{|R-L|^2}{s}} \int_0^{+\infty} e^{-\lambda_0 s}\,ds \\ &\leq \frac{1}{\lambda_0}\left(\frac{d+1}{2ce}\right)^{\frac{d+1}{2}} \frac{1}{|R-L|^{d+1}}.\end{aligned}$$

Thus, by putting all the results together, we get

$$\left| \int_0^T \int_{K_L} v^i(x,t)\theta^j(x,t)\mu_L(x)\,dx\,dt \right| \leq \frac{\tilde{C}_{2,b}}{|R-L|^2} \left\| v^i(\cdot,0) \right\|_{W^1_{per}(K)},$$

where

$$\tilde{C}_{2,b} = \frac{C_\mu \tilde{C}}{2\lambda_0^2}\left(\frac{d+1}{2ce}\right)^{\frac{d+1}{2}} \left(\frac{1}{(1-L/R)}\right)^{d-1} \left\| v^i(\cdot,0) \right\|_{L^2(K)}.$$

□

Remark 4.20. *Proposition 4.19 states that if T is chosen too large, then the convergence rate of*

the resonance error is only second order. Such a behaviour was noticed in numerical experiments and was particularly evident in the one-dimensional simulations, see Section 4.2.4.

4.2 Numerical experiments

In this section we present several numerical tests to support the theoretical results of Section 4.1 and experimentally verify the resonance error bound of Theorem 4.1. We illustrate the expected convergence rates by varying the regularity parameter q of the filters, in a periodic, smooth setting, as rigorously proven in the previous sections. Additionally, we compare the convergence rate of the resonance error for the parabolic scheme with that of standard numerical homogenization scheme. We also test non-smooth periodic and stochastic coefficients, which violate the theoretical assumptions in the analysis. Nevertheless, we obtain results as in the smooth periodic case.

In order to numerically assess the convergence rate of the resonance error, we compute the approximations of the homogenized tensor through the described parabolic corrector problems on domains of increasing size, $R \in [1,20]$, and calculate the Frobenius norm of the difference between such approximations and the exact a^0. In the case of periodic coefficients whose homogenized value could not be known exactly (i.e., without discretization error) the reference value is computed by solving the standard elliptic corrector problem eq. (3.3) with $R = 1$ and periodic boundary conditions and using formula eq. (3.4). In the random setting no approximation is available without some resonance error. In this case, we take as reference value for the homogenized tensor the one computed from the numerical approximation of the parabolic correctors over the largest domain $R_{max} = 20$.

To compute a numerical approximation of $a^{0,R,L,T}$, we use a Finite Elements (FE) discretization for the corrector problems eq. (3.11) in space, and a stabilised explicit Runge-Kutta method with adaptive time stepping for the time discretization. A high (fourth) order method, [1], is chosen in order to make the temporal discretization error negligible with respect to the resonance error. As we use explicit methods in time, we need a mass matrix that is cheap to invert. This is achieve by using either mass lumping (for low order FEMs) or discontinuous Galerkin methods (for arbitrary order FEMs).

As a second step, the upscaled tensor is approximated by a double integration in space and time. The spatial integral of the parabolic correctors is computed by using the FE filtered mass matrix of components

$$m_{ij} = \int_{K_L} \phi_i(x)\phi_j(x)\mu_L(x)\,dx,$$

where $\{\phi_i\}_i$ are the FE basis functions. The integration in time is performed by the use of Newton-Cotes formulae for non-uniform discretizations.

In order to optimize the convergence rate of the error with respect to the sampling domain size R, we take the optimal values of Theorem 4.1 for the averaging domain size L ($K_L \subset K_R$)

and for the final time T given by

$$L = k_o R, \text{ and } T = \frac{R-L}{\sqrt{8\beta\lambda_0}},$$

where β is the continuity constant of the tensor a and λ_0 is the smallest eigenvalue of the elliptic operator $-\nabla\cdot(a(\cdot)\nabla)$ with periodic boundary conditions. The oversampling ratio, $0 < k_o < 1$, and the order of filters, q, can be chosen freely.

4.2.1 Two-dimensional periodic case

We consider the upscaling of the 2×2 isotropic tensor:

$$a(x) = \left(\frac{2+1.8\sin(2\pi x_1)}{2+1.8\cos(2\pi x_2)} + \frac{2+\sin(2\pi x_2)}{2+1.8\cos(2\pi x_1)}\right) I \tag{4.43}$$

for which the homogenized tensor is

$$a^0 \approx \begin{pmatrix} 2.757 & -0.002 \\ -0.002 & 3.425 \end{pmatrix}.$$

Here, we compare the performances of the described parabolic approach ("par." in the legends) and the standard elliptic approach ("ell." in the legends). In comparing the two methods, we used a filtered version of (2.45), namely

$$a_{ij}^{0,R,L} := \int_{K_L} \mathbf{e}_i \cdot a(x)\left(\mathbf{e}_j + \nabla\chi_R^j(x)\right)\mu_L(x)\,dx, \tag{4.44}$$

that improves the error constant for the classical approach. However, we recall that the standard elliptic method provides a first order convergence rate, independently of the use of oversampling or filtering, as shown in [134]. By contrast, the use of high order filters in the parabolic scheme improves the convergence rate without affecting the computational cost. The two approaches are solved using $\mathbb{P}_1$ finite element discretization in space with 64 points per periodic cell. Mass lumping has been used in order to perform the time integration, which is carried out via the ROCK4 method, see [1], with $tol = 10^{-6}$. Finally Simpson's quadrature rule is used for computing the time integral defining homogenized coefficients.

Results are depicted in Figure 4.2. As expected, one cannot reach a convergence rate higher than 1 for the standard elliptic approach, in contrast to the parabolic method. We notice a longer "flat" region in the convergence plot for small values of k_o and high order filters. Intuitively, for any given R, the region where the filter is "not almost zero" decreases for smaller k_o and larger q. Hence, we need larger values of R for the averaging integral to contain enough data and the error to decrease with the expected rate.

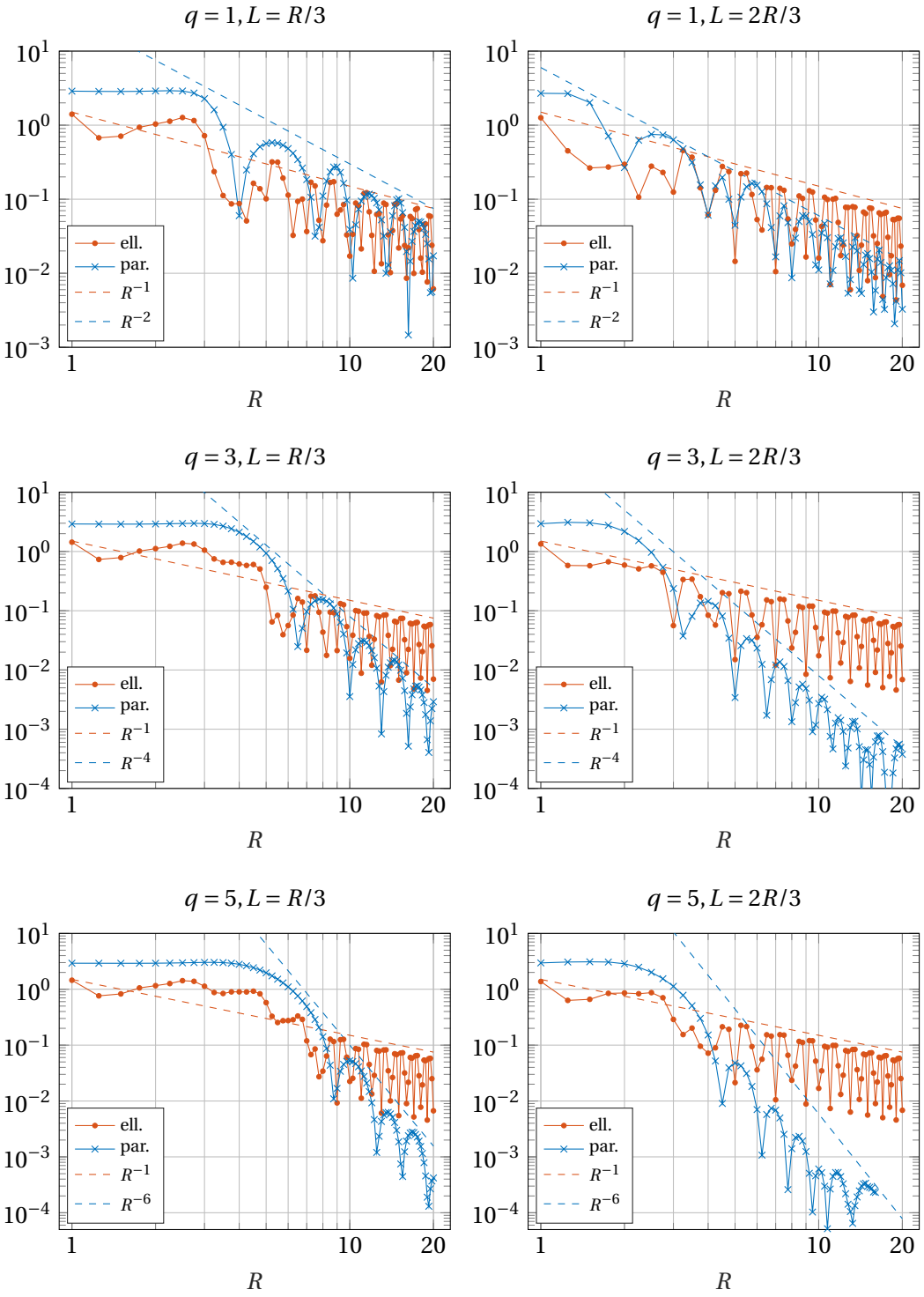


Figure 4.2 – Comparison of the resonance error in the elliptic and parabolic models for tensor eq. (4.43).

4.2.2 Discontinuous coefficients

In the error analysis, we made the assumption that the initial condition $\nabla \cdot (a(\cdot)\mathbf{e}_i) \in L^2(K_R)$. Nevertheless, the parabolic problem can also be solved for initial condition $\nabla \cdot (a(\cdot)\mathbf{e}_i) \in H^{-1}(K_R)$ and we are interested in verifying numerically if the provided *a priori* estimates for the resonance error hold also for this case. For simplicity, we consider the one dimensional periodic piecewise continuous coefficient

$$a(x) = \begin{cases} 1 & \frac{1}{4} < \{x\} < \frac{3}{4}, \\ 3 & \text{elsewhere}, \end{cases} \tag{4.45}$$

where $\{x\}$ is the fractional part of x, i.e. $\{x\} = x - \lfloor x \rfloor$. The homogenized coefficient, which can be computed analytically, is $a^0 = \frac{3}{2}$. Convergence plots pictured in Figure 4.3 show that the theoretical results also apply to the case of discontinuous coefficients. The test is done with $\mathbb{P}_2$ finite element discretization on a uniform grid of size $h = 1/1024$ and the ROCK4 time integration scheme with $tol = 10^{-6}$. The results are reported in Figure 4.3 where, for the sake of completeness, we also pictured the convergence plot for the elliptic scheme without filtering nor oversampling (this simplifying choice is motivated from the fact that filtering and oversampling have been proved to be ineffective for improving the convergence rate in the elliptic case, see Section 4.2.1). Also in this case, if the filter's order q is increased or the oversampling ratio k_o is decreased, the expected convergence rate will reached for larger values of R.

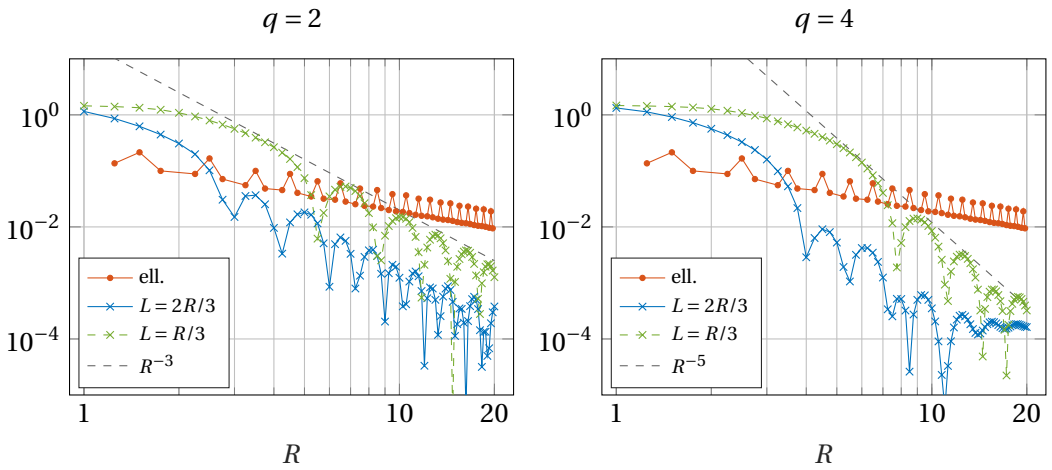
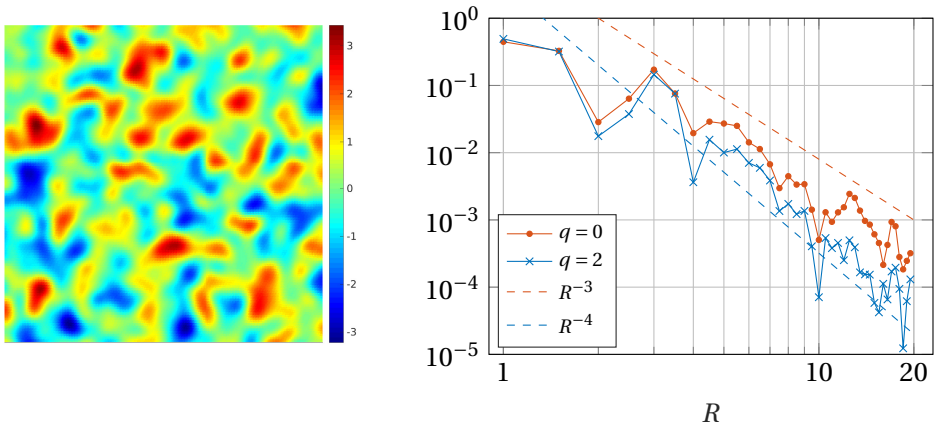


Figure 4.3 – Resonance error in the elliptic and parabolic models for the discontinuous tensor eq. (4.45). The elliptic approximation to a^0 is computed without filtering nor oversampling.



(a) Realization of the field on the square. The colour scale is logarithmic.

(b) Resonance error. $L = 2R/3$.

Figure 4.4 – Log-normal random field eq. (4.46) with $\mu = 0$, $\sigma^2 = 1$ and $\ell = 0.2$, and resonance error for the parabolic ell problem with filter order q and final time $T = \frac{|R-L|}{10}$.

4.2.3 A stochastic case

In the last numerical test, we provide an example for a stochastic tensor, which does not comply with the periodicity assumption made in Section 4.1. With this test, we do not aim at proving any theoretical convergence rate of the error, but rather to verify numerically that the periodicity assumption is not necessary for achieving fast decaying rates of the boundary error. We consider a single realization of a stationary log-normal random field with Gaussian isotropic covariance:

$$\log a(\cdot) \sim \mathcal{N}(\mu, Cov(x-y)), \quad Cov(z) = \sigma^2 e^{-\frac{|z|^2}{2\ell^2}}, \tag{4.46}$$

where μ and σ^2 are the mean and the variance of the field and ℓ is the correlation length. An example of such a field is depicted in Figure 4.4a. We are not interested in evaluating the statistical error, but only the boundary error, which is

$$\left\| a^{0,R,L,T} - a^{0,\infty,L,T} \right\|_F.$$

In practice, we will consider $a^{0,R_{max},L,T}$ for the large value $R_{max} = 20$ in place of $a^{0,\infty,L,T}$ as a reference for evaluating the resonance error. The new reference $a^{0,R_{max},L,T}$ is computed using the numerical approximation of the parabolic corrector on $K_{R_{max}}$ with periodic BCs. The test is done with a $\mathbb{P}_1$ finite element discretization on a uniform grid of size $h = 1/20$ and the ROCK4 time integration scheme with $tol = 10^{-5}$. In Figure 4.4b we show that the resonance error decays with a rate comprised between 3 and 4 with respect to R.

4.2.4 Numerical tests for long integration time

The aim of this section is to briefly illustrate the numerical results obtained for final integration time $T = 100$, so that it does not satisfy the condition $T \leq \frac{2c}{d+1}|R-L|^2$. It was proved in Section 4.1.6 that the error terms $I_{ij}^{2,b}$ scale as R^{-2}, thus overriding all other terms and limiting the convergence rate to 2. Since the error term $I_{ij}^{2,b}$ affects only the correction part of the G-limit approximation,

$$a_{corr}^{0,R,L,T} = -2\int_0^T \int_{K_L} u_R^i(x,t)\, u_R^j(x,t)\mu_L(x)\, dx\, dt,$$

we display the error committed in approximating this term:

$$e_{corr,ij} := \left| \fint_K \mathbf{e}_i \cdot a(x)\nabla\chi^j\, dx + 2\int_0^T \int_{K_L} u_R^i(x,t)\, u_R^j(x,t)\mu_L(x)\, dx\, dt \right|.$$

We tested both a one dimensional case,

$$a(y) = \frac{1}{2+\sin(2\pi y)}, \tag{4.47}$$

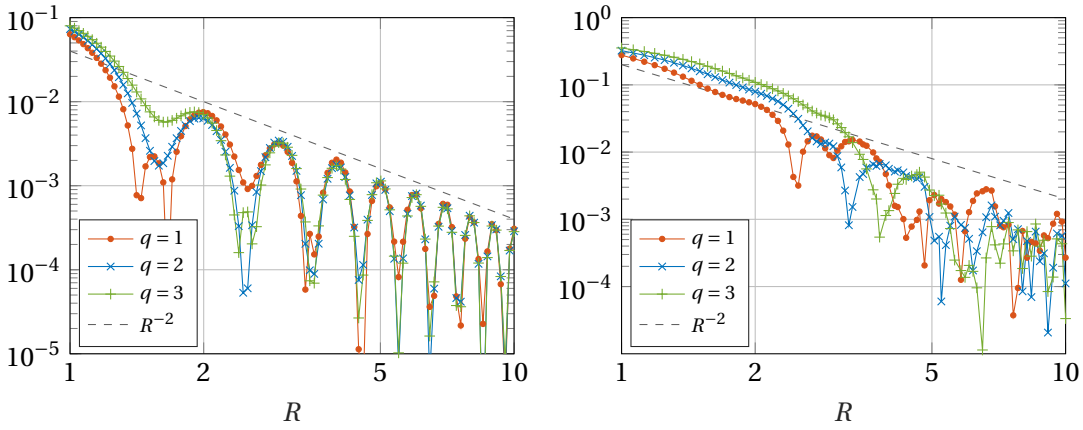
and a two dimensional case

$$a(y) = \begin{pmatrix} \left(3+\frac{2\sqrt{17}}{8\sin(2\pi y_1)+9}\right)^{-1} & 0 \\ 0 & \left(\frac{1}{20}+\frac{2\sqrt{17}}{8\cos(2\pi y_2)+9}\right)^{-1} \end{pmatrix}, \tag{4.48}$$

for R ranging from 1 to 10, $L = 2R/3$, $T = 100$ and with different filter orders q. The second order convergence rate is mostly clear in the one dimensional example depicted in Figure 4.5a, but also the convergence plot for the two dimensional case of Figure 4.5b decays more slowly than the plots of Figure 4.2.

4.3 Discussion over the computational cost

The goal of this section is to provide a theoretical estimate of the scaling of the computational cost with respect to the error tolerance for the proposed parabolic approach and to compare it to the standard elliptic approach. Since both discretization and resonance parameters play a role in the determination of the computational cost, in our analysis we will assume that both errors are smaller than a prescribed tolerance and we derive the computational cost under these constraints. Our analysis shows that, for sufficiently high order filters, the computational cost is lower for the parabolic model than for the elliptic one, i.e. the parabolic case is asymptotically less expensive.



(a) 1D case eq. (4.47)

(b) 2D case eq. (4.48)

Figure 4.5 – Error in the correction term of the homogenized tensor.

4.3.1 Standard elliptic case

Let us consider the standard elliptic homogenization scheme on the rescaled sampling domain of eq. (2.43). We partition the domain K_R with uniform simplicial elements of size h and we introduce a finite elements space $S_h \subset H^1_0(K_R)$ made of piecewise polynomial functions of degree s on the simplices. The finite elements discretization of the corrector problem reads: Find $\chi^i_{R,h} \in S_h$ such that

$$\int_{K_R} a(x)\left(\nabla \chi^i_{R,h} + \mathbf{e}_i\right) \cdot \nabla w_h \, dx = 0, \quad \forall w_h \in S_h, \quad i = 1, \dots, d, \tag{4.49}$$

and the upscaled tensor is defined as

$$a^{0,R,h}_{ij} = \fint_{K_R} \mathbf{e}_i \cdot a(x)\left(\nabla \chi^j_{R,h} + \mathbf{e}_j\right) dx. \tag{4.50}$$

Hence, the total error for the upscaled coefficients is:

$$|a^{0,R,h}_{ij} - a^0_{ij}| \leq C\left(h^{2s} + R^{-1}\right),$$

where the first term in the error estimate is the discretization error derived in [4], while the second term is the resonance error. The finite elements corrector $\chi^i_{R,h}$ is computed by solving the linear system

$$A_h \mathbf{v}_i = \mathbf{b}_i, \text{ for } i = 1, \dots, d, \tag{4.51}$$

where A_h is a $N \times N$ symmetric positive definite matrix and $\mathbf{v}_i$ and $\mathbf{b}_i$ are the coordinates of, respectively, $\chi^i_{R,h}$ and $-\nabla \cdot (a(x)\mathbf{e}_i)$ in the finite element space given a Lagrangian basis. Here, $N = \mathcal{O}(R^d h^{-d})$ is the dimension of the space S_h. The linear system can be solved in several ways using direct or iterative methods, whose cost depends on N. For example, for sparse

LU factorization the number of operations is[2] $\mathcal{O}(N^{3/2})$ [74], for Conjugate Gradient (CG) it is $\mathcal{O}(\sqrt{\kappa}N)$, where κ is the condition number, while for multigrid (MG) it is $\mathcal{O}(N)$, [125]. In the following analysis we will assume that the latter method is used for solving the linear system. We require the total errors to scale as a given tolerance tol, so $R = \mathcal{O}(tol^{-1})$ and $h = \mathcal{O}(tol^{1/2s})$. Hence, the total cost is

$$Cost = \mathcal{O}(N) = \mathcal{O}(R^d h^{-d}) = \mathcal{O}(tol^{-d-\frac{d}{2s}}).$$

4.3.2 Parabolic case with explicit stabilised time integration methods

Let us consider the parabolic corrector problem eq. (3.11) with the upscaling formula eq. (3.13). As in the elliptic case, one can discretize eq. (3.11) in space and compute an approximation $u_h^i(t)$ of $u_R^i(\cdot,t)$ in the N-dimensional finite elements space S_h. For simplicity of notation, we will omit the superscript i. For a given basis of S_h, the function $u_h(t)$ is uniquely determined by the vectorial function $\mathbf{w}_h : [0,T] \mapsto \mathbb{R}^N$, that solve the semi-discrete problem:

$$\frac{d}{dt}\mathbf{w}_h = -M_h^{-1} A_h \mathbf{w}_h. \tag{4.52}$$

We assume that the mass matrix M_h is easy to invert (which hold, e.g., in the case of mass lumping or discontinuous Galerkin FEs), so that the cost of the right-hand side evaluation is negligible with respect to the solution of the ODE system. The differential equation (4.52) is solved by an explicit stabilised time integration scheme of order r. Examples of second order methods are RKC2 [131] and ROCK2 [13], while ROCK4 [1] is a fourth order method. The fully discrete problem reads

$$\mathbf{W}_k = \Phi_h(\mathbf{W}_{k-1}), \text{ for } k = 1,\dots,N_t,$$

where the function Φ_h identifies the time integration method and N_t the number of time steps. The computed sequence $\{\mathbf{W}_k\}_{k=0}^{N_t} \subset \mathbb{R}^N$ is an approximation, at times $t_k = k\Delta t$, of $\mathbf{w}(t_k)$ and it determines (via the finite elements basis) a sequence $\{U_k\}_{k=0}^{N_t} \subset S_h$. The discrete approximation of the homogenized tensor is

$$a_{ij}^{0,R,h,\Delta t} = \int_{K_L} a_{ij}(y)\mu_L(x)\,dx - 2\mathcal{Q}\left(\int_{K_L} U_k U_k^j \mu_L(y)\,dx, \Delta t\right),$$

where $\mathcal{Q}(\cdot,\Delta t)$ is a quadrature rule on the discretization $t_k = k\Delta t$ of order at least r (where r is the order of the time integration scheme). Hence, the total error for the upscaled coefficients is:

$$|a_{ij}^{0,R,h,\Delta t} - a_{ij}^0| \le C\left(h^{s+1} + \Delta t^r + R^{-(q+1)}\right), \tag{4.53}$$

where we have assumed that, for sufficiently large R, the term $R^{-(q+1)}$ dominates the exponential term in the resonance error bound. This is also the convergence rate that we reported in the numerical examples of Sections 4.2.1 and 4.2.2. Here, the constant C grows linearly with

[2] The constant in this asymptotic rate depends on the sparsity pattern of the matrix, which is much worse for 3D problems than for diffusion problems in 2D.

Corrector problem	Parabolic	Standard Elliptic
Error	$R^{-q-1}+h^{s+1}+\Delta t^r$	$R^{-1}+h^{2s}$
Computational cost	$R^d h^{-d-1}\Delta t^{-\frac{1}{2}}$	$R^d h^{-d}$
Computational cost (tol)	$tol^{-\frac{d}{q+1}-\frac{d+1}{s+1}-\frac{1}{2r}}$	$tol^{-d-\frac{d}{2s}}$

Table 4.1 – Error and computational cost for the parabolic and standard homogenization approaches.

the final time T, whose optimal value scales as $R-L$. However, the ratio $(R-L)/\sqrt{8\beta\lambda_0}$ is in general $\mathcal{O}(1)$, so we can consider $T=\mathcal{O}(1)$ in the range of values used for R and L. In order for the error to scale as tol, we require that all the three summands in (4.53) scale as tol:

$$R=\mathcal{O}(tol^{-\frac{1}{q+1}}),\quad h=\mathcal{O}(tol^{\frac{1}{s+1}}),\quad \Delta t=\mathcal{O}(tol^{\frac{1}{r}}).$$

The global computational cost is $\mathcal{O}(Nn_SN_t)$, where $N_t=T/\Delta t$ is the number of time steps, n_S is the number of function evaluations (stages) per time step for a stabilised method and $N=\mathcal{O}(R^dh^{-d})$ is the cost of each function evaluation which, in the linear case, is the cost of multiplying a sparse $N\times N$ matrix by a vector in $\mathbb{R}^N$. Since we are using a stabilised method we need to satisfy the weak stability condition $\rho\Delta t=cn_S^2$, where ρ is the spectral radius of the Jacobian of the ODE (4.52) and n_S is the number of stages for each time step. As ρ is the spectral radius of $M_h^{-1}A_h$, it scales as h^{-2}. Therefore, $n_S=\mathcal{O}(\Delta t^{1/2}h^{-1})$. From the fact that $T=\mathcal{O}(1)$ one derives that the total cost is

$$Cost=\mathcal{O}(R^dh^{-d}\Delta t^{1/2}h^{-1}\Delta t^{-1})=\mathcal{O}(tol^{-\frac{d}{q+1}-\frac{d+1}{s+1}-\frac{1}{2r}}).$$

4.3.3 Comparison of the parabolic and the standard elliptic methods

Now, we are interested in evaluating under which condition the use of stabilised time integration methods is more efficient than the regularized elliptic approach. In Table 4.1, we summarize the dependency of computational cost and the error on resonance and discretization parameters, as well as the scaling of the cost for a given tolerance. In order for the parabolic approach to be competitive with respect to the elliptic one, the condition to satisfy is:

$$\frac{d}{q+1}+\frac{d+1}{s+1}+\frac{1}{2r}<d+\frac{d}{2s}.$$

In Figure 4.6 we display the theoretical increase of the computational cost for the two considered approaches. We observe that, for high order filters, the elliptic model is much more expensive than the parabolic corrector problem.

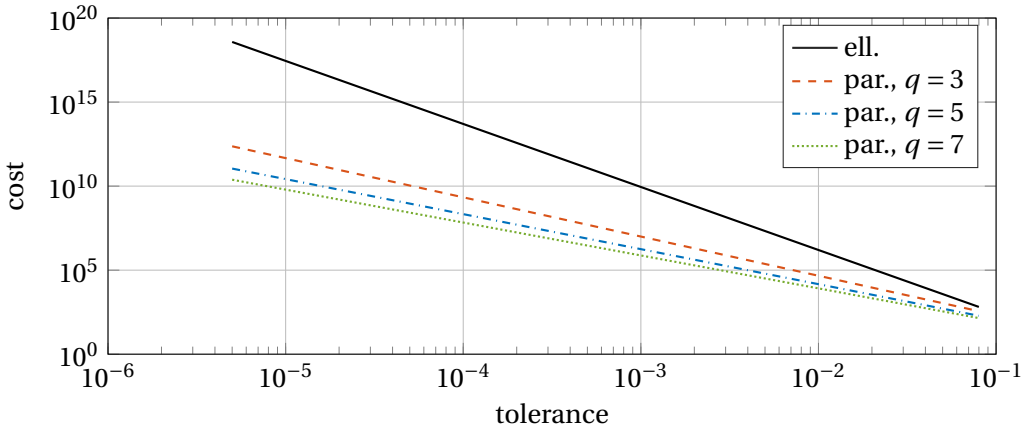


Figure 4.6 – Theoretical computational cost for $d = 3$, $\mathbb{P}_2$-FEM, 4-th order time integration, $q = 3, 5, 7$.

4.4 Conclusion

In this chapter, we have discussed an approach for numerical homogenization based on the solution of parabolic corrector problems. By assuming the coefficients to be sufficiently smooth and periodic and using Green's function estimates, we rigorously proved that the convergence rate of the resonance error can be arbitrarily high. Numerical tests demonstrate the same rates also for piecewise continuous and non-periodic cases. From the point of view of the computational complexity, the parabolic approximation obtained by high order filtering and using stabilised explicit solvers in time is asymptotically more efficient than the inversion of the discretized elliptic operator, required by elliptic approaches.

Despite the undoubted advantages of this method, it may seem cumbersome to solve a time independent problem, as eq. (3.1) by a time dependent approach, may it be the parabolic case presented here or the hyperbolic one proposed in [17, 18]. Besides having to deal with the fictitious time variable, time integration schemes have to be implemented to solve the local problems, which are stiff by nature. Moreover, the smooth periodic case is of purely academic interest, as realistic simulations often deal with materials with some degree of randomness. For this case, the analysis that is carried out in this chapter is not complete, as additional errors arise.

These issues motivated us to:

1. Develop an elliptic approach with arbitrary order of convergence, based on the results of this chapter and discussed in Chapter 5;

2. Explore the case of random media from both the theoretical and computational points of view, as done in Chapter 7.

5 Reduction of the resonance error via modified elliptic corrector problems

In Chapter 3 we discussed a novel elliptic corrector problem, with a modified right-hand side, to approximate the homogenized limit of multiscale coefficients in linear second order elliptic PDE. The new correctors, which can be equivalently constructed as time integral of the parabolic solutions of eq. (3.11), solve the equations:

$$\begin{cases} -\nabla\cdot\left(a(x)\nabla\chi^i_{T,R}\right) = g^i(x) - [e^{-AT}g^i](x) & \text{in } K_R, \\ \chi^i_{T,R} = 0 & \text{on } \partial K_R, \end{cases} \tag{5.1}$$

where $g^i(x) = \nabla\cdot(a(x)\mathbf{e}_i)$. Then, the effective coefficients at the macroscale are computed as

$$a^{0,R,L,T}_{ij} = \int_{K_L} \mathbf{e}_i \cdot a(x)\left(\mathbf{e}_j + \nabla\chi^j_{T,R}(x)\right)\mu_L(x)\,dx. \tag{5.2}$$

The arbitrary convergence rate of such a modified elliptic model is derived under the same assumptions on the multiscale tensor a^ε as in the parabolic model, i.e.:

i) $a^\varepsilon(x) = a(x/\varepsilon)$, for $a \in \mathscr{M}(\alpha,\beta)$;

ii) $a(\cdot)$ is K-periodic, with $K := [-1/2,1/2]^d$;

iii) $a(\cdot)\mathbf{e}_i \in H_{div}(K_R)$, for $i = 1,\dots,d$;

iv) $a(\cdot) \in \left[C^{1,\gamma}(K_R)\right]^{d\times d}$.

As in the parabolic case, we prove arbitrary convergence rates of the resonance error for the modified elliptic model.

Outline

The main result of this chapter is that the resonance error for this method decays with arbitrary convergence rate. The proof is given in Section 5.1 and it is based on previous results for

the parabolic case. In Section 5.2, we discuss some numerical techniques to pre-compute the additional term $e^{-AT}g^i$ without solving a time dependent PDE. This can be done by computing the matrix exponential, for which many algorithms are currently available, see [118]. This pre-computation step increases the computational cost of the model, but the gain in accuracy makes it more efficient than the standard elliptic method, as we discuss in Section 5.3. The theoretical convergence analysis is supported by several numerical experiments, see Section 5.4. The arbitrary convergence rate is also found for coefficients not satisfying the assumptions above, suggesting that they hold true under more general assumptions.

The content of this chapter is based on [8].

5.1 *A priori* analysis of the resonance error

The main result of this chapter is the following theorem, which gives an error bound for the difference between the exact homogenized coefficient a^0 and the approximation eq. (5.2) for a periodic coefficients $a(x)$.

Theorem 5.1. *Let $K_R \subset \mathbb{R}^d$, with $d \leq 3$ and $R \geq 1$. Let the coefficient matrix $a(\cdot)$ satisfy:*

i) $a(\cdot) \in \mathscr{M}(\alpha, \beta)$,

ii) $a(\cdot)$ *is K-periodic,*

iii) $a(\cdot)\mathbf{e}_i \in H_{div}(K_R)$, $i = 1, \dots, d$,

iv) $a(\cdot) \in \left[C^{1,\gamma}(K_R)\right]^{d\times d}$ *for some* $0 < \gamma \leq 1$.

Let $a^{0,R,L,T}$ and a^0 be defined, respectively, as in eq. (5.2) and eq. (3.4). Let $\mu_L \in \mathbb{K}^q(K_L)$ be a q-th order filter, with $0 < L < R - 3/2$ and $T \leq \frac{2c}{d+1}|R-L|^2$, with $c = 1/(4\beta)$. Then, there exists a constant $C > 0$ independent of R, L or T (but it may depend on d, $a(\cdot)$ and $\mu_L(\cdot)$) such that

$$\|a^{0,R,L,T} - a^0\|_F \leq C\left(\sqrt{T}L^{-(q+1)} + e^{-c_1 T} + \frac{R^{d-1}T^{\frac{5-d}{2}}}{|R-L|^3}e^{-c_2\frac{|R-L|^2}{T}}\right),$$

where $c_1 = \frac{\alpha\pi^2}{d}$ and $c_2 = \frac{1}{4\beta}$. Moreover, the choice

$$L = k_o R, \quad T = k_T R,$$

with $0 < k_o < 1$, and $k_T = \sqrt{\frac{c_2}{c_1}}(1-k_o)$ results in the following convergence rate in terms of R

$$\|a^{0,R,L,T} - a^0\|_F \leq C\left(R^{-q-\frac{1}{2}} + \gamma(R)e^{-\sqrt{c_1 c_2}(1-k_o)R}\right),$$

where $\gamma(R) = \left(R^{\frac{d-3}{2}} + 1\right)$, and C is a constant independent of R.

Remark 5.2. *Note that the exponent in the exponential term, $\sqrt{c_1 c_2} \approx \sqrt{\alpha/\beta}$, depends on the contrast ratio. So, the exponential part of the resonance error will be dominant for high contrast problems.*

In Theorem 5.1, The error $\sqrt{T}L^{-(q+1)}$ is the averaging error, which is obtained by using a filter $\mu_L \in \mathbb{K}^q(K_L)$. The order q of the filter can be chosen arbitrarily large with no additional computational cost. This allows to have better convergence rates for the resonance error. However, for higher order filters we witness a *plateau* in the convergence plot of the error, which is not present for low order filters, e.g., see Figure 5.3. The error $e^{-c_1 T}$ is related to the solution of the parabolic PDE eq. (4.2) for a finite T. Note that eq. (4.2) is introduced only for the analysis, but in practice, we don't solve it. The term $e^{-c\frac{|R-L|^2}{T}}$ along with its prefactor is an upper bound for the boundary error, and it will decay exponentially fast only if $T < |R-L|^2$.

The proof of Theorem 5.1 is split in four steps, in a way similar to the one adopted in the proof of Theorem 4.1. The four steps are:

Step 1: Decomposition of the resonance error into four terms:

$$\left| a_{ij}^{0,R,L,T} - a_{ij}^{0} \right| \leq e_{AV}(a_{ij}) + e_{AV}(\chi_{T,R}^{j}) + e_{TR} + e_{BC}.$$

Step 2: Estimation of the *averaging* errors $e_{AV}(a_{ij})$ and $e_{AV}(\chi_{T,R}^{j})$ by means of Lemma 4.3.

Step 3: Estimation of the *truncation* error e_{TR}.

Step 4: Estimation of the *boundary* error e_{BC} by means of upper bounds for the function θ^i derived in Section 4.1.4.

5.1.1 Error decomposition

The aim here is to show that the error can be split as

$$\left| a_{ij}^{0,R,L,T} - a_{ij}^{0} \right| \leq e_{AV}(a_{ij}) + e_{AV}(\chi_{T,R}^{j}) + e_{BC} + e_{TR}. \tag{5.3}$$

The terms $e_{AV}(a_{ij})$ and $e_{AV}(\chi_{T,R}^{j})$ are the averaging error which decreases by using filters $\mu_L \in \mathbb{K}^q(K_L)$ with higher values for q. The error e_{TR} is associated with truncation in time of the solutions of parabolic cell-problems. The boundary error e_{BC} quantifies the effect of boundary conditions. To see this, we use Theorem 3.8 and write

$$\begin{aligned} a_{ij}^{0,R,L,T} &:= \int_{K_L} a_{ij}(x)\mu_L(x)\,dx + \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla\chi_{T,R}^{j}(x)\mu_L(x)\,dx \\ &= \int_{K_L} a_{ij}(x)\mu_L(x)\,dx + \int_0^T \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla u_R^j(x,t)\mu_L(x)\,dx\,dt, \end{aligned}$$

where u_R^j is the solution of the parabolic corrector problem eq. (4.2). In the same way, by Theorem 3.2, the exact homogenized coefficient can be rewritten as

$$\begin{aligned} a_{ij}^0 &= \fint_K a_{ij}(x)\,dx + \fint_K \mathbf{e}_i \cdot a(x)\nabla\chi^j(x)\,dx \\ &= \fint_K a_{ij}(x)\,dx + \int_0^\infty \fint_K \mathbf{e}_i \cdot a(x)\nabla v^j(x,t)\,dx\,dt, \end{aligned}$$

where v^j is the periodic parabolic solution in eq. (3.5), and χ^j is the solution to the periodic corrector problem eq. (3.3) We exploit this equality to further decompose the error as follows

$$\begin{aligned} \left|a_{ij}^{0,R,L,T} - a_{ij}^0\right| \le & \underbrace{\left|\int_{K_L} a_{ij}(x)\mu_L(x)\,dx - \fint_K a_{ij}(x)\,dx\right|}_{e_{AV}(a_{ij})} \\ & + \underbrace{\left|\int_0^T \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla u_R^j(x,t)\mu_L(x)\,dx\,dt - \int_0^T \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla v^j(x,t)\mu_L(x)\,dx\,dt\right|}_{e_{BC}} \\ & + \underbrace{\left|\int_0^T \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla v^j(x,t)\mu_L(x)\,dx\,dt - \int_0^T \fint_K \mathbf{e}_i \cdot a(x)\nabla v^j(x,t)\,dx\,dt\right|}_{e_{AV}(\chi_{T,R}^j)} \\ & + \underbrace{\left|\int_0^T \fint_K \mathbf{e}_i \cdot a(x)\nabla v^j(x,t)\,dx\,dt - \int_0^\infty \fint_K \mathbf{e}_i \cdot a(x)\nabla v^j(x,t)\,dx\,dt\right|}_{e_{TR}}. \end{aligned} \tag{5.4}$$

In the following steps we give bounds for all the errors.

5.1.2 Averaging errors bounds

We now give an *a priori* estimate on ∇v^i, based on the spectral properties of the periodic cell-problem. This will be used in the proof of Lemma 5.4.

Lemma 5.3. *Let $a \in \mathcal{M}(\alpha,\beta)$ be K-periodic, $v^i \in C\left([0,+\infty), L_0^2(K)\right)$ be the weak solution of eq.* (3.5) *and let $g^i(x) := v^i(x,0) \in W_{per}^1(K)$. Then, there exist $C(\alpha,\beta) > 0$ such that*

$$\left\|\nabla v^i\right\|_{L^1([0,+\infty);L^2(K))} \le C\left\|\nabla g^i\right\|_{L^2(K)}. \tag{5.5}$$

Proof. Let us define the bilinear form $B : W_{per}^1(K) \times W_{per}^1(K) \mapsto \mathbb{R}$ as

$$B(w,\hat{w}) := \int_K \nabla\hat{w}(x) \cdot a(x)\nabla w(x)\,dx, \quad w,\hat{w} \in W_{per}^1(K),$$

and let us denote the eigenvalues and eigenfunctions of $B(\cdot,\cdot)$ by $\{\lambda_k\}_{k=0}^\infty$ and $\{\varphi_k\}_{k=0}^\infty$, respectively. It is well known that the sequence of eigenvalues is positive and non-decreasing,

i.e.,

$$0 < \lambda_0 \le \lambda_1 \le \lambda_2, \dots.$$

The eigenfunctions $\{\varphi_k\}_{k=0}^{\infty}$ are orthonormal in the L^2-sense and they satisfy:

$$B\left(\varphi_k, v\right) = \lambda_k \langle \varphi_k, v \rangle_{L^2(K)}, \quad \forall v \in W^1_{per}(K). \tag{5.6}$$

Since the eigenvalues form a basis of $W^1_{per}(K)$, we can write the function $v^i(t,x) = \sum_{k=0}^{\infty} v^i_k(t)\varphi_k(x)$. By plugging this expression into eq. (5.6), we conclude that the components of v^i decay exponentially in time:

$$v^i_k(t) = e^{-\lambda_k t} g^i_k, \quad \text{where } g^i_k := \langle g^i, \varphi_k \rangle_{L^2(K)}. \tag{5.7}$$

By eq. (5.7) and by coercivity of the bilinear form, we obtain

$$\begin{aligned}\alpha \left\| \nabla v^i(t,\cdot) \right\|^2_{L^2(K)} \le B\left(v^i, v^i\right)(t) &= \sum_{k,\ell=0}^{\infty} e^{-(\lambda_k+\lambda_\ell)t} g^i_k g^i_\ell B\left(\varphi_k, \varphi_\ell\right) \\ &= \sum_{k,\ell=0}^{\infty} e^{-(\lambda_k+\lambda_\ell)t} g^i_k g^i_\ell \lambda_k \langle \varphi_k, \varphi_\ell \rangle_{L^2(K)} = \sum_{k=0}^{\infty} e^{-2\lambda_k t} |g^i_k|^2 \lambda_k.\end{aligned}$$

Hence,

$$\begin{aligned}\left\| \nabla v^i(t,\cdot) \right\|_{L^1([0,+\infty);L^2(K))} &:= \int_0^{+\infty} \left\| \nabla v^i(t,\cdot) \right\|_{L^2(K)} \, dt \\ &\le \int_0^{+\infty} \sqrt{\frac{1}{\alpha} \sum_{k=0}^{\infty} e^{-2\lambda_k t} \lambda_k |g^i_k|^2} \, dt \le \sqrt{\frac{1}{\alpha} \sum_{k=0}^{\infty} \lambda_k \left|g^i_k\right|^2} \int_0^{+\infty} e^{-\lambda_0 t} \, dt \\ &= \frac{\alpha^{-1/2}}{\lambda_0} \sqrt{B\left(g^i, g^i\right)} \le \sqrt{\frac{\beta}{\alpha}} \frac{1}{\lambda_0} \left\| \nabla g^i \right\|_{L^2(K)}.\end{aligned}$$

□

The main result of this section is summarised in the following lemma.

Lemma 5.4. *Let $a \in \mathcal{M}(\alpha,\beta)$ be K-periodic, $a\mathbf{e}_i \in H_{div}(K)$ and $e_{AV}(a_{ij})$ and $e_{AV}(\chi^j_{T,R})$ be defined as in eq.* (5.4)*. Then, there exists a constant $C > 0$ independent of R, T, L (but it depends on a and μ_L)such that*

$$e_{AV}(a_{ij}) + e_{AV}(\chi^j_{T,R}) \le \begin{cases} C\sqrt{T}L^{-q-1} & \text{if} \nabla \cdot (a\mathbf{e}_i) \in L^2(K), \\ CL^{-q-1} & \text{if} \nabla \cdot (a\mathbf{e}_i) \in W^1_{per}(K). \end{cases}$$

Proof. By Lemma 4.3, we can immediately see that

$$\begin{aligned} e_{AV}(a_{ij}) &:= \left| \int_{K_L} a_{ij}(x)\mu_L(x) \, dx - \fint_K a_{ij}(x) \, dx \right| \\ &\le CL^{-q-1} \|a_{ij}\|_{L^2(K)} \le C\beta L^{-q-1}. \end{aligned}$$

Moreover,

$$e_{AV}(\chi^j_{T,R}) := \left| \int_0^T \left(\int_{K_L} \mathbf{e}_i \cdot a(x) \nabla v^j(x,t) \mu_L(x)\, dx - \fint_K \mathbf{e}_i \cdot a(x) \nabla v^j(x,t)\, dx \right) dt \right|$$
$$\leq CL^{-q-1} \int_0^T \|\mathbf{e}_i \cdot a(x) \nabla v^j(t,\cdot)\|_{L^2(K)}\, dt \leq CL^{-q-1} \beta \int_0^T \|\nabla v^j(t,\cdot)\|_{L^2(K)}\, dt.$$

If the tensor $a(x)$ has higher regularity, i.e. $\nabla \cdot (a\mathbf{e}_i) \in W^1_{per}(K)$, we can directly estimate $\|\nabla v^j\|_{L^1(0,T;L^2(K))} := \int_0^T \|\nabla v^j(t,\cdot)\|_{L^2(K)}\, dt$ by eq. (5.5) in Lemma 5.3 and obtain

$$e_{AV}(\chi^j_{T,R}) \leq C\beta L^{-q-1} \|\nabla \cdot \left(a\mathbf{e}_j\right)\|_{W^1_{per}(K)}.$$

Otherwise, if $\nabla \cdot \left(a\mathbf{e}_j\right) \in L^2(K)$ only, we will apply Cauchy-Schwarz inequality which yields

$$\int_0^T \|\nabla v^j(t,\cdot)\|_{L^2(K)}\, dt \leq \sqrt{T} \|\nabla v^j\|_{L^2(0,T;L^2(K))}.$$

Then, by employing Equation (3.6) in Proposition 3.1, we obtain

$$e_{AV}(\chi^j_{T,R}) \leq C\beta\sqrt{T} L^{-q-1} \|\nabla \cdot \left(a\mathbf{e}_j\right)\|_{L^2(K)}.$$

□

5.1.3 Truncation error bound

Lemma 5.5. *Let $a \in \mathcal{M}(\alpha,\beta)$ be K-periodic and $a\mathbf{e}_i \in H_{div}(K)$. Then the truncation error e_{TR} defined in eq.* (5.4) *satisfies the estimate*

$$e_{TR} \leq Ce^{-\frac{\alpha\pi^2}{d}T},$$

where α is the coercivity constant and C is a constant independent of T (but it depends on α and d).

Proof. By using integration by parts and the Cauchy-Schwarz inequality we have

$$\begin{aligned}
e_{TR} &:= \left| \int_T^\infty \fint_K \mathbf{e}_i \cdot a(x) \nabla v^j(t,x)\, dx\, dt \right| \\
&= \left| \int_T^\infty \fint_K \nabla \cdot (a\mathbf{e}_i)\, v^j(t,x)\, dx\, dt \right| \\
&\leq \int_T^\infty \|\nabla \cdot (a\mathbf{e}_i)\|_{L^2(K)} \|v^j(t,\cdot)\|_{L^2(K)}\, dt \\
&\leq \|\nabla \cdot (a\mathbf{e}_i)\|_{L^2(K)} \int_T^\infty e^{-\lambda_0 t} \|\nabla \cdot \left(a\mathbf{e}_j\right)\|_{L^2(K)}\, dt \\
&= \left\|\nabla \cdot \left(a\mathbf{e}_j\right)\right\|_{L^2(K)} \|\nabla \cdot (a\mathbf{e}_i)\|_{L^2(K)} \frac{1}{\lambda_0} e^{-\lambda_0 T},
\end{aligned}$$

where $\lambda_0 \geq \alpha C_p(K)^{-2}$, and $C_p(K)$ is the constant of the Poincaré-Wirtinger inequality in $W^1_{per}(K)$, which can be bounded by

$$C_p(K) \leq \frac{diam(K)}{\pi} = \frac{\sqrt{d}}{\pi},$$

see [124]. Hence, $\lambda_0 \geq \frac{\alpha\pi^2}{d}$ and the final result follows. □

5.1.4 Boundary error bound

Lemma 5.6. *Let $a \in \mathscr{M}(\alpha,\beta)$ be K-periodic, $a(\cdot) \in \left[C^{1,\gamma}(K_R)\right]^{d\times d}$ for some $0 < \gamma \leq 1$ and $\mu_L \in \mathbb{K}^q(K_L)$ with $L < \tilde{R}$, where $\tilde{R}$ is the largest integer such that $\tilde{R} \leq R - 1/2$. Then, there exists a constant $C > 0$ independent of R, L and T (but it depends on a and μ_L) such that the boundary error e_{BC} defined in eq.* (5.4) *satisfies*

$$e_{BC} \leq C\frac{R^{d-1}T^{\frac{5-d}{2}}}{|R-L|^3}e^{-c\frac{|R-L|^2}{T}},$$

where $c = \frac{1}{4\beta}$.

Proof. To estimate the boundary error, we define $\theta^j = u^j - \rho v^j$, where the smooth function $\rho \in C^\infty_c(K_R)$ is a cut-off function of Definition 4.13. Then, $\left(u^j_R - v^j\right)(x,t) = \theta^j(x,t)$ for any $t > 0$ and $x \in K_L \subset K_{\tilde{R}}$, hence

$$\begin{aligned} e_{BC} &:= \left|\int_0^T \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla\left(u^j_R - v^j\right)(x,t)\mu_L(x)\,dx\,dt\right| \\ &= \left|\int_0^T \int_{K_L} \mathbf{e}_i \cdot a(x)\nabla\theta^j(x,t)\mu_L(x)\,dx\,dt\right|. \end{aligned}$$

Next, by integration by parts it follows that

$$\begin{aligned} e_{BC} &\leq \int_{K_L} \left|\nabla\cdot\left(a\mathbf{e}_i\mu_L\right)\right|\,dx \sup_{x\in K_L}\int_0^T |\theta^j(x,t)|\,dt \\ &\leq \|\mu_L\|_{W^{1,2}(K_L)}\|a\mathbf{e}_i\|_{H_{div}(K_L)} \sup_{x\in K_L}\int_0^T |\theta^j(x,t)|\,dt \\ &\leq C_\mu L^{-d/2}L^{d/2}\|a\mathbf{e}_i\|_{H_{div}(K)} \sup_{x\in K_L}\int_0^T |\theta^j(x,t)|\,dt. \end{aligned}$$

By using the bound for $\sup_{x\in K_L}\left|\theta^j(x,t)\right|$ of Lemma 4.15, we bound $\sup_{x\in K_L}\int_0^T \left|\theta^j(x,t)\right|\,dt$ using the change of variable $s = c\frac{|R-L|^2}{t}$

$$\sup_{x\in K_L}\int_0^T \left|\theta^j(x,t)\right|\,dt \leq C\frac{R^{d-1}}{|R-L|}\|\nabla v^j\|_{L^2([0,\infty);L^2(K))}\int_0^T \left(\frac{1}{t} + \frac{1}{2c|R-L|^2}\right)^{\frac{d-1}{2}} e^{-c\frac{|R-L|^2}{t}}\,dt$$

$$\leq CR^{d-1}|R-L|^{2-d}\|\nabla g^j\|_{L^2(K)}\int_{\frac{c|R-L|^2}{T}}^{\infty}\frac{(s+\frac{1}{2})^{\frac{d-1}{2}}}{s^2}e^{-s}\,ds$$

$$\leq CR^{d-1}|R-L|^{2-d}\|\nabla g^j\|_{L^2(K)}\frac{T^2\left(\frac{c|R-L|^2}{T}+\frac{1}{2}\right)^{\frac{d-1}{2}}}{c^2|R-L|^4}\int_{\frac{c|R-L|^2}{T}}^{\infty}e^{-s}\,ds$$

$$\leq C\|\nabla g^j\|_{L^2(K)}\frac{R^{d-1}T^{\frac{5-d}{2}}}{|R-L|^3}e^{-c\frac{|R-L|^2}{T}}.$$

□

Remark 5.7. *We emphasize here that one of the key arguments in proving an exponentially decaying error bound for e_{BC} is the requirement that $L < R$, as done in Chapter 4.*

5.1.5 *A priori* bound on the resonance error for the modified elliptic approach

We can now prove Theorem 5.1, by recalling the results of Sections 5.1.1 to 5.1.4.

Proof of Theorem 5.1. The decomposition eq. (5.4) implies

$$\left\|a^{0,R,L,T}-a^0\right\|_F \leq d^2\max_{i,j}\left(\left|e_{AV}(a_{ij})\right|+\left|e_{AV}(\chi^j_{T,R})\right|+|e_{TR}|+|e_{BC}|\right).$$

By using the upper bounds in Lemmas 5.4 to 5.6 in the above inequality we get

$$\|a^{0,R,L,T}-a^0\|_F \leq C\left(\sqrt{T}L^{-(q+1)}+e^{-c_1T}+\frac{R^{d-1}T^{\frac{5-d}{2}}}{|R-L|^3}e^{-c_2\frac{|R-L|^2}{T}}\right)$$

for some constant C independent of R, L and T. □

5.2 Approximation of the exponential operator e^{-TA}

From a computational perspective, the right-hand side $e^{-TA}g^i$ in eq. (5.1) must be approximated in order to compute the modified corrector $\chi^i_{T,R}$. One can look at $e^{-TA}g^i$ as the solution at time T of a parabolic PDE with initial data g^i. Then, the naive approach would be to, first, discretize the equation in space and, then, solve the ensuing evolution problem by some time discretization scheme. With such a procedure, the approximation of $e^{-TA}g^i$ would *a priori* suffer from the discretization error in both space and time and it would not lead to any gain in the computational cost in comparison to the parabolic approach described in Chapter 4. We will not discuss the parabolic approach further, instead, we will describe other approaches that aim to approximate $e^{-TA}g^i$ without the use of time-advancing schemes.

One can look for an approximation of the exponential operator *before* or *after* the semi-discretization in space of the corrector problem. In the first case, the approximation of $e^{-TA}g^i$ is sought in a finite dimensional subspace $V_m \subset H^1_0(K_R)$ of dimension m. In the other case, we

have to approximate the matrix exponential e^{-TA_h}, where the matrix $A_h \in \mathbb{R}^{N\times N}$ depends on the discretization, and $V_m \subset \mathbb{R}^N$.

The subspace V_m can be chosen in several ways: an approach based on a continuous analogue of the Krylov subspace method has recently been proposed in [76]. The continuous Krylov subspaces are constructed by iterative action of the operator A on the initial condition g^i and of properly chosen projection/preconditioning operators that are meant to ensure that the product Ag^i does always make sense. In Section 5.2.1, we discuss an alternative approximation, based on *spectral truncation*, where V_m is generated by m eigenvalues of the second order operator A.

In the discrete setting, i.e. when the corrector problem is discretized and has the form of a linear equation in $\mathbb{R}^N$, the right-hand side term is a matrix exponential, which depends on the chosen discretization technique. For example, if the Finite Difference Method is used, the semi-discrete parabolic corrector problem eq. (3.11) will be

$$\frac{d}{dt}\mathbf{w}_h = -A_h\mathbf{w}_h, \quad \Longrightarrow \quad \mathbf{w}_h(T) = e^{-TA_h}\mathbf{g}_h^i,$$

where $\mathbf{w}_h : \mathbb{R}^+ \mapsto \mathbb{R}^N$, $A_h \in \mathbb{R}^{N\times N}$ is the stiffness matrix and $\mathbf{g}_h \in \mathbb{R}^N$ is the evaluation of the initial condition g^i on the mesh nodes[1]. Hence, the term $\mathbf{w}_h(T)$ will be used in the discrete version of (5.1) in place of $e^{-TA}g^i$. Instead, if the Finite Element Method is chosen to discretize the parabolic corrector problem, the following ODE system arises:

$$M_h\frac{d}{dt}\mathbf{w}_h = -A_h\mathbf{w}_h, \quad \Longrightarrow \quad \mathbf{w}_h(T) = e^{-TM_h^{-1}A_h}\mathbf{g}_h^i,$$

where we have to consider also the mass matrix M_h. The computation of the matrix exponential is crucial in many applications, such as the development of *exponential time integrators* for ODE systems [95]. Several methods are available to compute the matrices exponential, or more general matrix functions, see [85, 92, 117, 118, 125]. However, these methods may not be directly applicable due to the large size of A_h (and M_h), and approximations of the product $e^{-TA_h}\mathbf{g}_h^i$ (or $e^{-TM_h^{-1}A_h}\mathbf{g}_h^i$) must be sought in smaller subspaces $V_m \subset \mathbb{R}^N$. The space V_m can be chosen as the m-th dimensional Krylov subspace generated by $(A_h, \mathbf{g}_h^i)$, [94], or the rational/extended Krylov subspaces, [49, 50, 72, 86]. The convergence of the rational Krylov subspace method does not depend on the spectral radius of the matrix (which is linked to its size) but requires to solve m linear systems. In Section 5.2.2 we analyse the use of the standard Krylov subspace method to approximate e^{-TA_h}.

5.2.1 Spectral truncation

Here, we discuss a technique to approximate the exponential operator e^{-AT} by spectral decomposition. This method can be used for any operator A with compact, self-adjoint and

[1] We dropped the i superscript in $\mathbf{g}_h$ for the sake of simplicity in notation.

positive definite inverse A^{-1}, so we will not need to assume the coefficients a to be periodic.

The correction term $e^{-AT}g^i$ in eq. (5.1) corresponds to the solution (at time T) of the parabolic PDE, eq. (4.2), which can be expressed as

$$[e^{-AT}g^i](x) := \sum_{k=0}^{\infty} e^{-\lambda_k T}\hat{g}_k^i \varphi_k(x), \text{ with } \quad \hat{g}_k^i := \langle g^i, \varphi_k \rangle_{L^2(K_R)},$$

where $\{\lambda_k\}_{k=0}^{\infty} \subset \mathbb{R}^+$ and $\{\varphi_k\}_{k=0}^{\infty} \subset H_0^1(K_R)$ are the eigenvalues and eigenfunctions of the operator A. If T is not too small, most of the modes in the expansion can be neglected due to the exponential decay with respect to the eigenvalues. Hence solving a more expensive parabolic PDE can be avoided at the expense of computing a few dominant modes of the operator A. To this end, let

$$[S_m g^i](x) := \sum_{k=0}^{m-1} e^{-\lambda_k T}\hat{g}_k^i \varphi_k(x).$$

Then, the cell-problem eq. (5.1) can be approximated by

$$\begin{cases} -\nabla \cdot \left(a(x)\nabla \chi_{T,R,m}^i\right) = (I - S_m)\, g^i & \text{in } K_R, \\ \chi_{T,R,m}^j = 0 & \text{on } \partial K_R. \end{cases} \tag{5.8}$$

Similarly, the homogenized coefficient of eq. (5.2) is approximated by

$$a_{ij}^{0,R,L,T,m} = \int_{K_L} \mathbf{e}_i \cdot a(x)\left(\mathbf{e}_j + \nabla \chi_{T,R,m}^j\right)\mu_L(x)\, dx. \tag{5.9}$$

In the discretized version of eq. (5.9), the spectral truncation is performed on the matrix exponential e^{-TA_h} (or $e^{-TM_h^{-1}A_h}$), not on the exponential operator. However, the *a priori* error analysis for the continuous, spectrally truncated cell-problem in eq. (5.8) allows to derive error bound which are independent on the discretization. In the following lemma, we give a bound for the *spectral* error, defined as the difference between $a^{0,R,L,T}$ of eq. (5.2) and $a^{0,R,L,T,m}$ of eq. (5.9).

Lemma 5.8. *Let $a \in \mathcal{M}(\alpha, \beta, K_R)$, $a\mathbf{e}_i \in H_{div}(K_R)$, and $\mu_L \in \mathbb{K}^q(K_L)$. Moreover, let $a^{0,R,L,T}$ and $a^{0,,R,L,T,m}$ be defined as in eq.* (5.2) *and eq.* (5.9) *respectively. Then*

$$e_{SP} := \left| a_{ij}^{0,R,L,T} - a_{ij}^{0,R,L,T,m} \right| \leq C\left(\frac{R}{L}\right)^{\frac{d}{2}} R \exp\left(-\frac{c_d m^{2/d} T}{R^2}\right) \tag{5.10}$$

where $C(\alpha, \beta, d, \mu_L)$ and c_d are constants independent of R, L, T and m.

Proof. Let

$$[e^{-AT}g^i](x) = \sum_{k=0}^{\infty} e^{-\lambda_k T}\hat{g}_k^i \varphi_k(x), \quad [S_m g^i](x) = \sum_{k=0}^{m-1} e^{-\lambda_k T}\hat{g}_k^i \varphi_k(x),$$

where $\{\lambda_j, \varphi_j(x)\}_{j=0}^{\infty}$ are the eigenvalue-function pairs of the operator $A = -\nabla\cdot(a\nabla)$ with Dirichlet boundary conditions on the domain K_R. Moreover, let $E_m := e^{-AT}g^i - S_m g^i$, with $g^i := \nabla\cdot(a\mathbf{e}_i)$. The eigenvalues of second order symmetric elliptic operators satisfy

$$\lambda_k \geq c_d k^{2/d}|K_R|^{-2/d} = c_d k^{2/d} R^{-2}, \tag{5.11}$$

where c_d is a constant that depends on the dimension[2] d and the ellipticity constant α, see [110, 126]. Then

$$\begin{aligned}\|E_m\|^2_{L^2(K_R)} &\leq \sum_{\ell,k=m}^{\infty} e^{-\frac{c_d(\ell^{2/d}+k^{2/d})T}{R^2}} g^i_\ell g^j_k \langle\varphi_\ell\varphi_k\rangle_{L^2(K_R)} \\ &= \sum_{k=m}^{\infty} e^{-\frac{2c_d k^{2/d}T}{R^2}}|\hat{g}^j_k|^2 \leq e^{-\frac{2c_d m^{2/d}T}{R^2}}\|g^i\|^2_{L^2(K_R)}.\end{aligned}$$

Taking the square root of both sides, we arrive at

$$\|E_m\|_{L^2(K_R)} \leq e^{-\frac{c_d m^{2/d}T}{R^2}}\|g^i\|_{L^2(K_R)}.$$

Moreover, since the difference $\psi := \chi^i_{T,R} - \chi^i_{T,R,m}$ satisfies $-\nabla\cdot a(x)\nabla\psi(x) = E_m(x)$ with homogeneous Dirichlet BCs, standard elliptic regularity yields

$$\begin{aligned}\|\chi^i_{T,R} - \chi^i_{T,R,m}\|_{H^1_0(K_R)} &\leq \frac{C_p(K_R)}{\alpha}\|E_m\|_{L^2(K_R)} \\ &\leq CRe^{-\frac{c_d m^{2/d}T}{R^2}}\|g^i\|_{L^2(K_R)} \\ &\leq CR^{1+\frac{d}{2}}e^{-\frac{c_d m^{2/d}T}{R^2}}\|a\mathbf{e}_i\|_{H_{div}(K)},\end{aligned}$$

where we have used the fact that the Poincaré constant $C_p(K_R)$ is bounded by $C_p(K_R) \leq diam(K_R)/\pi = R2^{1/d}/\pi$, see [124], and $\|g^i\|_{L^2(K_R)} \leq |K_R|^{1/2}\|a\mathbf{e}_i\|_{H_{div}(K)}$. Finally,

$$\begin{aligned}\left|a^{0,R,L,T}_{ij} - a^{0,R,L,T,m}_{ij}\right| &= \int_{K_L}\mathbf{e}_i\cdot a(x)\left(\nabla\chi^i_{T,R} - \nabla\chi^i_{T,R,m}\right)\mu_L(x)\,dx \\ &\leq \alpha|K_L|^{1/2}\|\nabla\chi^i_{T,R} - \nabla\chi^i_{T,R,m}\|_{L^2(K_R)}\frac{1}{L^d}\|\mu\|_{L^\infty(K)} \\ &\leq C\frac{R^{1+\frac{d}{2}}}{L^{\frac{d}{2}}}e^{-\frac{c_d m^{2/d}T}{R^2}}.\end{aligned}$$

This completes the proof. □

In Theorem 5.1, the optimal value for the parameter T is $T = \mathcal{O}(R)$. In order to get an exponential decay rate, such as e^{-cR} for some positive c, in Lemma 5.8, we then need to compute $m = \mathcal{O}(R^d)$ eigenmodes. This growth of the number of eigenmodes with respect to the dimen-

[2]The constant c_d may depend on α and β too. The value of c_d can be approximated by computing a few eigenvalues λ_k and finding the largest constant so that the relation eq. (5.11) holds.

sion is the main drawback of the naive spectral truncation leading to a high computational burden in higher dimensions. Therefore, in the next subsection we propose a much more efficient method based on the Krylov subspace projection, and we show that the cost of the method will scale linearly in terms of the number of degrees of the freedom, while retaining the desired exponential accuracy for the approximation of the homogenized coefficient.

5.2.2 Approximation by the Krylov subspace method

In this section we discuss the computation of the matrix exponential $e^{-TA_h}\mathbf{g}_h$, which is the discrete counterpart of $e^{-TA}g^i$. The matrix $A_h \in \mathbb{R}^{N\times N}$ comes from the Finite Difference discretization in space of a partial differential equation and, thus, it is of large dimension and sparse. For these conditions, standard methods for computing e^{-TA_h}, like those reviewed in [117], may be very inefficient. Iterative methods based on the Krylov subspace method have successfully been applied to many problems involving large sparse matrices. When the modified corrector problem is discretize through the Finite Elements Method, the matrix exponential to be computed is $e^{-TM_h^{-1}A_h}$, as anticipated before. In this case, the sparsity property is lost, because M_h^{-1} is full. Hence, one may recur to mass-lumping in order to keep the sparsity pattern. We will not discuss the issues due to the presence of the mass matrix any further, instead we will focus on exploiting Krylov subspaces to approximate $e^{-TA_h}\mathbf{g}_h$.

The Krylov subspace method allows to find an approximation of $f(A_h)\mathbf{g}_h$ (for any matrix function f, matrix $A_h \in \mathbb{R}^{N\times N}$ and $\mathbf{g}_h \in \mathbb{R}^N$) within the subspace

$$V_m = \text{span}\left\{\mathbf{g}_h, A_h\mathbf{g}_h, A_h^2\mathbf{g}_h, \dots, A_h^{m-1}\mathbf{g}_h\right\},$$

with $m \ll N$. A basis of V_m is constructed by the Arnoldi algorithm[3] which gives the decomposition:

$$A_hQ_m = Q_mH_m + h_{m+1,m}\mathbf{q}_{m+1}\mathbf{e}_m^T \implies Q_m^TA_hQ_m = H_m,$$

where $Q_m \in \mathbb{R}^{N\times m}$ has orthonormal columns, $H_m \in \mathbb{R}^{m\times m}$ is an upper-Hessenberg matrix (tridiagonal, if A_h is symmetric), $\mathbf{q}_{m+1} \in V_{m+1} \subset \mathbb{R}^N$ is orthogonal to the basis Q_m and $\mathbf{e}_m$ is the m-th basis vector of $\mathbb{R}^m$. The term $e^{-TA_h}\mathbf{g}_h$ can then be approximated by

$$e^{-TA_h}\mathbf{g}_h \approx Q_me^{-TH_m}\mathbf{e}_{1,\mathbf{g}},$$

where $\mathbf{e}_{1,\mathbf{g}} = \left|\mathbf{g}_h\right|\mathbf{e}_1 \in \mathbb{R}^m$. Therefore, computing the computationally expensive exponential matrix function e^{-TA_h} of size $N\times N$ is avoided by instead computing e^{-TH_m}, with a smaller computational cost. The calculation of e^{-TH_m} can be performed, e.g., by the scaling and squaring algorithm.

An important question regards the approximation error coming from the Arnoldi algorithm. The following theorem from [94] provides an upper bound for such an approximation.

[3] In our case, it is more appropriate to use the Lanczos algorithm, as A_h is symmetric and positive definite.

Theorem 5.9. *(Hochbruck and Lubich [94]) Let $B \in \mathbb{R}^{N\times N}$ be a Hermitian positive semi-definite matrix with eigenvalues in $[0,\rho]$. Moreover, let $H_m = Q_m^T B Q_m$ be a unitary transformation of B via an Arnoldi procedure with $H_m \in \mathbb{R}^{m\times m}$ and $Q_m \in \mathbb{R}^{N\times m}$. Then, for any $\tau \geq 0$, it holds that*

$$\left|e^{-\tau B}\mathbf{g}_h - Q_m e^{-\tau H_m}\mathbf{e}_{1,\mathbf{g}}\right| \leq \begin{cases} 10\left|\mathbf{g}_h\right| e^{-4m^2/(5\rho\tau)}, & \sqrt{\rho\tau} \leq m \leq \frac{\rho\tau}{2}, \\ \dfrac{40\left|\mathbf{g}_h\right|}{\rho\tau} e^{-\rho\tau/4}\left(\dfrac{e\rho\tau}{4m}\right)^m & m \geq \frac{\rho\tau}{2}. \end{cases} \tag{5.12}$$

Corollary 5.10. *Let $A_h \in \mathbb{R}^{N\times N}$ be a second order centred finite difference approximation of the operator $-\nabla\cdot(a\nabla)$ on the domain K_R with homogeneous Dirichlet boundary conditions and let $\rho(A_h) \leq c_\rho h^{-2}$ be its spectral radius. Let $H_m = Q_m^T A_h Q_m$ be a unitary transformation of A_h via an Arnoldi procedure with $H_m \in \mathbb{R}^{m\times m}$, $Q_m \in \mathbb{R}^{m\times N}$. Then the following estimate holds*

$$\left|e^{-TA_h}\mathbf{g}_h - Q_m e^{-TH_m}\mathbf{e}_{1,\mathbf{g}}\right| \leq 10\left|\mathbf{g}_h\right| e^{-\frac{4m^2h^2}{5T}},$$

for $\sqrt{\rho(A_h)T} \leq m \leq \rho(A_h)T/2$. Moreover, when $m = \sqrt{\rho(A_h)}T/2$, the estimate reads as

$$\left|e^{-TA_h}\mathbf{g}_h - Q_m e^{-TH_m}\mathbf{e}_{1,\mathbf{g}}\right| \leq 10\left|\mathbf{g}_h\right| e^{-T/5}, \quad \text{for } T \geq 4. \tag{5.13}$$

Proof. The matrix A_h is symmetric and positive definite, so Theorem 5.9 can be directly applied and the proof follows from the fact that the spectral radius of A_h is bounded by $c_\rho h^{-2} = c_\rho N^{2/d} R^{-2}$. □

The advantage of using an approximation for the exponential correction term via the Arnoldi approach is that the number of basis functions required in the Arnoldi iteration is independent of the dimension of the problem. In other words, denoting the numbers of degrees of freedom in d-dimensions by $N = \mathcal{O}\left((R/h)^d\right)$, only $m = (\sqrt{c_\rho}T)/(2h)$ basis functions are needed to obtain an exponentially accurate approximation for the exponential correction $e^{-AT}g$ up to a discretization error, see the estimate of eq. (5.13). Moreover, an estimate for a fully discrete approximation of the homogenized coefficient can also be derived similar to the analysis in the spectral section, where the upper bound will include the exponential estimate in Corollary 5.10 in addition to an error coming from the spatial discretization.

Remark 5.11. *The error estimates of Theorem 5.9 depend on the spectral radius ρ of the matrix A_h. Thus, the error estimates deteriorates for $h \to 0$, since $\rho = \mathcal{O}(h^{-2})$, and the dimension of V_m must increase accordingly. This is related to the fact that the iterate application of the operator $A : H_0^1(H_R) \cap H^2(K_R) \mapsto L^2(K_R)$ may not be possible if $g^i \notin H_0^1(H_R) \cap H^2(K_R)$ (or $A^q g^i \notin H_0^1(H_R) \cap H^2(K_R)$ for some $q \geq 1$). The use of rational/extended Krylov subspaces is known to mitigate this issue, as it is possible to derive error estimates like eq.* (5.12) *that are independent on the spectral radius, [50, 86]. The extended Krylov subspaces are defined as*

$$V_{m,q}^{\gamma} := \operatorname{span}\left\{\left(\gamma I - A_h\right)^{-q+1}\mathbf{g}_h, \ldots, \left(\gamma I - A_h\right)^{-1}\mathbf{g}_h, \mathbf{g}_h, A_h\mathbf{g}_h, \ldots, A_h^{m-1}\mathbf{g}_h\right\},$$

and their use has been extended to the approximation of continuous operators in [72] with applications in exponential integrators.

5.3 Discussion over the computational cost

The Arnoldi iteration of Section 5.2.2 can be used in different ways to approximate, in the subspace V_m, the solution of the modified elliptic PDE eq. (5.1). A standard Finite Difference discretization of the problem eq. (5.1) results in the following system[4]

$$A_h \mathbf{w}_h = \mathbf{g}_h - e^{-TA_h}\mathbf{g}_h. \tag{5.14}$$

Here we present two different ways based on the Arnoldi iteration to solve eq. (5.14).

Approach 1. Let $f_1(z) = e^{-tz}$, then the system eq. (5.14) can be approximated by

$$A_h \mathbf{w}_h = \mathbf{g}_h - Q_m f_1(H_m) {Q_m}^* \mathbf{g}_h. \tag{5.15}$$

Approach 2. Let $f_2(z) = z^{-1}(1 - e^{-tz})$, then the system eq. (5.14) can be approximated by

$$\mathbf{w}_h = Q_m f_2(H_m) {Q_m}^* \mathbf{g}_h. \tag{5.16}$$

The matrix function $f_2(A_h)$ can be approximated by the Krylov subspace method with the same convergence properties of $f_1(A_h)$: as a matter of fact, the results of Theorem 5.9 holds if e^{-TA_h} is substituted with $f_2(A_h)$, as stated in [94]. Moreover, this second approach has the additional advantage of avoiding to solve the large linear system with the sparse matrix A_h.

Assuming that the systems in the approach is inverted by a linearly scaling algorithm, such as the multigrid, the overall computational costs of both formulations are dominated by computation of the Krylov subspace. The Arnoldi algorithm consists of an outer loop for $j = 1 : m$, where in total $m \ll N$ sparse matrix vector multiplications of the form $A_h \mathbf{g}_h$ are needed. Moreover, there is an orthogonalisation process which occurs at an inner loop for $i = 1 : j$, where the essential cost is due to a row-column vector multiplication. The overall cost of the Arnoldi iteration, exploiting the inherent sparsity of A_h, becomes

$$Cost_{Arnoldi} \approx \sum_{j=1}^{m} \left(C_d N + 4 \sum_{i=1}^{j} N \right) = \mathcal{O}(Nm^2).$$

If m is fixed *a priori* instead of following the scaling of Corollary 5.10, then the cost of the algorithm will grow linearly with N. A more rigorous analysis can be done by using the analysis in Section 5.2.2, where the optimal value of m for the approach 1 has been presented. Following the result of Corollary 5.10, we find that $m^2 = \mathcal{O}(T^2 h^{-2})$. Hence, using the relation

[4] For simplicity all the indices are skipped in this discussion.

Corrector problem	Modified Elliptic	Standard Elliptic
Error	$R^{-q-1/2}+h^s$	$R^{-1}+h^{2s}$
Computational cost	$R^{2+d}h^{-d-2}$	$R^d h^{-d}$
Computational cost (tol)	$tol^{-\frac{2d+4}{2q+1}-\frac{d+2}{s}}$	$tol^{-d-\frac{d}{2s}}$

Table 5.1 – Error and computational cost for the modified and standard homogenization approaches.

$N = R^d h^{-d}$ and the quasi-optimal choice $T = \mathcal{O}(R)$, the overall cost becomes

$$Cost_{Arnoldi} = \mathcal{O}(Nm^2) = \mathcal{O}(R^d T^2 h^{-d-2}) = \mathcal{O}(R^{d+2}h^{-d-2}).$$

In the same way, the computational cost of the standard elliptic upscaling approach is estimated as $\mathcal{O}(R^d h^{-d})$, see Table 5.1. The global errors, which are composed of the resonance and the discretization errors, are also reported in Table 5.1. The resonance error scales as $R^{-q-1/2}$ for the modified elliptic approach, see Theorem 5.1, while it decays as R^{-1} for the standard elliptic case, according to eq. (2.43). The discretization error is assumed to be of order $\mathcal{O}(h^s)$ in both cases. In order to derive the scaling of the cost with respect to the accuracy, we impose the global error to be smaller than a prescribed tolerance tol. So, for the modified elliptic case, we choose R and h such that $R^{-q-1/2} \approx tol$ and $h^s \approx tol$, while $R^{-1} \approx tol$ and $h^{2s} \approx tol$ for the standard elliptic case. Therefore, the modified elliptic approach has a lower cost to reach a certain tolerance tol when

$$\frac{d+2}{q+1/2} + \frac{d+2}{s} < d + \frac{d}{2s},$$

which is easily achieved by using filters with better regularity properties (large q), as well as high order numerical methods for the approximation of the elliptic PDE. In conclusion, the scaling of the computational cost in comparison to the tolerance for the standard elliptic, the parabolic and the modified elliptic methods is depicted in Figure 5.1 .

5.4 Numerical experiments

In this section, we provide examples in two dimensions to verify the theoretical results stated in Theorem 5.1. We illustrate the expected convergence rates by varying the regularity parameter q of the filters, in a periodic, smooth setting. Moreover, additional numerical tests are provided to show that the method performs equally well even when the regularity and structural assumptions of the theorem are violated. In particular, the test cases include a periodic medium, a discontinuous layered medium, a quasi-periodic medium, as well as a random medium. These results are discussed in separate subsections below. We compute the approximations of the homogenized tensor through the described modified elliptic cell problems on domains of increasing size, $R \in [1, 12]$, and calculate the Frobenius norm of the

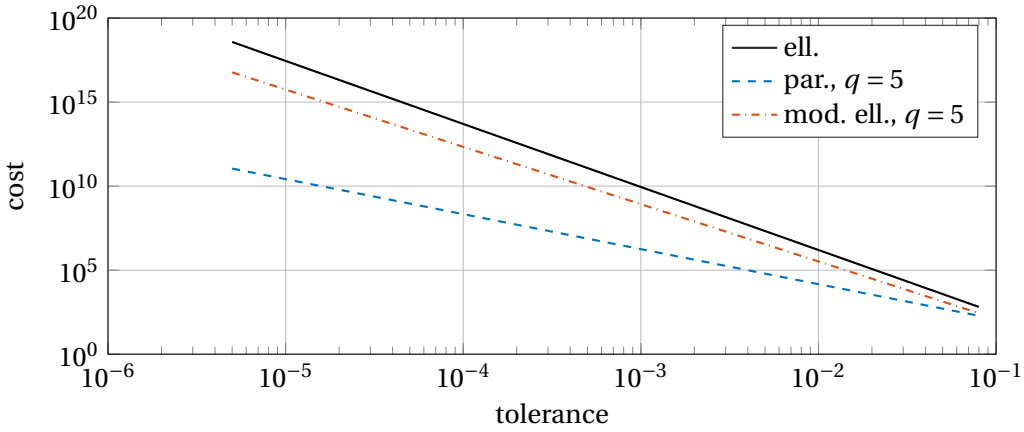


Figure 5.1 – Theoretical computational cost for $d = 3$, $\mathbb{P}_2$-FEM, $q = 5$.

difference between such approximations and the exact a^0. In the case of periodic coefficients whose homogenized value could not be known exactly (i.e., without discretization error) the reference value is computed by solving the standard elliptic corrector problem with $R = 1$ and periodic boundary conditions. In the random setting, for which an exact, computable formulation for the homogenized coefficients does not exists, we took as reference value for the homogenized tensor the one computed from the numerical approximation of the parabolic correctors over the largest domain $R_{max} = 12$.

To compute a numerical approximation of $a^{0,R,L,T}$, we use a Finite Difference (FD) discretization that allows us to compute the micro correctors through the second approach of Section 5.3, since the mass matrix is the identity.

In order to optimize the convergence rate of the error with respect to the sampling domain size R, we take the optimal values of Theorem 5.1 for the averaging domain size L ($K_L \subset K_R$) and for the final time T given by

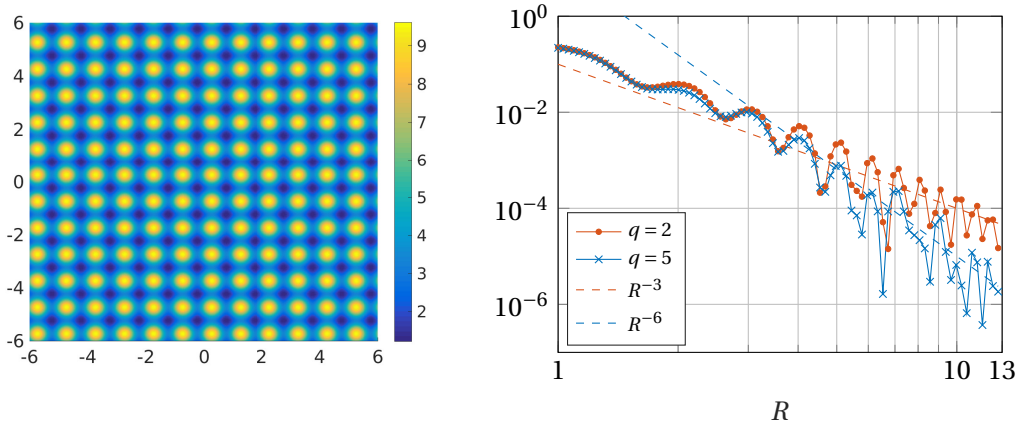
$$L = k_o R, \text{ and } T = \frac{\pi}{2\sqrt{d}}\sqrt{\frac{\alpha}{\beta}}(R - L),$$

where α, β are, respectively, the ellipticity and continuity constants for the tensor a. The oversampling ratio, $0 < k_o < 1$, and the order of filters, q, can be chosen freely.

5.4.1 A smooth periodic example

As our first example, we consider the following two-dimensional coefficient

$$a(x) = \prod_{j=1}^{2} \left(2.1 + \sin(2\pi x_j)\right) I, \tag{5.17}$$



(a) $a(x)$ of eq. (5.17)

(b) The resonance error.

Figure 5.2 – A two dimensional smooth medium.

where I is the 2×2 identity matrix, see the left picture in Figure 5.2 for a graphical representation of a. In this case, the homogenized coefficient is constant and given by[5]

$$a^0 = \left(2.1\sqrt{2.1^2 - 1}\right) I.$$

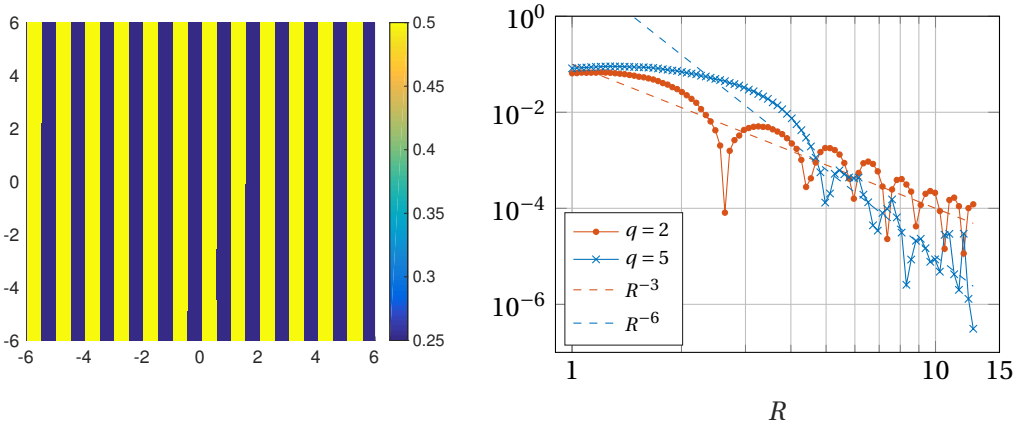
In Figure 5.2, the upscaling error $\|a^{0R,L,T} - a^0\|_F$ is shown for increasing values of R. The parameter values T and L are chosen optimally as stated in Theorem 5.1, with $k_0 = \frac{2}{3}$, $\alpha = \min_{x\in K} a(x)$, $\beta = \max_{x\in K} a(x)$. The number of basis functions in Arnoldi algorithm to approximate the right hand side is $m = \min(700, N^{1/d})$ (where N is the total number of degrees of freedom) for all values of R since the Arnoldi's error is typically much smaller than the rest of the errors. Two different kernels with $q = 2$ and $q = 5$ are used in the simulations. The cell-problem eq. (5.8) is approximated by a second order finite difference scheme with the stepsize $h = 1/120$. The numerical results show that the overall error is dominated by the filtering error even for moderate values of R, and that arbitrarily high convergence rates are obtained by using kernels with better regularity properties.

5.4.2 Discontinuous coefficients

The second example is a layered medium characterised by the coefficient

$$a(x) = \tilde{a}(x) I, \quad \text{with} \quad \tilde{a}(x) = \begin{cases} \frac{1}{2} & 0 \le x_1 < \frac{1}{2}, \\ \frac{1}{4} & \frac{1}{2} \le x_1 < 1. \end{cases}$$

[5]The diagonal component of a^0 can be computed by $a_{ii} = \left(\fint_0^1 \left(2.1 + \sin(2\pi x_i)\right)^{-1} dx_i\right)^{-1} \fint_0^1 2.1 + \sin(2\pi x_j)\, dx_j$ for $i \neq j$



(a) $\tilde{a}(x)$ of section 5.4.2

(b) The resonance error.

Figure 5.3 – A two dimensional periodic discontinuous medium

Such a choice is to test the generality of the method when the regularity assumption on the coefficient is relaxed. The exact homogenized coefficient is again constant and given by

$$a^0 = \begin{pmatrix} 1/3 & 0 \\ 0 & 3/8 \end{pmatrix}.$$

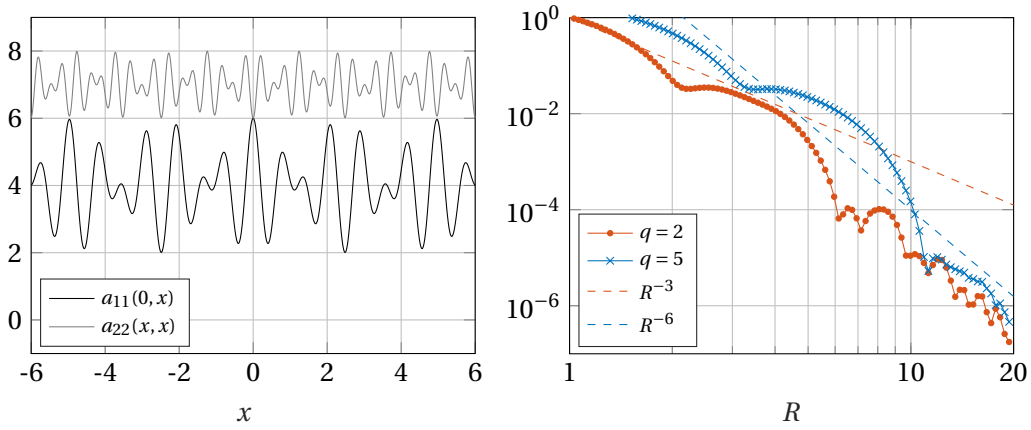
All the numerical parameters are chosen identical to those in example 1, with an obvious adaptation of α and β. Similar to example 1, higher order convergence rates are achieved upon using higher order kernels, showing the generality of the method also for problems in discontinuous media.

5.4.3 The quasi-periodic case

To test the applicability of the method beyond the periodic setting, we consider a quasi-periodic coefficient given by

$$a(x) = \begin{pmatrix} 4 + \cos(2\pi(x_1 + x_2)) + \cos(2\pi\sqrt{2}(x_1 + x_2)) & 0 \\ 0 & 6 + \sin^2(2\pi x_1) + \sin^2(2\pi\sqrt{2}x_2) \end{pmatrix}. \quad (5.18)$$

The very same coefficient has been used also in the elliptic approach proposed in [78]. In this paper, such a choice for the coefficient has been intentional as it allows for a comparison between the two methods. In this particular setting, the homogenized coefficient is not easy to compute and therefore the value of $a^{0,R,L,T}$ with the largest R is used instead of a^0 (similar to [78]). All the parameter values are chosen identical as in Section 5.4.1. Figure 5.4 shows a fast decay of the error down to 10^{-6} for moderate values of R, i.e., $R \approx 10$. It is worth mentioning that such an error tolerance is achieved only for $R \approx 40$ in the zero-order approach from [78].



(a) $\tilde{a}(x)$ of eq. (5.18)

(b) The resonance error.

Figure 5.4 – A two dimensional quasi-periodic medium

5.4.4 Random coefficients

As yet another example of a non-periodic medium, we construct a random medium as follows: we start by choosing a large computational grid, which corresponds to a discretization of the domain $K_{R_{\max}}$ with $R_{\max} = 40$. We then generate a sequence of uniformly distributed random variables taking values in the interval $[1,2]$, and assign these random numbers on each grid point. Next, we set a correlation length σ (here $\sigma = 0.25$ is chosen), and construct the random coefficient at each discretization point $x_i \in K_R$ (for a given $R < R_{\max}$) by taking the average of the generated random values associated to the points $x_j \in B_\sigma(x_i)$. Since, the interest here is not to study the statistical error, we compute only the error

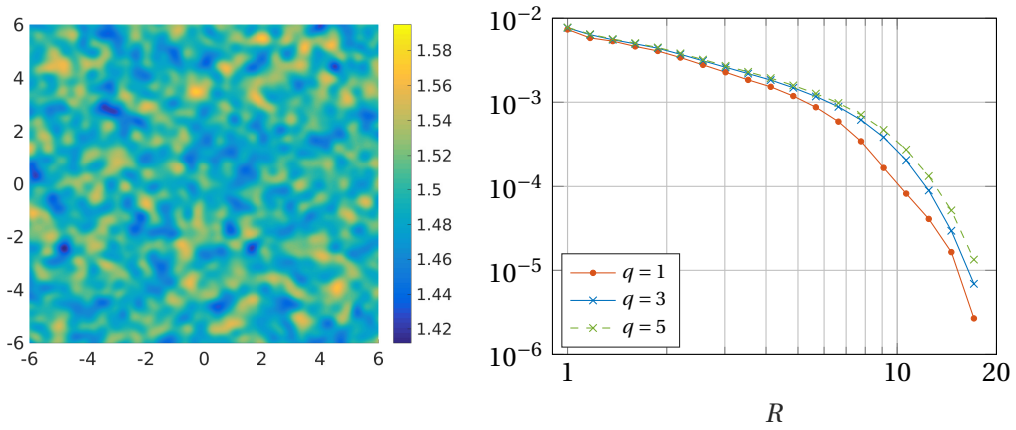
$$e_{BC} := \| a^{0,R,L,T} - a^{0,R_{\max},L,T} \|_F,$$

which sees the deterministic part of the overall error only. In Figure 5.5, the generated random coefficient along with the boundary error is depicted. All the parameter values except $h = 1/40$, $\alpha = 1$, and $\beta = 2$ are the same of Section 5.4.1. An exponential decay for the boundary error is observed for three different choices of filters, consistently with the fact that the observed error corresponds to the boundary error, and not the filtering error.

5.5 Conclusion

In this chapter we used the properties of the parabolic model of Chapter 4 to construct elliptic corrector problems to approximate the homogenized coefficients. Such corrector problems are characterised by the presence of an additional term, which is the solution of the parabolic model at time T, i.e. the action of the semigroup e^{-AT} on $g^i = \nabla \cdot (a\mathbf{e}_i)$.

Under the same regularity assumptions of Chapter 4, we can derive an upper bound of



(a) $a(x)$ (b) The upscaling error

Figure 5.5 – A two dimensional random medium

the resonance error for periodic coefficients. The error bound is composed of three terms depending on the averaging domain K_L, the truncation at time T and the mismatching boundary conditions on K_R. The first term decays with an arbitrary rate L^{-q}, while the other two show exponential decays, so their contribution is eventually negligible for sufficiently high values of R and T. Finally, by balancing all the terms, we show that the quasi-optimal parameter scaling is obtained for $L, T = \mathcal{O}(R)$.

Several numerical experiments supported the theoretical results. We tested the modified elliptic method also for coefficients not satisfying the assumptions used in the theory, e.g. discontinuous, quasi-periodic and stochastic coefficients. The results for these cases are in agreement with the theoretical bounds, showing that it could be possible to relax the regularity assumptions on the coefficients.

The additional term $e^{-AT}g^i$ in the corrector problem needs to be pre-calculated to derive a full space discretization. Many strategies are available for this task, also thanks to the double interpretation of the additional term. One strategy is based on the use of time-advancing schemes for the parabolic equation, up to time T. The exponential matrix $e^{-A_h T}$ can be computed exactly (up to finite arithmetic errors) only for small matrices, while in our case A_h is a large, sparse matrix. The use of standard/extended Krylov subspaces method allows to accurately approximate the exponential matrix, with a limited computational cost. The cost of this latter approach is compared to one of the standard elliptic and parabolic strategies. The parabolic one is the most efficient, provided that we use a sufficiently high order time integration scheme. The difference between the computational costs of the modified and standard elliptic methods is less striking, even though the modified one still performs better. The reason for such a small difference is mainly due to the assumption that the linear system can be solve in $\mathcal{O}(N)$ operations, which may not always be the case.

With this chapter we conclude the study of homogenization of periodic media, as we will consider the case of random media in Chapters 6 and 7.

6 Homogenization of diffusion problems in random media

In this chapter we consider the homogenization of multiscale equations with random coefficients. Let us consider the partial differential equation on the domain $D \subset \mathbb{R}^d$:

$$\begin{cases} -\nabla \cdot (a^\varepsilon(x,\omega)\nabla u^\varepsilon) = f & \text{in } D, \\ u^\varepsilon = 0 & \text{on } \partial D, \end{cases} \tag{6.1}$$

where $f \in H^{-1}(D)$ and $a^\varepsilon \in L^2(\Omega, L^\infty(D))^{d\times d}$ is symmetric, uniformly elliptic and bounded. The multiscale diffusion coefficient is the realization of a random field and it oscillates at the ε-scale, i.e.

$$a^\varepsilon(x,\omega) = a\left(\frac{x}{\varepsilon},\omega\right).$$

Under suitable assumptions, one can prove that $a^\varepsilon \xrightarrow{G} a^0$ in the limit for $\varepsilon \to 0$. Under the assumptions of *stationarity* and *ergodicity*, the sequence of random coefficients converges almost surely to a deterministic, constant tensor a^0, i.e., at the large scale, the heterogeneous random medium appears as a deterministic homogeneous medium.

The corrector equations to reconstruct a^0 can be derived heuristically through the standard asymptotic expansion procedure: we suppose that the solution u^ε satisfies the asymptotic expansion

$$u^\varepsilon(x,\omega) = u_0(x) + \varepsilon u_1\left(x,\frac{x}{\varepsilon},\omega\right) + \varepsilon^2 u_2\left(x,\frac{x}{\varepsilon},\omega\right) + o\left(\varepsilon^2\right). \tag{6.2}$$

The change of variables $y = x/\varepsilon$ leads to three differential problems at different orders of magnitude:

$$-\nabla_y \cdot \left(a\left(y,\omega\right)\nabla_y u_0\right) = 0, \qquad \left(\mathcal{O}\left(\varepsilon^{-2}\right)\right)$$

$$-\nabla_y \cdot \left(a\left(y,\omega\right)\nabla_y u_1\right) - \nabla_y \cdot \left(a\left(y,\omega\right)\nabla_x u_0\right) - \nabla_x \cdot \left(a\left(y,\omega\right)\nabla_y u_0\right) = 0, \qquad \left(\mathcal{O}\left(\varepsilon^{-1}\right)\right)$$

$$-\nabla_y \cdot \left(a\left(y,\omega\right)\nabla_y u_2\right) = f + \nabla_y \cdot \left(a\left(y,\omega\right)\nabla_x u_1\right) + \nabla_x \cdot \left(a\left(y,\omega\right)\nabla_y u_1\right) + \nabla_x \cdot \left(a\left(y,\omega\right)\nabla_x u_0\right). \tag{$\mathscr{O}(1)$}$$

Having assumed that u_0 does not depend on y, the first equation eq. ($\mathscr{O}\left(\varepsilon^{-2}\right)$) is directly satisfied. Moreover, by linearity, we can write the solution u_1 of eq. ($\mathscr{O}\left(\varepsilon^{-1}\right)$) as

$$u_1\left(x,y,\omega\right) = \sum_{i=1}^{d} \chi^i(y,\omega)\frac{\partial u_0}{\partial x_i}(x),$$

where χ^i solves

$$-\nabla_y \cdot \left(a\left(y,\omega\right)\left(\nabla_y \chi^i + \mathbf{e}_i\right)\right) = 0. \tag{6.3}$$

At this point, in the periodic setting one would solve eq. ($\mathscr{O}\left(\varepsilon^{-1}\right)$), plug the solution into eq. ($\mathscr{O}(1)$) and, by imposing the solvability conditions, find that the leading order term u_0 satisfies the homogenized equation eq. (2.3) with a^0 defined as in eq. (2.8). In the stochastic setting this procedure is not as straight-forward as in the periodic case because it is still unclear in which sense we solve the corrector equation eq. (6.3) and we do not have any solvability condition yet for eq. ($\mathscr{O}(1)$). Then, the problem reduces to the analysis of the general stochastic partial differential equation:

$$-\nabla \cdot \left(a\left(y,\omega\right)\nabla \chi^i\right) = g(y,\omega),$$

with a and g stationary random fields. In the coming sections we will provide the abstract framework to derive the solvability conditions for the stochastic corrector problem.

Outline

This chapter is structured as follows: in Section 6.1 we provide the mathematical framework in which stochastic homogenization is studied. The first qualitative results in stochastic homogenization are due to the pioneering works [104, 123] and extensions to non-linear equations were later provided by [46, 47]. In Section 6.2 the main results in quantitative stochastic homogenization is described. Besides being interesting *per se* quantitative results are useful in deriving efficient numerical methods to estimate the homogenized coefficients. Some of these methods are discussed in Section 6.3.

6.1 The mathematical framework

In this section we explain our notation and provide a precise formulation of the proposed corrector problems in the stochastic setting. Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space and let $\mathbb{R}^d$ be endowed with the Borel σ-algebra $\mathscr{B}$. We denote as *random variables* all the measurable functions $X : (\Omega,\mathscr{F},\mathbb{P}) \mapsto \mathbb{R}^d$ and as *random fields* all the measurable functions $f : \left(\mathbb{R}^d,\mathscr{B}\right) \times (\Omega,\mathscr{F},\mathbb{P}) \mapsto \mathbb{R}$ and we define stationarity as follows.

Definition 6.1. *A random field $f : \mathbb{R}^d \times \Omega \mapsto \mathbb{R}$ is said* stationary *if, for any $h \in \mathbb{R}^d$ and any*

points $y_1, \dots, y_n \in \mathbb{R}^d$,

$$\big(f(y_1,\omega),\dots,f(y_n,\omega)\big) \quad \textit{and} \quad \big(f(y_1+h,\omega),\dots,f(y_n+h,\omega)\big)$$

have the same joint distribution.

Definition 6.2. *A* translation group *(or d-dimensional dynamical system) is a family of invertible measurable maps, indexed by $x \in \mathbb{R}^d$, $\tau_x : \Omega \mapsto \Omega$ such that*

i) $\tau_{x+y} = \tau_x \tau_y$, $\tau_0 = Id$;

ii) τ_x *preserves the measure* $\mathbb{P}$: $\mathbb{P}(\tau_x F) = \mathbb{P}(F)$, *for any* $F \in \mathscr{F}$ *and any* $x \in \mathbb{R}^d$;

iii) for any random variable $X : \Omega \mapsto \mathbb{R}$, *the function* $X(\tau_x a(\cdot))$ *is* $\mathbb{R}^d \times \Omega$ *measurable with respect to the product* σ*-algebra.*

Then, Definition 6.1 can be rephrased as

Proposition 6.3. *A random variable $f(y,\omega)$ is stationary if and only if, for any $h \in \mathbb{R}^d$,*

$$f(y,\tau_h\omega) = f(y+h,\omega), \quad \textit{a.e. } y \in \mathbb{R}^d, \mathbb{P}\textit{-a.s.}$$

We now introduce the concept of ergodicity and explain how this property is used in stochastic homogenization.

Definition 6.4. *A measurable set $F \in \mathscr{F}$ is called* invariant *if $\tau_x F \subset F$ and a random variable $X : \Omega \mapsto \mathbb{R}$ is called* invariant *if $X(\tau_x\omega) = X(\omega)$ almost everywhere in Ω.*

A translation group τ_x is called ergodic *if the only invariant sets F have either $\mathbb{P}(F) = 0$ or $\mathbb{P}(F) = 1$ or, alternatively, if all invariant random variables are constant almost everywhere in Ω.*

If a random field $f(\cdot,\omega)$ is generated by a stationary extension $f(0,\tau_x\omega)$ through an ergodic dynamical system, then we call f an ergodic random field. An important tool of ergodic theory is the Birkhoff's ergodic theorem, see Theorem 6.5, which establishes the equivalence between the probabilistic and the spatial averages of a stationary ergodic random field f.

Let $f \in L^1_{loc}(\mathbb{R}^d)$. We denote the mean value of f as $M(f)$ given by

$$M(f) := \lim_{R\to\infty} \fint_{K_R} f(x)\,dx, \tag{6.4}$$

where, for simplicity, we considered $K_R := (-R/2, R/2)^d$ as integration domain, but one can equally use any domain $Q_R = \{x \in \mathbb{R}^d, x/R \in Q\}$ that is the homothetic dilatation of a given

measurable set $Q \subset \mathbb{R}^d$. Let us define $f_\varepsilon(x) = f(x/\varepsilon)$; we say that $f_\varepsilon \in L^p_{loc}\left(\mathbb{R}^d\right)$ has the *mean value property* if

$$f_\varepsilon \overset{\varepsilon\to 0}{\rightharpoonup} M(f), \text{ weakly in } L^p_{loc}(\mathbb{R}^d). \tag{6.5}$$

The Birkhoff's ergodic theorem states that, for stationary random variables defined through ergodic translation group, the spatial and probability averages are identical.

Theorem 6.5 (Birkhoff's ergodic theorem)**.** *Let $f \in L^p(\Omega)$, $p \geq 1$. Then, for almost all $\omega \in \Omega$, the realization $f(\tau_x\omega)$ possesses the mean value property. Moreover, the mean value $M(f(\tau_x\omega))$ considered as a function of $\omega \in \Omega$ is invariant and*

$$\mathbb{E}\left[f\right] := \int_\Omega f(\omega)\, d\mathbb{P} = \int_\Omega M(f(\tau_x\omega))\, d\mathbb{P}.$$

In particular, if τ_x is ergodic, then

$$\mathbb{E}\left[f\right] = M(f(\tau_x\omega)), \quad \mathbb{P}\text{-a.s.}$$

With the space Ω we aim to take into account all possible realizations of a random coefficient field. Therefore, it is convenient to identify the space Ω with the set of all possible coefficients $a(x)$ that allow the solvability of eq. (6.1), i.e.

$$\Omega = \mathscr{M}\left(\alpha, \beta\right).$$

Then, we endow Ω with Borel set-indexed family of σ-algebras $\mathscr{F}_B$ generated by

$$\left\{ a \in \Omega \mapsto \int_{\mathbb{R}^d} a_{ij}(x)\varphi(x)\, dx, \varphi \in C_c^\infty(B), i, j = 1, \ldots, d \right\}.$$

The largest of these σ-algebras is denoted by $\mathscr{F}$; thus, two realizations $\check{a}$ and $\hat{a}$ are considered identical if they differ only on a set $B \subset \mathbb{R}^d$ of vanishing measure.

Finally, we take the translation group τ_x defined by:

$$\tau_x a(y) := a(x + y).$$

Remark 6.6. *The random tensor field $x \mapsto a(x)$ is stationary.*

The basis to study the correctors in stochastic homogenization is given by the extension of the Weyl decomposition of square-integrable functions into their potential and solenoidal parts. However, in the present situation we deal with stationary random vector fields. Stationary random fields can be represented by their values at a given point (for instance, the origin) and, then, extended to the whole $\mathbb{R}^d$ through the mapping $f(x, \omega) = f(0, \tau_x\omega)$. From now on, with a slight abuse of notation we will denote in the same way the random variable $f(\tau_x\omega)$ and its stationary extension $f(x, \omega)$. Let U_x be a d-parameters strongly continuous group of unitary

operators in $L^2(\Omega)$ associated with τ_x :

$$\left(U_x f\right)(\omega) = f(\tau_x \omega), \quad f \in L^2(\Omega).$$

The infinitesimal generator of U_x along the i-th direction are the closed and densely defined operators ∂^i_ω:

$$\left(\partial^i_\omega f\right)(\omega) := \lim_{h\to 0} \frac{f(\tau_{h\mathbf{e}_i}\omega) - f(\omega)}{h}.$$

The domains $\mathscr{D}_i$ of the operators ∂^i_ω are dense in $L^2(\Omega)$, and so is their intersection, see [51] which is denoted by:

$$\mathscr{H}^1 = \bigcap_{i=1}^{d} \mathscr{D}_i. \tag{6.6}$$

Thus, we can define the *stochastic* gradient and divergence for functions $f, g_1, \ldots, g_d \in \mathscr{D}$:

$$\nabla_\omega f = \left(\partial^1_\omega f, \ldots, \partial^d_\omega f\right), \quad \text{and} \quad \nabla_\omega \cdot \mathbf{g} = \sum_{i=1}^{d} \partial^i_\omega g_i.$$

Since the group U_x is unitary, the infinitesimal generators ∂^i_ω are skew-symmetric operators, thus,

$$\mathbb{E}\left[\partial^i_\omega f g\right] = -\mathbb{E}\left[f \partial^i_\omega g\right], \quad \text{and in particular,} \quad \mathbb{E}\left[\partial^i_\omega f\right] = 0 \quad \text{for } f, g \in \mathscr{D}.$$

Moreover, the stochastic derivatives of a random variable are related to the spatial derivative of its stationary extension:

$$\left(\partial^i_\omega f\right)(\tau_x \omega) = \frac{\partial}{\partial x_i}\left(f(\tau_x \omega)\right).$$

So, we can say that $f(\cdot) \in \mathscr{H}^1$ if and only if $f(\tau_x \cdot) \in H^1(D)$ $\mathbb{P}$-a.s.

Let us introduce the spaces of, respectively, *potential* and *solenoidal* random fields:

$$\mathbf{L}^2_{pot}(\Omega) = \overline{\left\{\mathbf{f} \in \mathbf{L}^2(\Omega) : \mathbf{f} = \nabla_\omega u, \text{ for some } u \in \mathscr{H}^1\right\}},$$

$$\mathbf{L}^2_{sol}(\Omega) = \overline{\left\{\mathbf{f} \in \bigotimes_{i=1}^{d} \mathscr{D}_i : \nabla_\omega \cdot \mathbf{f} = 0\right\}},$$

where the overline symbol means the closure in $\mathbf{L}^2(\Omega)$. From the closedness of the operator ∇^i_ω, it follows that

$$\mathbb{E}[\mathbf{f}] = 0 \text{ for all } \mathbf{f} \in \mathbf{L}^2_{pot}(\Omega) \quad \text{and} \quad \mathbb{E}\left[f_i \partial^j_\omega g\right] = \mathbb{E}\left[f_j \partial^i_\omega g\right] \text{ for } i, j = 1, \ldots, d.$$

We have now provided a setting in which the auxiliary problem eq. (6.3) can be solved. One can write the weak form of eq. (6.3) as: Find $\psi^i \in \mathbf{L}^2_{pot}(\Omega)$ such that

$$\mathbb{E}\left[\varphi \cdot a\left(\psi^i + \mathbf{e}_i\right)\right] = 0, \quad \forall \varphi \in \mathbf{L}^2_{pot}(\Omega). \tag{6.7}$$

The auxiliary stochastic problem eq. (6.7) has a unique solution, as a consequence of the Lax-Milgram theorem on the Hilbert space $\mathbf{L}^2_{pot}(\Omega)$. We underline that the bilinear form in this case is

$$\begin{aligned} B: \quad \mathbf{L}^2_{pot}(\Omega) \times \mathbf{L}^2_{pot}(\Omega) &\mapsto \mathbb{R} \\ (\varphi, \psi) &\mapsto \mathbb{E}\left[\varphi \cdot a(x)\psi\right], \end{aligned}$$

and the coercivity condition is satisfied thanks to the ellipticity of $a(\cdot)$:

$$\alpha \left\|\varphi\right\|^2_{\mathbf{L}^2(\Omega)} \leq \mathbb{E}\left[\varphi \cdot a(x)\varphi\right].$$

This bilinear form does not involve the weak (stochastic) gradients of $\mathscr{H}^1$-functions and the reason is that the Poincaré inequality is not valid in this setting. In contrast to the periodic case, we cannot write the corrector problem in a way to ensure the existence and uniqueness of the correctors χ^i as defined in eq. (6.3). As a matter of fact, the correctors χ^i cannot be uniquely defined, as they are defined up to additive constants. In the following theorem we summarize the well-posedness of eq. (6.7) and explain how one can select in a unique way the corrector χ^i.

Theorem 6.7 ([123])**.** *Let $a(\cdot)$ be a stationary and ergodic tensor field. Then, for any direction $\mathbf{e}_i$, $i = 1, \dots, d$, there exists a unique ψ^i that satisfies*

$$\mathbb{E}\left[\varphi \cdot a(x)\left(\psi^i + \mathbf{e}_i\right)\right] = 0, \quad \forall \varphi \in \mathbf{L}^2_{pot}(\Omega).$$

Moreover, there exist uniquely defined processes $\chi^i(x,\omega) \in H^1_{loc}(\mathbb{R}^d, L^2(\Omega))$ such that

$$\chi^i(0,\omega) = 0, \text{ and } \quad \frac{\partial \chi^i}{\partial x_j}(x,\omega) = \psi^i_j(\tau_x \omega),$$

so that the gradients of χ^i are stationary, but not the functions themselves. Additionally, the correctors χ^i grow sub-linearly at infinity: for every compact set $K \subset \mathbb{R}^d$,

$$\lim_{R\to\infty} \sup_{x\in K} \mathbb{E}\left[\left(\frac{\chi(Rx)}{R}\right)^2\right] = 0.$$

Remark 6.8. *In the general case, the corrector χ^i is not statistically stationary. It is possible to prove (see, e.g. [19, Chapter 4] or [83]) that, in dimensions $d > 2$, there exists a stationary corrector χ^i, uniquely defined by the condition $\mathbb{E}\left[\chi^i\right] = 0$.*

The G-limit of the stationary ergodic random tensor $a^\varepsilon(x)$ of eq. (6.1) has components:

$$a^0_{ij} = \mathbb{E}\left[\mathbf{e}_i \cdot a(x)\left(\psi^j + \mathbf{e}_j\right)\right] = \fint_{\mathbb{R}^d} \mathbf{e}_i \cdot a(x)\left(\nabla \chi^j(x,\omega) + \mathbf{e}_j\right) dx, \tag{6.8}$$

where the second identity follows from the Birkhoff's ergodic theorem. The first identity can also be written as

$$a^0_{ij} = \mathbb{E}\left[\left(\psi^i + \mathbf{e}_i\right) \cdot a(x)\left(\psi^j + \mathbf{e}_j\right)\right].$$

Finally, it is possible to prove the convergence $u^\varepsilon \to u^0$, for $\varepsilon \to 0$.

Theorem 6.9 (Weak convergence, [123]). *Let $a(\cdot)$ be a stationary ergodic random tensor field. Let $u^\varepsilon \in H_0^1(D)$ solve eq.* (6.1) *with $f \in H^{-1}(D)$. Then, $u^\varepsilon \rightharpoonup u^0$ weakly in $H_0^1(D)$, where $u^0 \in H_0^1(D)$ solves*

$$\begin{cases} -\nabla \cdot \left(a^0 \nabla u^0\right) = f & \text{in } D, \\ u^0 = 0 & \text{on } \partial D. \end{cases} \tag{6.9}$$

Theorem 6.10 (Strong convergence, [123]). *Let $u^\varepsilon \in H_0^1\left(D, L^2(\Omega)\right)$ be the solution of eq.* (6.1), *and let $u^0 \in H_0^1(D)$ be the solution of eq.* (6.9). *Then*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left\| u^\varepsilon - u^0 \right\|_{L^2(D)}\right] = 0, \text{ and}$$

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left\| u^\varepsilon - u^0 - \varepsilon \sum_{k=1}^{d} \chi^k\left(\frac{\cdot}{\varepsilon}\right) \frac{\partial u^0}{\partial x_k} \right\|_{H_0^1(D)}\right] = 0.$$

6.2 From qualitative to quantitative results in homogenization of random media

Up to now, we only have shown that the solution u^ε converges to the solution u^0 of the homogenized problem, but we have not discussed the rate of convergence, i.e. whether the difference $\left\| u^\varepsilon - u^0 \right\|_{L^2}$ can be bounded as a function of ε. Quantifying such bounds is the goal of *quantitative* stochastic homogenization. In connection to this problem (or, as an intermediate step to solve it) one can also investigate the error rate in the approximation of a^0. Besides the theoretical interest, this is important also from the numerical point of view. Indeed, the stochastic auxiliary problems are stated in an abstract probability space, thus it does not give any practical recipe for constructing or approximating the effective coefficients. This is in contrast to the periodic case for which many efficient numerical homogenization procedures are available. The rate of convergence of the homogenization procedure is therefore a key tool to derive convergence rates for multiscale numerical methods.

An important result in the derivation of convergence rate in stochastic homogenization was obtained in [135], where boundary value problems for a second order divergence form operator were studied and, under proper mixing condition, polynomial bounds for the convergence rate were achieved.

A widely used mathematical technique to derive *a priori* bounds on the correctors and, consequently, on the approximations of a^0 is the addition of a zero-th order term in the corrector equation:

$$\frac{1}{T}\chi_T^i - \nabla_\omega \cdot \left(a(x)\left(\nabla_\omega \chi_T^i + \mathbf{e}_i\right)\right) = 0. \tag{6.10}$$

This problem has a unique solution in $\chi_T^i \in \mathscr{H}^1$, differently to the problem eq. (6.7). The random variable χ_T^i can, then, generate a stationary random field that satisfies the spatial

version of eq. (6.10)[1]: Find $\chi_T^i \in L^2\left(\Omega, H^1_{loc}(\mathbb{R}^d)\right)$ such that

$$\mathbb{E}\left[\int_{\mathbb{R}^d} \frac{1}{T}\chi_T^i \varphi + \nabla\varphi \cdot a(x)\left(\nabla\chi_T^i + \mathbf{e}_i\right) dx\right] = 0, \quad \forall \varphi \in L^2\left(\Omega, C_0^\infty\left(\mathbb{R}^d\right)\right). \tag{6.11}$$

In contrast, the correctors χ^i of eq. (6.3) are not stationary. Since the modified correctors χ_T^i exist and are unique, they can be used to define the first order approximation:

$$u^\varepsilon(x) = u^0(x) + \varepsilon \sum_{i=1}^d \chi_T^i\left(\frac{x}{\varepsilon}\right)\frac{\partial u^0}{\partial x_i}(x) + o(\varepsilon).$$

Yurinskii [135] provided the first result in quantitative homogenization by proving that there exist $C, \gamma > 0$ such that

$$\left\|\nabla_\omega \chi_T^i - \psi^i\right\|_{L^2(\Omega)} \leq CT^{-\gamma} \quad \text{and} \quad \left|\mathbb{E}\left[\mathbf{e}_i \cdot a(x)\left(\nabla_\omega \chi_T^j + \mathbf{e}_j\right)\right] - a_{ij}^0\right| \leq CT^{-\gamma},$$

where γ depends on the dimension d and on the exponent of the *uniformly strong intermixing condition*:

$$\left|\mathbb{E}\left[\xi\eta\right] - \mathbb{E}\left[\xi\right]\mathbb{E}\left[\eta\right]\right| \leq Cr^{-\gamma_1}\mathbb{E}\left[\xi^2\right]^{\frac{1}{2}}\mathbb{E}\left[\eta^2\right]^{\frac{1}{2}}, \tag{6.12}$$

where ξ is a $\mathscr{F}(A)$-measurable function, η is a $\mathscr{F}(B)$-measurable function, $r = \inf_{x\in A, y\in B}\left|x-y\right|$ and $A, B \subset \mathbb{R}^d$. This condition is satisfied, for instance, by random chequerboard structures. This result allows to derive an *a priori* convergence bound to the homogenized limit:

$$\mathbb{E}\left[\left\|u^\varepsilon - u^0 - \varepsilon\sum_{i=1}^d \chi_T^i\left(\frac{\cdot}{\varepsilon}\right)\frac{\partial u^0}{\partial x_i}\right\|_{H^1(D)}\right] \leq C\varepsilon^{\gamma_2},$$

upon choosing $T = \varepsilon^{-2+\delta}$, for sufficiently small $\delta > 0$.

Neither the auxiliary problem eq. (6.10) nor its stationary extension can be solved directly by numerical methods, since the former is posed over the abstract space $\mathscr{H}^1$ and the latter over the whole space $\mathbb{R}^d$. Therefore, the discrepancy in various cut-off approximation procedures must be considered. In numerical applications, under the same intermixing conditions as above and by Green's function estimates, [35] proved *a priori* bounds on the approximation error for the auxiliary problem eq. (6.10) when it is posed over a bounded domain with homogeneous Dirichlet, homogeneous Neumann or periodic BCs are proved in [35]. The bounds for the cut-off error in eq. (6.10) can then be used to prove

$$\mathbb{E}\left[\left\|a^{0,R} - a^0\right\|_F^2\right]^{\frac{1}{2}} \leq CR^{-\gamma}, \tag{6.13}$$

[1] We denote in the same way the random variable and its stationary extension in order to keep the notation simple.

where $a^{0,R}$ is computed through the cut-off auxiliary equation

$$-\nabla\cdot\left(a(x,\omega)\left(\nabla\chi_R^i+\mathbf{e}_i\right)\right)=0 \quad \text{in } K_R \quad + \text{ BCs.} \tag{6.14}$$

The error estimate eq. (6.13) is proved upon choosing $T=R^{2-\delta}$, for sufficiently small $\delta>0$. We remark that, in the analysis, one has to "pass through" the penalized equation eq. (6.10) on cut-off domains K_R because the gradient of the correctors, $\nabla\chi_R^i$, (and the correctors as well) are not statistically stationary, which is a handicap from the analytical point of view. *A priori* bounds similar to eq. (6.13), but with explicit exponents, were derived in [55] to estimate the HMM resonance error for random media.

In order to simplify the analysis, several authors focused on the problem of stochastic homogenization over discrete networks with random conductances. This can be seen as a Finite Difference approximation of eq. (6.11) and it has several advantages in comparison to its continuous version. For example, one can define the conductances over the edges as i.i.d. random variables, which directly entails stationarity and ergodicity of the coefficients. In [81, 82], the authors consider the regularised problem eq. (6.11) over a $\mathbb{Z}^d$ lattice with random conductances defined on the edges $e\in E\subset\mathbb{Z}^d\times\mathbb{Z}^d$, which reads: Find $\chi_T^i:\mathbb{Z}^d\mapsto\mathbb{R}$ such that

$$\frac{1}{T}\chi_T^i(x)+\sum_{y\in\mathbb{Z}^d,|x-y|=1}a(e)\left(\chi_T^i(x)-\chi_T^i(y)+1\right)=0, \tag{6.15}$$

where e is the edge connecting x and y. In order to be practically computable, the model needs to be defined over a bounded network of size R, but, thanks to the exponential decay of the related Green's function, the cut-off error decays exponentially with respect to $\frac{|R-L|}{\sqrt{T}}$ and it is thus negligible for sufficiently large R, in the regime $\sqrt{T}\lesssim R-L$. Thus, the whole-domain problem is analysed, knowing that it can be approximated with arbitrary accuracy orders by bounded domain solutions. For the estimation of the error, the total mean square error is first split into two contributes, a variance term (the *statistical* error) and a bias term (the *systematic* error):

$$\begin{aligned}\mathbb{E}\left[\left|\sum_e\left(\mathbf{e}_i+\nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j+\nabla\chi_T^j\right)\mu_L-a_{ij}^0\right|^2\right]^{\frac{1}{2}}\leq&\operatorname{Var}\left[\sum_e\left(\mathbf{e}_i+\nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j+\nabla\chi_T^j\right)\mu_L\right]^{\frac{1}{2}}\\&+\left|\mathbb{E}\left[\left(\nabla\chi_T^i-\nabla\chi^i\right)\cdot a(e)\left(\nabla\chi_T^j-\nabla\chi^j\right)\right]\right|,\end{aligned}$$

where μ_L is the sampling on the vertices of $\mathbb{Z}^d$ of a smooth averaging function vanishing outside $K_L\cap\mathbb{Z}^d$ and such that $\sum_x\mu_L=1$. The main result of [81] is the proof of the following inequality: there exist an exponent $q(\alpha,\beta)>0$ and a constant $C>0$, independent of L and T, such that:

$$\operatorname{Var}\left[\sum_e\left(\mathbf{e}_i+\nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j+\nabla\chi_T^j\right)\mu_L\right]\leq C\begin{cases}\left(L^{-2}+T^{-2}\right)(\log T)^q & \text{for } d=2,\\ L^{-d}\left(1+\frac{L}{T}\right) & \text{for } d>2;\end{cases} \tag{6.16}$$

in the regime of interest, $L \lesssim T \lesssim L^2$, the two bounds can be simplified as $L^{-2}(\log T)^q$ and L^{-d}, respectively. The proof of eq. (6.16) is based on the use of a *spectral gap* estimate to control the variance of random variables $\zeta \in L^2(\Omega)$:

$$\mathrm{Var}\,[X] \leq C \sum_e \mathbb{E}\left[\left(\frac{\partial \zeta}{\partial a(e)}\right)^2\right].$$

This type of estimate can be seen as a sort of Poincaré inequality with mean value zero with respect to the infinite product measure that describes the distribution of the coefficients. In the discrete setting, the spectral gap inequality is a direct consequence of the independence of the conductances values (see Lemma 2.3 in [81]). An estimate for the systematic error is derived in the companion paper [82]: there exist an exponent $q(\alpha,\beta) > 0$ and a constant $C > 0$, independent of T, such that:

$$\left|\mathbb{E}\left[\left(\nabla\chi_T^i - \nabla\chi^i\right)\cdot a(e)\left(\nabla\chi_T^j - \nabla\chi^j\right)\right]\right| \leq C \begin{cases} T^{-1}(\log T)^q & \text{for } d=2, \\ T^{-3/2} & \text{for } d=3, \\ T^{-2}\log T & \text{for } d=4, \\ T^{-2} & \text{for } d>4. \end{cases} \tag{6.17}$$

Hence, in the regime $T \sim L^2$, the total mean square error for the homogenization of random network can be bounded by:

$$\mathbb{E}\left[\left|\sum_e \left(\mathbf{e}_i + \nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j + \nabla\chi_T^j\right)\mu_L - a_{ij}^0\right|^2\right]^{\frac{1}{2}} \begin{cases} L^{-1}(\log L)^q & \text{for } d=2, \\ L^{-d/2} & \text{for } 3 \leq d \leq 7, \\ L^{-4}\log L & \text{for } d=4, \\ L^{-4} & \text{for } d>8. \end{cases}$$

Further results can be found in [80], whose main findings is the proof of an optimal decay in time of the semigroup associated with the corrector problem (i.e. of the generator of the process called "random environment as seen from the particle"). As a corollary the existence of stationary correctors (in dimensions $d > 2$) is recovered and new optimal estimates for the penalized correctors (in dimensions $d \geq 2$) are proved. These convergence rates were confirmed numerically in [61].

In [83] the results of [80, 81, 82] are adapted to the *continuum* context. Like in the discrete case, the mean square error for the approximation of the homogenized coefficients by the zero-th order regularised corrector problem eq. (6.10) is split into a statistical and systematic error terms:

$$\mathbb{E}\left[\left|\int_{K_L}\left(\mathbf{e}_i + \nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j + \nabla\chi_T^j\right)\mu_L - a_{ij}^0\right|^2\right]^{\frac{1}{2}} \leq \underbrace{\mathrm{Var}\left[\int_{K_L}\left(\mathbf{e}_i + \nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j + \nabla\chi_T^j\right)\mu_L\right]^{\frac{1}{2}}}_{\text{statistical error}}$$

$$+\underbrace{\left|\mathbb{E}\left[\left(\nabla\chi_T^i-\nabla\chi^i\right)\cdot a(e)\left(\nabla\chi_T^j-\nabla\chi^j\right)\right]\right|}_{\text{systematic error}}.$$

As in the discrete case, we can neglect the error due to the boundary conditions because it converges with an exponential rate in $\frac{|R-L|}{\sqrt{T}}$. The systematic error is controlled by the same upper bound of the discrete case, eq. (6.17):

$$\left|\mathbb{E}\left[\left(\nabla\chi_T^i-\nabla\chi^i\right)\cdot a(e)\left(\nabla\chi_T^j-\nabla\chi^j\right)\right]\right|\leq C\begin{cases}T^{-1} & \text{for } d=2,\\ T^{-3/2} & \text{for } d=3,\\ T^{-2}\log T & \text{for } d=4,\\ T^{-2} & \text{for } d>4.\end{cases}$$

and the proof follows from the spectral properties of the operator $-\nabla_\omega\cdot(a(x)\nabla_\omega)$. On the other hand, the statistical error can be bounded as

$$\mathrm{Var}\left[\sum_e\left(\mathbf{e}_i+\nabla\chi_T^i\right)\cdot a(e)\left(\mathbf{e}_j+\nabla\chi_T^j\right)\mu_L\right]\leq C\begin{cases}L^{-2}\log\left(2+\frac{\sqrt{T}}{L}\right) & \text{for } d=2,\\ L^{-d} & \text{for } d>2.\end{cases}$$

This bound can be proved by using a spectral gap inequality, which, in the continuum context, takes the form:

Definition 6.11. *A probability space $(\Omega,\mathscr{F},\mathbb{P})$ satisfies the Spectral Gap if there exists $\rho>0$ and $\ell<\infty$ such that for all random variables $\zeta\in L^2(\Omega)$ we have*

$$\mathrm{Var}\left[\zeta\right]\leq\frac{1}{\rho}\int_{R^d}\mathbb{E}\left[\left(\operatorname*{osc}_{a|_{B_\ell(z)}}\zeta\right)^2\right]dz, \tag{SG}$$

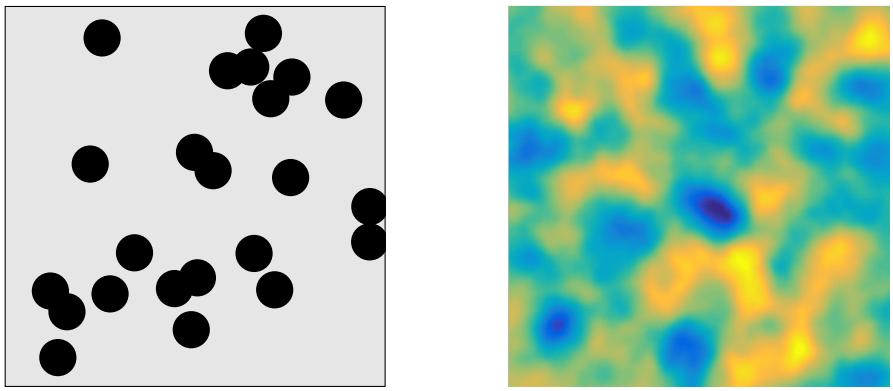
where $\operatorname*{osc}_{a|_{B_\ell(z)}}\zeta$ denotes the oscillation of the random field ζ with respect to the values of $a(x)$ on the ball $B_\ell(z)$:

$$\begin{aligned}\left(\operatorname*{osc}_{a|_{B_\ell(z)}}\zeta\right)(a(x))&=\left(\sup_{a|_{B_\ell(z)}}\zeta\right)(a(x))-\left(\inf_{a|_{B_\ell(z)}}\zeta\right)(a(x))\\&=\sup\left\{\zeta(\tilde{a}):\tilde{a}\in\Omega,\tilde{a}|_{\mathbb{R}^d\setminus B_\ell(z)}=a|_{\mathbb{R}^d\setminus B_\ell(z)}\right\}\\&\quad-\inf\left\{\zeta(\tilde{a}):\tilde{a}\in\Omega,\tilde{a}|_{\mathbb{R}^d\setminus B_\ell(z)}=a|_{\mathbb{R}^d\setminus B_\ell(z)}\right\}.\end{aligned}$$

The condition eq. (SG) is stronger than ergodicity, as stated below.

Lemma 6.12 (Lemma 2.3 in [83])**.** *If $(\Omega,\mathscr{F},\mathbb{P})$ satisfies eq.* (SG)*, then the translation group τ_x is ergodic.*

The assumption that the probability space satisfies a spectral gap inequality can be by-passed by taking stronger assumptions such as the mixing condition of eq. (6.12), or the unit range of



(a) Poisson random inclusions. (b) Smoothed Gaussian noise

Figure 6.1 – Two examples of random media with finite range of dependence.

dependence, as in [19]:

> $\mathscr{F}_A$ and $\mathscr{F}_B$ are $\mathbb{P}$-independent for every pair $A, B \subset \mathbb{R}^d$
> of Borel subsets such that $\operatorname{dist}(A,B) \geq 1$.

Classes of random coefficient fields that satisfies the unit range of dependence (and thus the spectral gap inequality) are, for example, the Poisson random inclusions (where spherical inclusions are randomly placed in a uniform matrix) or the smoothed Gaussian noise, depicted in Figure 6.1. The main drawback of the corrector problem eq. (6.11) is that the reconstruction of a^0 yields an error that saturates at $T^{-d/2}$ (T^{-2} for dimension $d \geq 4$). In order to reduce this drawback, *Richardson iterates* are proposed in [79, 84]. They are defined as:

$$\begin{aligned} \chi^i_{T,1} &= \chi^i_T, \\ \chi^i_{T,k+1} &= \frac{1}{2^k - 1}\left(2^k \chi^i_{2T,k} - \chi^i_{T,k}\right). \end{aligned}$$

This allows to reduce the systematic error to

$$\left|\mathbb{E}\left[\left(\nabla\chi^i_{T,k} - \nabla\chi^i\right)\cdot a(e)\left(\nabla\chi^j_{T,k} - \nabla\chi^j\right)\right]\right| \leq C \begin{cases} T^{-d/2} & \text{for } 2 \leq d \leq 5, \\ T^{-3}\log(T) & \text{for } d = 6, \\ T^{-3} & \text{for } d > 6, \end{cases}$$

with an improvement of one order of convergence with respect to the case without the Richardson iteration. Albeit this method allows to improve the decay rate, it is not as effective as in the periodic case where the improvement was from T^{-2} to T^{-2k}.

The recent contribution [19] provides new results in stochastic homogenization for what concerns the existence and uniqueness of correctors, the convergence rates for the approximation of a^0 and bounds on the homogenization error.

6.3 Computational approaches in stochastic homogenization

The aim of numerical upscaling schemes is to compute the *best* approximation of a^0 with the *least* possible computational cost. Above, we described how the regularised corrector problem eq. (6.11) can be used to compute approximate homogenized matrices and discuss the convergence rates of such an approach. In the context of stochastic homogenization, the main source of error is due to the statistical error, with the exception of the standard cut-off model ($T = \infty$) for $d \leq 2$. In [77], the author adopts a Monte Carlo (MC) approach to reduce the statistical error and addresses the problem of optimizing the computational cost. The homogenized coefficients are approximated by averaging the upscaled coefficients obtained by N independent samplings of the conductances. The problem is solved for both the case of a single processor and of multiple processors and the optimal value of N will depend on the number of available processors.

6.3.1 The embedded method

We have already discussed the use of embedded corrector problems to approximate a^0 in Section 2.4.1. The method [41, 42] was developed for materials with random inclusions, as the one depicted in Figure 6.1a, where the size of the inclusions may vary as well. The *embedded corrector problem* is

$$-\nabla \cdot \left(a_R(y)\left(\chi_R + \xi\right)\right) = 0 \quad \text{in } \mathscr{D}'\left(\mathbb{R}^d\right). \tag{6.18}$$

where the coefficient tensor $a_R(y)$ is defined as

$$a_R(y) = \begin{cases} a(y) & y \in K_R, \\ \bar{a} & y \in \mathbb{R}^d \setminus K_R. \end{cases}$$

The constant matrix $\bar{a} \in \mathbb{R}^{d\times d}$ is *a priori unknown*, as it approximates the homogenized matrix. For any $\bar{a} \in \mathbb{R}^{d\times d}$, the matrix G is defined by $\xi \cdot G(\bar{a})\xi := 2\mathscr{J}_p(\bar{a})$, where $\mathscr{J}_p(\bar{a})$ is a functional defined in Section 2.4.1. The following three approximations of a^0 are proposed:

$$a_1^{0,R} = \underset{\bar{a}\in\mathscr{M}(\alpha,\beta)}{\operatorname{argmax}} \operatorname{Tr}\left(G(\bar{a})\right), \quad a_2^{0,R} = G(a_1^{0,R}), \quad a_3^{0,R} = G(a_3^{0,R}).$$

Numerical tests for randomly arranged inclusions showed that the errors for the second and third method are smaller, and we have less oscillations in the convergence to a^0. However, convergence rates are difficult to evaluate in this setting.

6.3.2 Variance reduction techniques

Since the variance error is often the dominant source of error, several works addressed the problem of reducing it. Many of those adapted *variance reduction* techniques to the context of homogenization of random media, see e.g. [106]. In [34], three variance reduction techniques

are presented:

- The method of antithetic variables, proposed in [32, 109], achieves a reduction of the upscaled coefficients' variance by sampling from an "antithetic" random coefficient field, i.e. a random variable with the same distribution as the original field, but such that their covariance is negative. For example, if we sample a random conductivity field, $a(x)$ from a log-normal distribution, an antithetic variable can be its inverse, $1/a(x)$. This approach has a substantial efficiency, but is also limited in particular because the technique does not fully exploit the specifics of the problem considered.
- The method of control variates, proposed in [108], requires a better knowledge of the problem at hand. In this approach, a surrogate problem, simpler to simulate and close to the original problem, has to be considered and concurrently solved. The technique uses that knowledge to, effectively, obtain a much better reduction of the variance in comparison to the method of antithetic variables. For example, in the case of periodic structures with random defects, the surrogate model is the periodic problem without defects, which can be solved easily on a single periodic cell.
- The method of special quasi-random structures, proposed in [107], consists in selecting only some realizations of the random environment. For instance, one may discard realizations of the random field if the empirical statistics (e.g., the spatial average or the volume fraction) are sufficiently far from the moments of the random field. Mathematically, this selection of suitable configurations among all possible ones amounts to replacing the computation of an expectation by that of a conditional expectation. This approach allows to neglect, via a cheap pre-evaluation, to compute the solution for a very unlikely configuration and, consequently, reduces the overall cost of the Monte Carlo sampling. This approach has been fully analysed in [70].

A different approach, to reduce the MC sampling cost, is to adopt a Multi Level Monte Carlo (MLMC) approach, [59]: many inexpensive computations with the smallest cell size are combined with fewer expensive computations performed on larger cells. An important remark is that MLMC approaches are interesting when the exact homogenized properties are stochastic, therefore, the MLMC approach is preferable for non-ergodic random fields[2]. In the case of ergodic coefficients, the homogenized matrix is deterministic and the MLMC cost equals the MC's one. The mathematical reason is that the variance of the estimators of a^0 decays not only with the number of MC samples, but also with the cell's size as R^{-d}, as a consequence of the Central Limit Theorem.

6.3.3 An iterative method

In Section 2.4.1, we described an iterative approach to approximate the homogenized limit of random coefficients which was developed in the context of random networks, [119], and

[2] Wiener processes are examples of non-ergodic random fields.

later adapted to the continuous case, [87]. The aim of [119] is to develop a method for the computation of the homogenized coefficients with optimal computational complexity, having proved that no algorithm can output an approximation a_δ of a^0 with error

$$\mathbb{E}\left[\left|a_\delta - a^0\right|^2\right]^{\frac{1}{2}} \leq \delta,$$

with less than $\mathcal{O}(\delta^{-2})$ operations. The iterative scheme proposed reaches this limit and is, thus, optimal under the point of view of computational complexity. A bounded domain problem is proposed for practical purposes as an approximation of the original one eq. (2.56), which is posed on the entire $\mathbb{Z}^d$. The approximation error due to the spatial cut-off, which is of exponential order and can thus be neglected. The peculiarity of this method is that the computational domain can be reduced as the iterations go on. The auxiliary problems take the form:

$$\begin{cases} 2^{-k}\chi^i_{R,k} - \nabla\cdot\left(a(y)\nabla\chi^i_{R,k}\right) = 2^{-k}\chi^i_{R,k-1} & \text{in } K_{R_k}, \\ \chi_{R,k} = 0 & \text{on } \partial K_{R_k}, \end{cases}$$

where

$$\chi^i_{-1} = \nabla\cdot\left(a(y)\mathbf{e}_i\right), \quad \text{and} \quad R_k = 2^{n-\left(\frac{1}{2}-\varepsilon\right)k} + C(1+n)2^{\frac{k}{2}}.$$

The homogenized coefficients are reconstructed as

$$a^{0,R,L,n}_{ij} := \fint_{K_{L_0}} a_{ij}(y)\,dy + \sum_{k=0}^{n} 2^k \fint_{K_{L_k}} \chi^i_{R,k-1}\chi^j_{R,k} + \chi^i_{R,k}\chi^j_{R,k}\,dy.$$

with

$$L_k = R_k - C(1+n)2^{\frac{k}{2}}.$$

In this case, the mean square resonance error is bounded by

$$\mathbb{E}\left[\left|a^0_{ij} - a^{0,R,L,n}_{ij}\right|^2\right]^{\frac{1}{2}} \leq CL_n^{-\frac{d}{2}}, \tag{6.19}$$

which is referred as the Central Limit Theorem error.

6.4 Conclusion

The subject of stochastic homogenization is far more complicated than the periodic case. First of all, first order correctors cannot be uniquely defined. Additionally, further assumptions are needed in order to derive convergence rates of the approximation error, for example the mixing condition or the finite range of dependence. Even in such a case the rates of convergence are not explicitly known, [135]. A way to overcome this gap is the use of a regularised corrector problem, for which explicit convergence rates for the bounds on the statistical and systematic errors were proved by Gloria and Otto, under the assumption of spectral gap inequality. The recent monograph [19] gives a broad overview of the analysis of stochastic homogenization.

These last researches indicate that the best approximation error estimate that can be achieved in this context is given by the Central Limit Theorem and scales as $R^{-d/2}$.

From the computational point of view, many past studies focused on verifying the error bounds derived analytically, without proposing alternative numerical schemes to solve stochastic multiscale problems. Legoll et al. [34, 106, 107, 108, 109] proposed several variance reduction approaches to reduce the statistical error in the approximation of a^0, with the aim of reducing the computational cost, rather than improving the convergence rates.

7 Modified elliptic corrector problems for random media

As discussed in Chapter 5, the modified elliptic corrector problem eq. (5.1) can be used to approximate the homogenized coefficients with higher convergence rates of the boundary error. The analysis in Chapter 5 only covers periodical deterministic micro-structures, so we aim to extend those results to the homogenization of stationary ergodic random media. Our final goal is to assess the convergence rate of the error associated to the approximation of a^0 via the modified elliptic scheme and compare it to the standard corrector problem (6.14). For the latter case, we already know that accuracy scales at most as R^{-1} [35, 135].

Let $\xi \in \mathbb{R}^d$, $|\xi| = 1$ and e^{-AT} be the semigroup, evaluated at time T, generated by the second order elliptic operator $A : H_0^1(K_R) \mapsto H^{-1}(K_R)$ defined by $Au := -\nabla \cdot (a(x)\nabla u)$. Let us consider the cell problem

$$\begin{cases} -\nabla \cdot \left(a(x)\left(\nabla \chi_{R,T} + \xi\right)\right) = -e^{-AT}\left[\nabla \cdot (a(x)\xi)\right] & \text{in } K_R, \\ \chi^i_{T,R}(x) = 0 & \text{on } \partial K_R, \end{cases} \tag{7.1}$$

where the diffusion coefficients $a_{ij}(x)$ are the realization of a stationary ergodic random tensor field, as seen in Chapter 6. The choice of homogeneous Dirichlet boundary conditions is completely arbitrary; as a matter of fact, we could replace it by, e.g., periodic or homogeneous Neumann boundary conditions, provided that the space from which A operates is changed accordingly. We will assume that, as proved in the periodic setting, the modified corrector over the bounded cell, $\chi_{R,T}$, approximates with an infinite order of accuracy the modified corrector over the unbounded cell, χ_T, which is defined as the solution of

$$-\nabla \cdot \left(a(x)\left(\nabla \chi_T + \xi\right)\right) = -u(\cdot, T), \quad \text{in } \mathbb{R}^d, \tag{7.2}$$

where $u(\cdot, T)$ is the solution of the Cauchy problem (7.3), evaluated at time T:

$$\begin{cases} \dfrac{\partial u}{\partial t} - \nabla \cdot (a(x)\nabla u) = 0 & \text{in } \mathbb{R}^d \times (0, +\infty), \\ u(x, 0) = \nabla \cdot (a(x)\xi) & \text{in } \mathbb{R}^d. \end{cases} \tag{7.3}$$

The modified corrector functions $\chi_{R,T}$ and χ_T are employed to upscale the multiscale tensor by the *corrector average formulas*:

$$\xi \cdot a^{0,R,L,T}\xi := \fint_{K_L} \left(\nabla\chi_{R,T} + \xi\right) \cdot a(x) \left(\nabla\chi_{R,T} + \xi\right) dx, \text{ and} \tag{7.4}$$

$$\xi \cdot a^{0,L,T}\xi := \fint_{K_L} \left(\nabla\chi_T + \xi\right) \cdot a(x) \left(\nabla\chi_T + \xi\right) dx, \tag{7.5}$$

where the restriction over the smaller box K_L is necessary in order to achieve the exponential decay of the boundary error, whose upper bound depends on $(R-L)$, as proved in Chapter 5 for the periodic setting.

Outline

In Section 7.1, we prove the well-posedness of the modified corrector problem on the unbounded domain (7.2) in the space of stationary random fields. An *a priori* upper bound on the systematic error $a^{0,T} - a^0$ ($a^{0,T}$ being defined in (7.18)) is proved in Section 7.2 as a consequence of the time decay of parabolic solutions. Along with the systematic error, the statistical error is discussed as well. Finally, we demonstrate the decay of the global resonance error $a^{0,R,L,T} - a^0$ by means of numerical experiments in Section 7.3. In particular, this makes it possible to choose optimal values for the parameters T, R, L in the model problem (7.1), which is needed for computationally efficient and accurate approximations of the homogenized tensor in random media.

The content of this chapter is based on [7].

7.1 Well-posedness of the corrector problem

In this section we prove that the corrector problem eq. (7.2) is well-posed and that $\nabla\chi_T$ is a stationary random field. The well-posedness proof is based on the equivalence between the gradient of the modified corrector $\nabla\chi_T$ and the time integral of ∇u, for which we rely on time decay properties of parabolic solutions. The stationarity of $\nabla\chi_T$, which can be compared to the stationarity of $\nabla\chi$ of eq. (6.3), is essential for applying the ergodic theorem in the definition of $a^{0,T}$. In this chapter we will use the following notation:

- $\mathscr{L}^2$ denotes the space of stationary extension of square-integrable random variables:

$$\mathscr{L}^2 = \left\{u(x,\omega) = u(\tau_x\omega) : u \in L^2(\Omega)\right\}.$$

- $\mathscr{L}^2_{pot}$ denotes the space of stationary extension of square-integrable potential random variables:

$$\mathscr{L}^2_{pot} = \left\{\mathbf{v}(x,\omega) = \mathbf{v}(\tau_x\omega) : \mathbf{v} \in \mathbf{L}^2_{pot}(\Omega)\right\}.$$

- $\mathscr{H}^1$ denotes the space of ω-differentiable random variable, as defined in eq. (6.6).

Theorem 7.1. *Let $u \in C\left([0,+\infty), \mathscr{L}^2\right)$ be the solution of eq.* (7.3)*. Then, there exists a unique $\nabla\chi_T \in \mathscr{L}^2_{pot}$ such that*

$$-\nabla\cdot\left(a(\nabla\chi_T+\xi)\right) = -u(\cdot,T), \ in\, \mathscr{D}'(\mathbb{R}^d), \mathbb{P}\text{-}a.s.. \tag{7.6}$$

Remark 7.2. *In [19] the authors proved that for ∇u and $\nabla\chi$ as above the following relation holds true:*

$$\nabla\chi(x) = \int_0^{+\infty} \nabla u(x,t)\,dt, \quad \mathbb{P}\text{-}a.s.$$

This identity entails the stationarity of $\nabla\chi$, as a consequence of the stationarity of ∇u. Moreover, the identity does not hold true for u and χ, because the time integral of u does not converge. We will use a similar identity to prove Theorem 7.1.

The proof of Theorem 7.1 is based on the decay in time of the parabolic solution u, which are collected in Section 7.1.1.

7.1.1 Decay of parabolic solutions

In this section we collect some results about the decay in time of the solutions to parabolic PDEs in $\mathbb{R}^d$ with stationary random coefficients. The decay properties will eventually be used in the proofs of Theorems 7.1 and 7.9. Existence and uniqueness of the solution u to eq. (7.3) is a classical results of the theory of linear parabolic partial differential equations, [71]. First of all, we recall a classical result on the time decay of the solutions to parabolic problems and deduce the results of Lemma 7.4. These results are not new, for example they are proved in [19] for the case of $\mathbb{Z}^d$-stationary random fields. The proof is based on writing the solution $u(x,t)$ as a time integral involving the fundamental solution $\Gamma(x,y,t)$:

$$u(x,t,a) = -\int_{\mathbb{R}^d} \nabla_y\Gamma(x,y,t)\cdot a(y)\xi\,dy. \tag{7.7}$$

We recall that the $\Gamma(x,y,t)$ solves:

$$\begin{cases} \dfrac{\partial\Gamma}{\partial t}(\cdot,y,\cdot) - \nabla_x\cdot(a(x)\nabla_x\Gamma(\cdot,y,\cdot)) = 0, \\ \Gamma(\cdot,y,0) = \delta_y(\cdot), \end{cases} \text{ and } \begin{cases} \dfrac{\partial\Gamma}{\partial t}(x,\cdot,\cdot) - \nabla_y\cdot(a(y)\nabla_y\Gamma(x,\cdot,\cdot)) = 0, \\ \Gamma(x,\cdot,0) = \delta_x(\cdot), \end{cases}$$

where δ_z is the *Dirac delta function* centered in $z \in \mathbb{R}^d$.

Lemma 7.3. *Let u be the solution of eq.* (7.3)*. Then, there exists a constant $C(\alpha,\beta,d) > 0$ such that, for every $t > 0$,*

$$\|u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)} + t^{\frac{1}{2}}\|\nabla u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)} \le C t^{-\frac{1}{2}}. \tag{7.8}$$

Lemma 7.4. *Let u be the solution of eq.* (7.3)*. Then, u is a stationary random field and*

$$\mathbb{E}[u(x,t)] = 0, \quad \forall\, t > 0, \forall\, x \in \mathbb{R}^d. \tag{7.9}$$

Proof. Step 1. We prove the stationarity of u. Let us recall that u can be expressed by formula eq. (7.7). Then, by the fact that $\Gamma(x+z, y+z, t, a) = \Gamma(x, y, t, \tau_z a)$

$$\begin{aligned} u(x+z,t,a) &= -\int_{\mathbb{R}^d} \nabla_y \Gamma(x+z, y, t, a) \cdot a(y)\xi \, dy \\ &= -\int_{\mathbb{R}^d} \nabla_y \Gamma(x+z, y+z, t, a) \cdot a(y+z)\xi \, dy \\ &= -\int_{\mathbb{R}^d} \nabla_y \Gamma(x, y, t, \tau_z a) \cdot \tau_z a(y)\xi \, dy = u(x,t,\tau_z a). \end{aligned}$$

Step 2. Let $B_1 \subset \mathbb{R}^d$ be the unit ball centred in 0, $\psi \in C_0^\infty(B_1)$ with unit mass in $L^1(B_1)$ and $\psi_R(x) := R^{-d}\psi(x/R)$. Let us write eq. (7.3) in weak form with ψ_R as test function and integrate in time for $0 < t_1 < t < t_2$:

$$\mathbb{E}[u(\cdot,t_1)] - \mathbb{E}[u(\cdot,t_2)] = \lim_{R\to+\infty} \mathbb{E}\left[\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \nabla u(x,t)\cdot a(x)\nabla\psi_R(x)\, dx\, dt\right].$$

By the Hölder inequality, we bound the absolute value of the right-hand side from above:

$$\begin{aligned} \left|\mathbb{E}\left[\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \nabla u(x,t)\cdot a(x)\nabla\psi_R(x)\, dx\, dt\right]\right| &\le \mathbb{E}\left[\int_{t_1}^{t_2} \beta \|\nabla u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)} \left\|\nabla\psi_R\right\|_{L^1(\mathbb{R}^d)} dt\right] \\ &\le \beta R^{-1} \left\|\nabla\psi\right\|_{L^1(\mathbb{R}^d)} \mathbb{E}\left[\int_{t_1}^{t_2} \|\nabla u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)}\, dt\right] \end{aligned}$$

The term $\mathbb{E}\left[\int_{t_1}^{t_2} \|\nabla u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)}\, dt\right]$ is uniformly bounded in R thanks to the decay of $\|\nabla u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)}$ of Lemma 7.3. So,

$$\lim_{R\to+\infty} \mathbb{E}\left[\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \nabla u(x,t)\cdot a(x)\nabla\psi_R(x)\, dx\, dt\right] = 0,$$

and we deduce that $\mathbb{E}[u(\cdot,t)]$ is constant in time. From the fact that $\|u(\cdot,t)\|_{L^\infty(\mathbb{R}^d)}$ decays to zero and from the stationarity of u, we conclude that $\mathbb{E}[u(x,t)] = 0$ for any $t > 0$ and any $x \in \mathbb{R}^d$. □

Time decay rates of $\mathbb{E}[|u|^p]$ and $\mathbb{E}[|\nabla u|^p]$ for homogenization problems over discrete networks were proved in several works, e.g. [80, Theorem 1] and [119, Lemma 9.7]:

$$\mathbb{E}\left[|u|^p\right]^{\frac{1}{p}} \le C(t+1)^{-\left(\frac{1}{2}+\frac{d}{4}\right)} \text{ for any } p \ge 1 \quad \text{and} \quad \mathbb{E}\left[|\nabla u|^2\right]^{\frac{1}{2}} \le C(t+1)^{-\left(1+\frac{d}{4}\right)}.$$

More recently, similar estimates were also derived for the continuous case in [19]. Theorem 7.5 and Corollary 7.6 provide time decay bounds on the moments $\mathbb{E}[|u|^p]$.

Theorem 7.5 ([19, Theorem 9.1])**.** *For every $\sigma \in (0,2)$, there exists a constant $C(\sigma,d,\alpha,\beta) < +\infty$ such that the following holds. Let $a(\cdot) \in L^\infty(\mathbb{R}^d)$ be a stationary random field such that, for every $x \in \mathbb{R}^d$, $a(x)\xi$ is $\mathscr{F}$-measurable and let $u \in C\left([0,+\infty), \mathscr{L}^2\right)$ be the solution of eq.* (7.3)*. Then, for every $t \in [1,+\infty)$ and $x \in \mathbb{R}^d$,*

$$\mathbb{E}\left[\exp\left(\left(C^{-1} t^{\frac{1}{2}+\frac{d}{4}} |u(x,t)|\right)^\sigma\right)\right] \leq 2. \tag{7.10}$$

Corollary 7.6. *Let the assumptions of Theorem 7.5 be satisfied. Then, for any $p \geq 1$, there exists a constant $C(p,d,\alpha,\beta) < +\infty$ such that, for every $t \in [1,+\infty)$ and $x \in \mathbb{R}^d$*

$$\mathbb{E}\left[|u(x,t)|^p\right]^{\frac{1}{p}} \leq C t^{-\left(\frac{1}{2}+\frac{d}{4}\right)} \tag{7.11}$$

Proof. From Theorem 7.5, by taking $\sigma = 1$, we know that there exist $C(d,\alpha,\beta) < +\infty$ such that

$$\mathbb{E}\left[\exp\left(C^{-1} t^{\frac{1}{2}+\frac{d}{4}} |u(x,t)|\right)\right] \leq 2$$

for every $t \in [1,+\infty)$ and $x \in \mathbb{R}^d$. Since the exponential of a random variable X grows faster than $|X|^p$ for any p, the integrability of e^X implies the integrability of any power of X. Therefore, there exists a constant $C(p) < +\infty$ such that

$$\mathbb{E}\left[|X|^p\right] \leq C(p)\mathbb{E}\left[e^X\right].$$

By taking $X = C^{-1} t^{\frac{1}{2}+\frac{d}{4}} |u(x,t)|$ in the previous inequality we conclude that there exists a constant $C(p,d,\alpha,\beta) < +\infty$ such that

$$\mathbb{E}\left[|u(x,t)|^p\right]^{\frac{1}{p}} \leq C t^{-\left(\frac{1}{2}+\frac{d}{4}\right)}.$$

□

Corollary 7.6 shows that there is a clear difference between the time decay of parabolic solutions set in bounded domains (as, for instance, in the case of periodic correctors) and in unbounded domains (as in the stochastic homogenization setting). Indeed, in the periodic (or bounded domain) setting, the Poincaré inequality entails exponential decay in time of the spatial L^2-norm. Such a property is fundamental in the derivation of exponential order convergence rates of the resonance error in Chapters 4 and 5. In the stochastic setting we do not necessarily have such an inequality in $\mathscr{H}^1$.

Proposition 7.7. *Let u be the solution of eq.* (7.3) *with $\nabla \cdot (a(x)\xi) \in \mathscr{L}^2$. Then*

$$u \in L^2\left((0,+\infty), \mathscr{H}^1\right) \cap C\left([0,+\infty), \mathscr{L}^2\right).$$

Proof. We first prove that $u \in L^2\left((0,+\infty), \mathscr{H}^1\right)$ and, then, that $u \in C\left([0,+\infty), \mathscr{L}^2\right)$.
Step 1 - $u \in L^2\left((0,+\infty), \mathscr{H}^1\right)$:

We already know from Lemma 7.4 that $u(\cdot,t)$ is stationary for any $t \geq 0$. So, we only have to prove that

$$\int_0^{+\infty} \mathbb{E}\left[u(\cdot,t)^2\right] dt < +\infty, \quad \text{and} \quad \int_0^{+\infty} \mathbb{E}\left[|\nabla u(\cdot,t)|^2\right] dt < +\infty.$$

The function $\mathbb{E}\left[u(\cdot,t)^2\right]$ is decreasing in time, indeed, from eq. (7.3),

$$\frac{d}{dt}\mathbb{E}\left[u^2\right] = 2\mathbb{E}\left[u\partial_t u\right] = -2\mathbb{E}\left[\nabla u \cdot a(x)\nabla u\right] < 0.$$

So, we can bound the integral using $\nabla \cdot (a(x)\xi) \in \mathscr{L}^2$ and the result of Corollary 7.6:

$$\begin{aligned}\int_0^{+\infty} \mathbb{E}\left[u(\cdot,t)^2\right] dt &\leq \int_0^1 \mathbb{E}\left[u(\cdot,t)^2\right] dt + \int_1^{+\infty} \mathbb{E}\left[u(\cdot,t)^2\right] dt \\ &\leq \mathbb{E}\left[|\nabla \cdot (a(x)\xi)|^2\right] + C\int_1^{+\infty} t^{-\left(1+\frac{d}{2}\right)} dt < +\infty.\end{aligned} \tag{7.12}$$

Next, from the ellipticity of $a(\cdot)$, we have:

$$\alpha\mathbb{E}\left[|\nabla u(\cdot,t)|^2\right] \leq \mathbb{E}\left[\nabla u \cdot a(x)\nabla u\right] = -\frac{1}{2}\frac{d}{dt}\mathbb{E}\left[u(\cdot,t)^2\right].$$

So, since $\mathbb{E}\left[u(\cdot,t)^2\right]$ vanishes for $t \to +\infty$,

$$\int_0^{+\infty} \mathbb{E}\left[|\nabla u(\cdot,t)|^2\right] dt \leq \frac{1}{2\alpha}\mathbb{E}\left[u(\cdot,0)^2\right] < +\infty. \tag{7.13}$$

From eq. (7.12) and eq. (7.13) we conclude that $u \in L^2\left((0,+\infty), \mathscr{H}^1\right)$.

Step 2 - $u \in C\left([0,+\infty), \mathscr{L}^2\right)$:

Let $t \geq 0$. Since $f(z) = \sqrt{z}$ is continuous in $[0,+\infty)$, it is sufficient to prove the continuity of $\mathbb{E}\left[u^2\right]$:

$$\begin{aligned}\mathbb{E}\left[u(\cdot,t+h)^2\right] - \mathbb{E}\left[u(\cdot,t)^2\right] &= \int_t^{t+h} \frac{d}{dt}\mathbb{E}\left[u(\cdot,t)^2\right] dt \\ &= -\int_t^{t+h} \mathbb{E}\left[\nabla u \cdot a(x)\nabla u\right] dt \underset{h\to 0}{\longrightarrow} 0,\end{aligned}$$

and the proof is concluded. □

Now, we state a result on the time decay of the second moment of ∇u. The proof follows from the one of [119, Lemma 9.7].

Proposition 7.8. *Let $a(\cdot) \in \Omega$ and let u be the solution of eq.* (7.3)*. Then, there exist a positive constant $C(d,\alpha,\beta) < +\infty$ such that, for every $t \in [2,+\infty)$ and $x \in \mathbb{R}^d$,*

$$\mathbb{E}\left[|\nabla u(x,t)|^2\right]^{\frac{1}{2}} \leq Ct^{-\left(\frac{d}{4}+1\right)}. \tag{7.14}$$

Proof. Let us begin by proving that the map $t \mapsto \mathbb{E}[\nabla u(x,t) \cdot a(x) \nabla u(x,t)]$ is nonincreasing. Indeed, its time derivative can be expressed as:

$$\begin{aligned}\partial_t \mathbb{E}[\nabla u(x,t) \cdot a(x) \nabla u(x,t)] &= 2\mathbb{E}[\nabla(\partial_t u)(x,t) \cdot a(x) \nabla u(x,t)] \\ &= 2\mathbb{E}[\nabla(\nabla \cdot (a(x) \nabla u(x,t))) \cdot a(x) \nabla u(x,t)] \\ &= -2\mathbb{E}\left[|\nabla \cdot (a(x) \nabla u(x,t))|^2\right] \leq 0.\end{aligned}$$

Thus, from the weak formulation of eq. (7.3) with u as test function and inequality eq. (7.11) for $t/2 \geq 1$ and $p = 2$, we can write

$$\begin{aligned}\mathbb{E}[\nabla u(x,t) \cdot a(x) \nabla u(x,t)] &\leq \frac{2}{t} \int_{\frac{t}{2}}^{t} \mathbb{E}[\nabla u(x,s) \cdot a(x) \nabla u(x,s)] \, ds \\ &\leq -\frac{1}{t} \int_{\frac{t}{2}}^{t} \partial_t \mathbb{E}\left[|u(x,s)|^2\right] ds \\ &\leq \frac{1}{t} \left[\mathbb{E}\left[|u(x,t)|^2\right] - \mathbb{E}\left[\left|u\left(x, \frac{t}{2}\right)\right|^2\right]\right] \\ &\leq C t^{-(\frac{d}{2}+2)}.\end{aligned}$$

Then, eq. (7.14) follows from the assumption of uniform ellipticity of the coefficients. □

Now we are ready to prove that the differential problem of eq. (7.2) is well-posed.

Proof of Theorem 7.1. Let us define the stationary function

$$\Psi := \int_0^T \nabla u(\cdot, t) \, dt. \tag{7.15}$$

Then, $\Psi \in (\mathscr{L}^2)^d$. Indeed, by Minkowski integral inequality and Proposition 7.7 we have:

$$\mathbb{E}\left[|\Psi(x)|^2\right]^{\frac{1}{2}} := \mathbb{E}\left[\left|\int_0^T \nabla u(x,t) \, dt\right|^2\right]^{\frac{1}{2}} \leq \int_0^T \mathbb{E}\left[|\nabla u(x,t)|^2\right]^{\frac{1}{2}} dt < +\infty.$$

The weak form of eq. (7.3) is: Find $u \in L^2\left((0,+\infty), \mathscr{H}^1\right)$ such that

$$\frac{d}{dt}\mathbb{E}\left[u\phi\right] + \mathbb{E}\left[\nabla\phi \cdot a(x) \nabla u\right] = 0, \quad \forall \phi \in \mathscr{H}^1.$$

By integration in time and eq. (7.15), we get

$$\mathbb{E}\left[\nabla\phi \cdot a(x) \left(\Psi + \xi\right)\right] = -\mathbb{E}\left[u(\cdot, T)\phi\right], \quad \forall \phi \in \mathscr{H}^1.$$

To conclude, we have to prove that $\Psi \in \mathscr{L}^2_{pot}$. The function Ψ is trivially vortex-free, since it is the gradient of $\int_0^T u(\cdot, t) \, dt$. Hence, we are allowed to define Ψ as $\nabla \chi_T$.

The uniqueness of $\nabla\chi_T$ trivially follows from uniqueness of solution for the standard corrector problem eq. (6.3), proved in [123]. □

7.2 *A priori* error bounds

In this section we will discuss the *a priori* bounds on the resonance error. The error can be measured in a strong sense, by the $L^2(\Omega)$-norm, or in a weak sense, as the absolute value of the mean difference. The difference between the two is the presence/absence of the so-called *statistical error*, which accounts for the fact that the we have a random approximation of a^0.

7.2.1 Error decomposition

We define two measures of the resonance error: the mean square resonance error and the mean resonance error, respectively defined as

$$e_{MS} := \sup_{\xi\in\mathbb{R}^d,\|\xi\|=1} \left(\mathbb{E}\left[\left(\xi\cdot(a^{0,R,L,T}-a^0)\xi\right)^2\right]\right)^{\frac{1}{2}}, \quad \text{and}$$
$$e_M := \sup_{\xi\in\mathbb{R}^d,\|\xi\|=1} \left|\xi\cdot\mathbb{E}\left[a^{0,R,L,T}-a^0\right]\xi\right|.$$

Both errors can be decomposed into several contributions that can be estimated separately. By the triangle inequality, the mean square error can be decomposed as

$$\left(\mathbb{E}\left[\left(\xi\cdot(a^{0,R,L,T}-a^0)\xi\right)^2\right]\right)^{\frac{1}{2}} \le \left(\mathbb{E}\left[\left(\xi\cdot(a^{0,R,L,T}-a^{0,L,T})\xi\right)^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\left(\xi\cdot(a^{0,L,T}-a^{0,T})\xi\right)^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\left(\xi\cdot(a^{0,T}-a^0)\xi\right)^2\right]\right)^{\frac{1}{2}},$$

where $a^{0,T}$ is defined in eq. (7.18). The difference $a^{0,R,L,T}-a^{0,L,T}$ accounts for the error due to mismatching conditions of the corrector functions on ∂K_R. Such an error has been analysed for the periodic deterministic case and computed numerically for the periodic and the stochastic cases [8]. The rate of convergence of the boundary error is exponential, whose exponent depends on the contrast of coefficients. We assume that the error $a^{0,R,L,T}-a^{0,L,T}$ can be bounded by a deterministic bound depending on R, L, T and we denote it by $e_{BD}(R,L,T)$. Hence, the mean square error is thus bounded by the sum:

$$e_{MS} \le e_{BD}(R,L,T) + \underbrace{\sup_{\xi\in\mathbb{R}^d,\|\xi\|=1}\sqrt{\mathrm{Var}\left[\xi\cdot a^{0,T,L}\xi\right]}}_{\text{statistical error}} + \underbrace{\sup_{\xi\in\mathbb{R}^d,\|\xi\|=1}\left|\xi\cdot(a^{0,T}-a^0)\xi\right|}_{\text{systematic error}}.$$

The advantage of the error in mean is that it allows to remove the statistical error as the difference in mean can be bounded by:

$$\left|\xi\cdot\mathbb{E}\left[a^{0,R,L,T}-a^0\right]\xi\right| \le \left|\xi\cdot\mathbb{E}\left[a^{0,R,L,T}-a^{0,L,T}\right]\xi\right|$$

$$+\underbrace{\left|\xi\cdot\mathbb{E}\left[a^{0,L,T}-a^{0,T}\right]\xi\right|}_{=0,\text{ by eq. (7.18)}}+\left|\xi\cdot\mathbb{E}\left[a^{0,T}-a^{0}\right]\xi\right|.$$

So, the error in mean can be bounded by

$$e_M \le e_{BD}(R,L,T)+\underbrace{\sup_{\xi\in\mathbb{R}^d,\|\xi\|=1}\left|\xi\cdot(a^{0,T}-a^0)\xi\right|}_{\text{systematic error}}. \tag{7.16}$$

In Section 7.2.2 we prove an *a priori* bound for the systematic error, while in Section 7.2.3 the statistical error is discussed.

7.2.2 Systematic error

The systematic error is

$$e_{SYS}:=\sup_{\xi\in\mathbb{R}^d,\|\xi\|=1}\left|\xi\cdot(a^{0,T}-a^0)\xi\right|. \tag{7.17}$$

We will rely on the result on the time decay of $\nabla u(\cdot,t)$ and on the definition of $\nabla\chi$ and $\nabla\chi_T$ as time integral of $\nabla u(\cdot,t)$ in order to bound the systematic error.

Theorem 7.9 (systematic error)**.** *Let $a(x)\in\Omega$, a^0 and $a^{0,T}$ be defined, respectively, as in eq.* (6.8) *and by*

$$\xi\cdot a^{0,T}\xi:=\lim_{L\to+\infty}\fint_{K_L}\left(\nabla\chi_T+\xi\right)\cdot a(y)\left(\nabla\chi_T+\xi\right)dy=\mathbb{E}\left[\xi\cdot a^{0,T,L}\xi\right]. \tag{7.18}$$

Then, there exists a positive constant $C(d,\alpha,\beta)<+\infty$ such that

$$\sup_{\xi\in\mathbb{R}^d,\|\xi\|=1}\left|\xi\cdot(a^{0,T}-a^0)\xi\right|\le CT^{-\frac{d}{2}}. \tag{7.19}$$

Proof. We first notice that the two identities contained in eq. (7.18) follow from the stationarity of $a(\cdot)$ and the Birkhoff ergodic theorem. Next, we prove that

$$\xi\cdot(a^{0,T}-a^0)\xi=\mathbb{E}\left[\left(\nabla\chi_T-\nabla\chi\right)\cdot a(\cdot)\left(\nabla\chi_T-\nabla\chi\right)\right]. \tag{7.20}$$

By definition of $a^{0,T}$ and a^0,

$$\begin{aligned}
\xi\cdot(a^{0,T}-a^0)\xi&=\mathbb{E}\left[\left(\nabla\chi_T+\xi\right)\cdot a(\cdot)\left(\nabla\chi_T+\xi\right)-\left(\nabla\chi+\xi\right)\cdot a(\cdot)\left(\nabla\chi+\xi\right)\right]\\
&=\mathbb{E}\left[\left(\nabla\chi_T+\xi\right)\cdot a(\cdot)\left(\nabla\chi_T+\xi\right)-\left(\nabla\chi+\xi\right)\cdot a(\cdot)\left(\nabla\chi_T+\xi\right)\right]\\
&\quad+\mathbb{E}\left[\left(\nabla\chi_T+\xi\right)\cdot a(\cdot)\left(\nabla\chi+\xi\right)-\left(\nabla\chi+\xi\right)\cdot a(\cdot)\left(\nabla\chi+\xi\right)\right]\\
&=\mathbb{E}\left[\left(\nabla\chi_T-\nabla\chi\right)\cdot a(\cdot)\left(\nabla\chi_T+\xi\right)+\left(\nabla\chi_T-\nabla\chi\right)\cdot a(\cdot)\left(\nabla\chi+\xi\right)\right]\\
&=\mathbb{E}\left[\left(\nabla\chi_T-\nabla\chi\right)\cdot a(\cdot)\left(\nabla\chi_T+\xi\right)-\left(\nabla\chi_T-\nabla\chi\right)\cdot a(\cdot)\left(\nabla\chi+\xi\right)\right]\\
&=\mathbb{E}\left[\left(\nabla\chi_T-\nabla\chi\right)\cdot a(\cdot)\left(\nabla\chi_T-\nabla\chi\right)\right],
\end{aligned}$$

where the fourth inequality comes from

$$\mathbb{E}\left[\left(\nabla\chi_T - \nabla\chi\right)\cdot a(x)\left(\nabla\chi+\xi\right)\right] = 0 = -\mathbb{E}\left[\left(\nabla\chi_T - \nabla\chi\right)\cdot a(x)\left(\nabla\chi+\xi\right)\right],$$

for any $x\in\mathbb{R}^d$[1]. Thus, by the uniform boundedness of $a(\cdot)$ and the Hölder inequality we have:

$$\left|\xi\cdot\left(a^{0,T}-a^0\right)\xi\right| \le \beta\mathbb{E}\left[\left|\nabla\chi_T-\nabla\chi\right|^2\right].$$

Now, we recall that

$$\nabla\chi = \int_0^{+\infty}\nabla u(\cdot,t)\,dt, \quad \text{and } \nabla\chi_T = \int_0^{T}\nabla u(\cdot,t)\,dt,$$

we subsitute these equivalences in the expression above and use the Minkowski integral inequality to switch the two integrations:

$$\left|\xi\cdot\left(a^{0,T}-a^0\right)\xi\right| \le \beta\mathbb{E}\left[\left|\int_T^{+\infty}\nabla u(\cdot,t)\,dt\right|^2\right] \le \beta\left(\int_T^{+\infty}\mathbb{E}\left[|\nabla u(\cdot,t)|^2\right]^{\frac{1}{2}}\,dt\right)^2.$$

Finally, from the time decay result for $\nabla u(\cdot,t)$ eq. (7.14) we conclude that

$$\left|\xi\cdot\left(a^{0,T}-a^0\right)\xi\right| \le C\left(\int_T^{+\infty} t^{-\left(\frac{d}{4}+1\right)}\,dt\right)^2 \le CT^{-\frac{d}{2}}.$$

and eq. (7.19) follows from the fact that C does not depend on ξ. □

7.2.3 On the bound of the statistical error

In this section, we describe a possible approach to bound the statistical error for the presented homogenization model. By following a technique proposed in [83], we rely on the spectral gap inequality eq. (SG) of Definition 6.11, applied on the random variable $\xi\cdot a^{0,L,T}\xi$:

$$\mathrm{Var}\left[\xi\cdot a^{0,L,T}\xi\right] \le \frac{1}{\rho}\int_{R^d}\mathbb{E}\left[\left(\operatorname*{osc}_{a|_{B_\ell(z)}}\xi\cdot a^{0,L,T}\xi\right)^2\right]dz. \tag{7.21}$$

Unfortunately, the approach that we followed was not successful to determine *a priori* bounds on the statistical error, due to a lack of control on

$$\int_{B_\ell(x)}\left|\operatorname*{osc}_{a|_{B_\ell(z)}}\nabla\chi_T(y)\right|^2 dy,$$

resulting in incomplete estimates. Nonetheless, we describe here the followed approach to stimulate further research in this direction or to warn other researchers on following the same path.

[1] To prove it, one can test the standard corrector equation against $\theta(\chi_T-\chi)$, where $\theta\in C_0^\infty\left(\mathbb{R}^d\right)$ is such that $\theta\equiv 1$ on the ball B_R, then, pass to the limit for $R\to+\infty$ and use the ergodic theorem.

In order to estimate the right-hand side of the spectral gap inequality we need to provide a preliminary bound on the quantity $\int_{B_\ell(x)} \left|\nabla\chi_T + \xi\right|^2$.

Lemma 7.10. *Let $a \in \Omega$ and let $\nabla\chi_T \in \mathscr{L}^2_{pot}$ be the unique solution of eq.* (7.6)*. Let us assume that there exists a positive $C(\alpha,\beta,\ell) < +\infty$ such that*

$$\int_{B_\ell(z)} \nabla\cdot\left(a(y)\xi - \tilde{a}(y)\xi\right)^2 dy \leq C.$$

Then, there exists $C > 0$ independent of L and T such that

$$\sup_{a|_{B_\ell(z)}} \int_{B_\ell(x)} \left|\nabla\chi_T(y) + \xi\right|^2 dy \leq C\left(1 + \int_{B_\ell(x)} \left|\nabla\chi_T\right|^2 dy + \int_{B_\ell(z)} \left|\nabla\chi_T\right|^2 dy\right). \tag{7.22}$$

Proof. Let us consider a variation of $a(\cdot)$ over $B_\ell(z)$, let us denote it by $\tilde{a}(\cdot)$ and the corresponding corrector function by $\tilde{\chi}_T$. We start by estimating $\int_{B_\ell(x)} \left|\nabla\chi_T(y) - \nabla\tilde{\chi}_T(y)\right|^2 dy$. The difference $\chi_T - \tilde{\chi}_T$ satisfies

$$-\nabla\cdot\left[a(x)\nabla\left(\chi_T - \tilde{\chi}_T\right)\right] = -\nabla\cdot\left[\left(a(x) - \tilde{a}(x)\right)\left(\nabla\tilde{\chi}_T + \xi\right)\right] + u(x,T) - \tilde{u}(x,T), \tag{7.23}$$

and the difference $u - \tilde{u}$ satisfies:

$$\begin{aligned} &\partial_t(u-\tilde{u}) - \nabla\cdot\left(a(y)\nabla\left(u-\tilde{u}\right)\right) = \nabla\cdot\left((a-\tilde{a})(y)\nabla\tilde{u}\right), \\ &(u-\tilde{u})(y,0) = \nabla\cdot\left((a-\tilde{a})(y)\xi\right). \end{aligned} \tag{7.24}$$

The right-hand side of eq. (7.23) is bounded in the $H^{-1}(\mathbb{R}^d)$-norm, hence the difference $\chi_T - \tilde{\chi}_T$ is well defined in $H^1(\mathbb{R}^d)$. The weak form thus becomes:

$$\begin{aligned} &\int_{\mathbb{R}^d} \left(\nabla\chi_T - \nabla\tilde{\chi}_T\right)\cdot a(y)\left(\nabla\chi_T - \nabla\tilde{\chi}_T\right) dy \\ &= \int_{\mathbb{R}^d} \left(\nabla\chi_T - \nabla\tilde{\chi}_T\right)\cdot\left(a(y) - \tilde{a}(y)\right)\left(\nabla\tilde{\chi}_T + \xi\right) dy + \int_{\mathbb{R}^d} \left(u(y,T) - \tilde{u}(y,T)\right)\left(\chi_T(y) - \tilde{\chi}_T(y)\right) dy. \end{aligned} \tag{7.25}$$

By application of the Cauchy-Schwarz and Young inequalities we can bound the first term on the right hand side as

$$\begin{aligned} &\int_{\mathbb{R}^d} \left(\nabla\chi_T - \nabla\tilde{\chi}_T\right)\cdot\left(a(y) - \tilde{a}(y)\right)\left(\nabla\tilde{\chi}_T + \xi\right) dy \\ &\qquad \leq 2\beta \left\|\nabla\chi_T - \nabla\tilde{\chi}_T\right\|_{L^2(\mathbb{R}^d)} \left\|\nabla\tilde{\chi}_T + \xi\right\|_{L^2(B_\ell(z))} \\ &\qquad \leq \frac{\alpha}{4}\left\|\nabla\chi_T - \nabla\tilde{\chi}_T\right\|^2_{L^2(\mathbb{R}^d)} + \frac{2\beta^2}{\alpha}\left\|\nabla\tilde{\chi}_T + \xi\right\|^2_{L^2(B_\ell(z))}. \end{aligned} \tag{7.26}$$

The last term can be bounded by the weak form of eq. (7.24) with $\chi_T - \tilde{\chi}_T$ as test function:

$$\int_{\mathbb{R}^d} \big(u(y,T)-\tilde{u}(y,T)\big)\big(\chi_T(y)-\tilde{\chi}_T(y)\big)\,dy + \int_0^T \int_{\mathbb{R}^d} a(y)\big(\nabla u(y,t)-\nabla\tilde{u}(y,t)\big)\big(\nabla\chi_T(y)-\nabla\tilde{\chi}_T(y)\big)\,dy\,dt$$
$$= \int_{\mathbb{R}^d} \big(a(y)-\tilde{a}(y)\big)\xi\big(\nabla\chi_T(y)-\nabla\tilde{\chi}_T(y)\big)\,dy + \int_0^T \int_{\mathbb{R}^d} \big(a(y)-\tilde{a}(y)\big)\nabla\tilde{u}(y,t)\big(\nabla\chi_T(y)-\nabla\tilde{\chi}_T(y)\big)\,dy\,dt.$$

By recalling that $\int_0^T \nabla u(\cdot,t)\,dt = \chi_T$ and $\int_0^T \nabla\tilde{u}(\cdot,t)\,dt = \tilde{\chi}_T$ and from uniform ellipticity and continuity of a, we derive that

$$\begin{aligned}
&\int_{\mathbb{R}^d} \big(u(y,T)-\tilde{u}(y,T)\big)\big(\chi_T(y)-\tilde{\chi}_T(y)\big)\,dy \\
&\qquad \le 2\beta|B_\ell|^{1/2}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|_{L^2(B_\ell(z))} + 2\beta\left\|\nabla\tilde{\chi}_T\right\|_{L^2(B_\ell(z))}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|_{L^2(B_\ell(z))} \\
&\le \frac{2\beta^2}{\alpha}|B_\ell| + \frac{\alpha}{4}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(B_\ell(z))} + \frac{2\beta^2}{\alpha}\left\|\nabla\tilde{\chi}_T\right\|^2_{L^2(B_\ell(z))} + \frac{\alpha}{4}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(\mathbb{R}^d)}. \qquad (7.27)
\end{aligned}$$

The ellipticity of $a(\cdot)$, eqs. (7.25) to (7.27) lead to the bound

$$\begin{aligned}
\alpha\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(B_\ell(x))} &\le \alpha\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(\mathbb{R}^d)} \\
&\le \frac{\alpha}{4}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(\mathbb{R}^d)} + \frac{2\beta^2}{\alpha}\left\|\nabla\tilde{\chi}_T+\xi\right\|^2_{L^2(B_\ell(z))} \\
&+ \frac{2\beta^2}{\alpha}|B_\ell| + \frac{\alpha}{4}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(B_\ell(z))} + \frac{2\beta^2}{\alpha}\left\|\nabla\tilde{\chi}_T\right\|^2_{L^2(B_\ell(z))} + \frac{\alpha}{4}\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

Thus,

$$\left\|\nabla\chi_T-\nabla\tilde{\chi}_T\right\|^2_{L^2(B_\ell(x))} \le C\left\|\nabla\tilde{\chi}_T+\xi\right\|^2_{L^2(B_\ell(z))} + C(\alpha,\beta,\ell,d). \qquad (7.28)$$

Finally, by the triangle inequality we have that

$$\int_{B_\ell(x)}\left|\nabla\chi_T+\xi\right|^2 dy \le \int_{B_\ell(x)}\left|\nabla\tilde{\chi}_T+\xi\right|^2 dy + \int_{B_\ell(x)}\left|\nabla\chi_T-\nabla\tilde{\chi}_T\right|^2 dy.$$

We conclude by plugging eq. (7.28) into the line above. □

We can now prove a first result to bound the statistical error. Such a result is not conclusive in the sense that the upper bound on the statistical error are not derived explicitly with respect to the parameters L and T.

Theorem 7.11 (statistical error)**.** *Let $a^{0,T,L}$ be defined as in eq.* (7.5)*, and let*

$$S(x,z) := \left(\sup_{a|_{B_\ell(z)}}\int_{B_\ell(x)}\left|\nabla\chi_T(y)+\xi\right|^2 dy\right)^{1/2}, \quad \textit{and} \quad O(x,z) := \left(\int_{B_\ell(x)}\left|\operatorname*{osc}_{a|_{B_\ell(z)}}\nabla\chi_T(y)\right|^2 dy\right)^{1/2}.$$

Then, there exists $C>0$, independent of L and T, such that

$$\operatorname{Var}\left[\xi\cdot a^{0,T,L}\xi\right] \le CL^{-2d}\mathbb{E}\left[\int_{\mathbb{R}^d}\int_{K_L}\int_{K_L} S(x,z)O(x,z)S(x',z)O(x',z)\,dx\,dx'\,dz + \int_{\mathbb{R}^d} S(z,z)^2\,dz\right].$$

Proof. We first estimate $\underset{a|_{B_\ell(z)}}{\operatorname{osc}} \xi \cdot a^{0,T,L}\xi$ and, then, use the spectral gap inequality eq. (7.21). We start by defining $\tilde{a}^{0,T,L}$ as the homogenized version of tensor $\tilde{a}$ which is a random field that coincide with $a(\cdot)$ outside the ball $B_\ell(z)$ (we will omit the subscript of $\underset{a|_{B_\ell(z)}}{\operatorname{osc}}$), and we evaluate the difference

$$\begin{aligned}\xi\cdot\left(\tilde{a}^{0,T,L}-a^{0,T,L}\right)\xi &= \fint_{K_L}\left(\nabla\chi_T+\xi\right)\cdot a(x)\left(\nabla\chi_T+\xi\right)dx-\fint_{K_L}\left(\nabla\tilde{\chi}_T+\xi\right)\cdot\tilde{a}(x)\left(\nabla\tilde{\chi}_T+\xi\right)dx\\ &=\fint_{K_L}\left(\nabla\chi_T+\xi\right)\cdot a(x)\left(\nabla\chi_T-\nabla\tilde{\chi}_T\right)dx+\fint_{K_L}\left(\nabla\tilde{\chi}_T+\xi\right)\cdot\tilde{a}(x)\left(\nabla\chi_T-\nabla\tilde{\chi}_T\right)dx\\ &\quad+\fint_{K_L}\left(\nabla\chi_T+\xi\right)\cdot\left(a(x)-\tilde{a}(x)\right)\left(\nabla\tilde{\chi}_T+\xi\right)dx.\end{aligned}$$

Hence, we have the bound

$$\begin{aligned}\left|\xi\cdot\left(\tilde{a}^{0,T,L}-a^{0,T,L}\right)\xi\right| \le C\Bigg(&\fint_{K_L}\left|\nabla\chi_T+\xi\right|\left(\underset{a|_{B_\ell(z)}}{\operatorname{osc}}\nabla\chi_T\right)dx+\fint_{K_L}\left|\nabla\tilde{\chi}_T+\xi\right|\left(\underset{a|_{B_\ell(z)}}{\operatorname{osc}}\nabla\chi_T\right)dx\\ &+\frac{1}{|K_L|}\int_{B_{\ell(z)}}\left(\nabla\chi_T+\xi\right)\left(\nabla\tilde{\chi}_T+\xi\right)dx\Bigg).\end{aligned}$$

Before taking the supremum over all coefficients such that $\tilde{a}|_{\mathbb{R}^d\setminus B_\ell(z)}=a|_{\mathbb{R}^d\setminus B_\ell(z)}$ we introduce the new variable y in the first integral term via the bound $\fint_{K_L}dx\le C\fint_{K_L}\int_{B_\ell(x)}dy\,dx$, in order to use the estimate of Lemma 7.10, and then we use Cauchy-Schwarz inequality:

$$\begin{aligned}\underset{a|_{B_\ell(z)}}{\operatorname{osc}}\,\xi\cdot a^{0,T,L}\xi \le C\fint_{K_L}&\left(\sup_{a|_{B_\ell(z)}}\int_{B_\ell(x)}\left|\nabla\chi_T(y)+\xi\right|^2dy\right)^{1/2}\left(\int_{B_\ell(x)}\left|\underset{a|_{B_\ell(z)}}{\operatorname{osc}}\nabla\chi_T(y)\right|^2dy\right)^{1/2}dx\\ &+\frac{C}{|K_L|}\sup_{a|_{B_\ell(z)}}\int_{B_{\ell(z)}}\left|\nabla\chi_T+\xi\right|^2dx.\end{aligned}\tag{7.29}$$

By definition of $S(x,z)$ and $O(x,z)$, we conclude that

$$\underset{a|_{B_\ell(z)}}{\operatorname{osc}}\,\xi\cdot a^{0,T,L}\xi\le C\fint_{K_L}S(x,z)O(x,z)\,dx+\frac{C}{|K_L|}S(z,z).$$

Then, by applying the spectral gap inequality, we conclude:

$$\operatorname{Var}\left[\xi\cdot a^{0,T,L}\xi\right]\le\frac{C}{|K_L|^2}\mathbb{E}\left[\int_{\mathbb{R}^d}\int_{K_L}\int_{K_L}S(x,z)O(x,z)S(x',z)O(x',z)\,dx\,dx'\,dz+\int_{\mathbb{R}^d}S(z,z)^2\,dz\right].$$

□

The lack of explicit estimates for the term

$$\mathbb{E}\left[\int_{\mathbb{R}^d}\int_{K_L}\int_{K_L}S(x,z)O(x,z)S(x',z)O(x',z)\,dx\,dx'\,dz+\int_{\mathbb{R}^d}S(z,z)^2\,dz\right]$$

prevents us to bound the statistical error explicitly in terms of L and T. Nevertheless, from numerical test, we expect the convergence rate of the Central Limit Theorem:

$$\mathrm{Var}\left[\xi \cdot a^{0,T,L}\xi\right] \leq CL^{-d}.$$

Thus, we make the following conjecture:

Conjecture 1. *Let $S(x,z)$ and $O(x,z)$ be defined as in Theorem 7.11. There exists $C>0$, independent on T and L such that*

$$\mathbb{E}\left[\int_{\mathbb{R}^d}\left(\int_{K_L} S(x,z)O(x,z)\,dx\right)^2 dz\right] + \mathbb{E}\left[\int_{\mathbb{R}^d} S(z,z)^2\,dz\right] \leq CL^d.$$

7.3 Numerical experiments

In this section we collect the results of numerical experiments performed in order to verify numerically the correctness of the proved bound on the systematic error, Theorem 7.9, and to compare the convergence of the global resonance error for the standard numerical homogenization scheme and for the modified elliptic approach. In the coming numerical test we plot the mean resonance error, with the aim to eliminate the statistical error. This allows to display only the contribution of the boundary error, which is of higher order, and of the systematic error. Computing *exactly* the error in mean is not possible, because it requires to solve the integrate over the probability space Ω, but we approximate it by computing the empirical average

$$\bar{a}^{0,R,L,T,N} = \frac{1}{N}\sum_{k=1}^{N} a^{0,R,L,T,k}, \tag{7.30}$$

which converges in probability to $\mathbb{E}\left[a^{0,R,L,T}\right]$, by the weak law of large numbers. Upon choosing a sufficiently large number of samples, N, the difference between $\bar{a}^{0,R,L,T,N}$ and $\mathbb{E}\left[a^{0,R,L,T}\right]$ becomes negligible. Computing the mean average does not take into account the effect of stochastic variability of the approximations of a^0, and it does not allow to conclude anything about the convergence in the probability space. The strong resonance error includes the convergence in the probability space, at the cost of having the additional statistical error term in the upper bounds. The mean square error cannot be computed exactly, so we approximate it by its empirical version

$$s_{R,L,T,N} := \sqrt{\frac{\sum_{k=1}^{N}\left\|a^{0,R,L,T,k} - \bar{a}^0\right\|_F^2}{N-1}}, \tag{7.31}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

7.3.1 The covariance function of random fields

For simplicity, we will consider only isotropic media, for which the heterogeneous tensor can be written as $a(x) = f(x)I$, where $f : \mathbb{R}^d \mapsto \mathbb{R}$ and $I \in \mathbb{R}^{d\times d}$ is the identity matrix. With a slight abuse of notation we will denote both the matrix of coefficients and the function f above by $a(x)$. Stationarity of the random fields implies that $\mathbb{E}[a(\cdot)] = \mu I_{d\times d}$ does not depend on the spatial variable x and that the covariance (matrix-valued) function $\mathrm{Cov}(x, y)$ defined as

$$\mathrm{Cov}(x, y) := \mathbb{E}\left[a(x)a(y)\right] - \mu^2$$

only depends on the distance $|x - y|$. So, for stationary random fields, there exist a function $r : \mathbb{R}^d \mapsto \mathbb{R}$ such that

$$\mathrm{Cov}(x, y) = r(x - y).$$

When the function $r(\cdot)$ is radial, i.e. $r(t) = r(|t|)$ the medium is said to be *statistically isotropic*. Several choices for the covariance function are possible. For example, widely used classes of covariance functions for one dimensional isotropic random fields are the exponential covariance function:

$$r(t) = \sigma^2 e^{-\left|\frac{t}{l}\right|},$$

and the Matèrn covariance function:

$$r(t) = \sigma^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\sqrt{2\nu}\left|\frac{t}{l}\right|\right)^{\nu} K_{\nu}\left(\sqrt{2\nu}\left|\frac{t}{l}\right|\right),$$

where σ^2 is the variance, l is the correlation length, Γ is the gamma function, K_ν is the modified Bessel function of the second kind and ν is a smoothness parameter. Another choice is the long-range covariance function:

$$r(t) = (1 + |t|)^{-1/2}. \tag{7.32}$$

All random fields considered in the numerical experiments are generated by the circulant embedding method described in [48].

7.3.2 Optimal scaling of T vs. R

Here we briefly discuss the optimal scaling of the T parameter as a function of R (and L) with the aim of maximizing the rate of decay of the error in mean. We will assume that the boundary error term decays exponentially, as in the periodic case:

$$\mathbb{E}\left[\left\|a^{0,R,L,T} - a^{0,L,T}\right\|_F^2\right]^{\frac{1}{2}} \leq C_1 \exp\left(-c_2 \frac{|R - L|^2}{T}\right),$$

for some constants $C_1, c_2 > 0$. Then, we find the regime under which none of the boundary and systematic errors is dominating but the two are (approximately) equal by imposing that

they are equal:

$$C_1 \exp\left(-c_2 \frac{|R-L|^2}{T}\right) = C_2 T^{-\frac{d}{2}} \implies T \log\left(\left(\frac{C_1}{C_2}\right)^{\frac{2}{d}} T\right) = \frac{2c_2}{d} |R-L|^2 .$$

The constants C_1, C_2 are unknown and problem-dependent, but we can conclude that the optimal scaling is obtained for

$$c\,|R-L| \leq T \leq C\,|R-L|^2 , \tag{7.33}$$

with $c, C > 0$. In the numerical experiments we will use these scaling values for T, and we take $L = k_o R$, with $0 < k_o < 1$.

7.3.3 One dimensional logit-normal random coefficients

We test the convergence of the approximate homogenized coefficient for a random diffusion coefficient distributed according to the logit-normal law. A logit-normal random field is an isotropic random field $a(\cdot) \in \Omega$ of the form

$$a(x) = \frac{b + e^{-\kappa(Z(x)-z_0)}}{c + e^{-\kappa(Z(x)-z_0)}};$$

where $b, c, \kappa, z_0 \in \mathbb{R}$ and Z is a Gaussian random field of zero average. For this example, we set $b = 2$, $c = 1$, $\kappa = 1$ and $x_0 = 0$ and used a Gaussian random field with Matèrn covariance function of order $\nu = 3/2$. A representation of such a field is depicted in Figure 7.1b. In the one dimensional case, the homogenized coefficient can be computed by the harmonic mean: $a^0 = \mathbb{E}\left[a(\cdot)^{-1}\right]^{-1}$. Hence, in the logit-normal case,

$$a^0 = \left(\int_{\mathbb{R}} a(y)^{-1} f_Z(y)\, dy\right)^{-1},$$

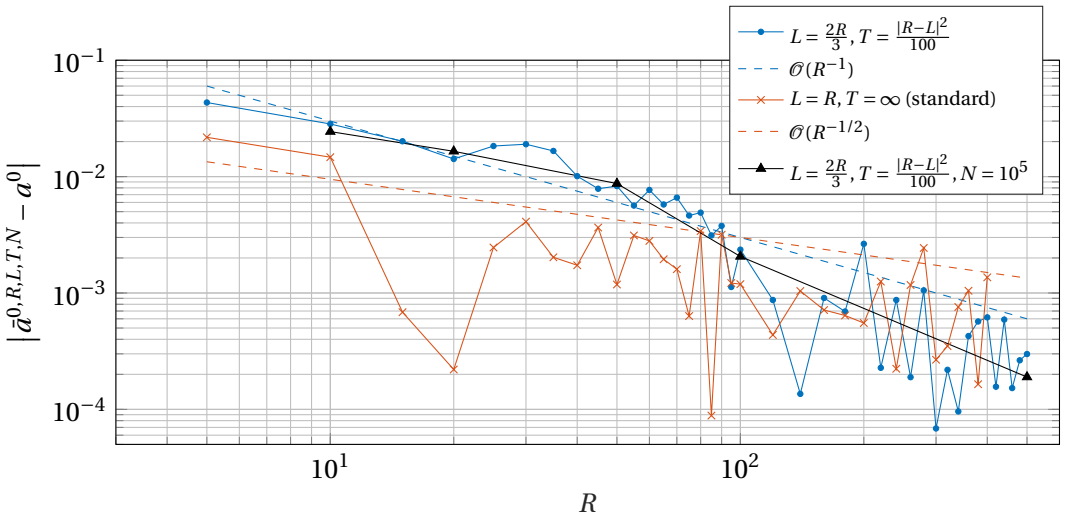
where f_Z is the Gaussian probability density function.

We computed the approximation to the homogenized coefficients by Finite Elements (FE) discretization on a grid with mesh size $h = 2^{-8}$. The modified auxiliary problem eq. (7.1) is solved over the domain is $K_R := (-R/2, R/2)$ with periodic boundary conditions, with the values of R ranging from 5 to 500. The other parameters are $L = 2R/3$, for the size of the averaging domain K_L, and T, for the modified forcing term. The approximation to the homogenized coefficients are computed as in eq. (7.4). As an approximation of the quasi-optimal scaling eq. (7.33), we choose

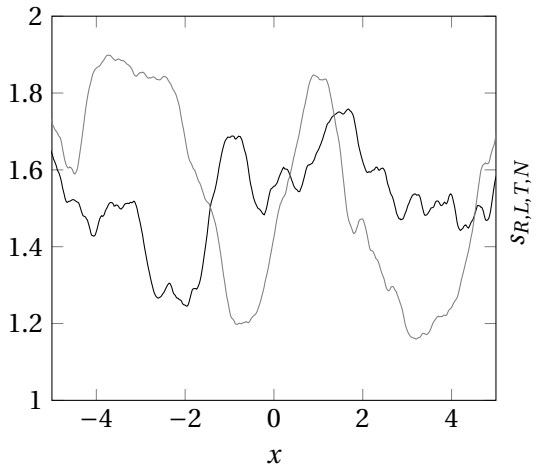
$$T = \frac{|R-L|^2}{100}.$$

The right-hand side of eq. (7.1) is approximated in the FE space by the exponential matrix $e^{-M_h^{-1}AT}\mathbf{g}$, where $\mathbf{g}$ is the vector of components of the projection of $g(\cdot) = \frac{d}{dx}a(\cdot)$ in the FE space, M_h is the lumped mass matrix and A_h is the stiffness matrix. The exponential matrix is

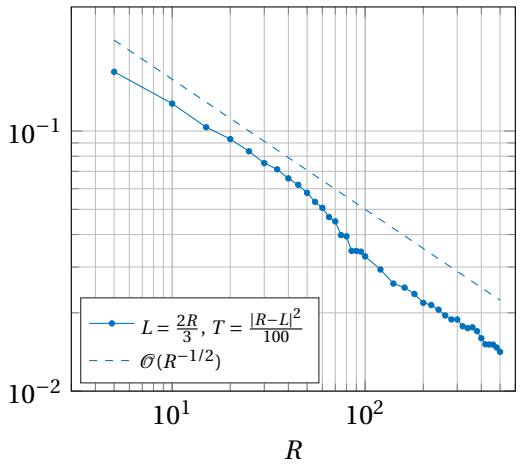
not computed exactly, but it is approximated in the Krylov subspace generated by $M_h^{-1}A_h\mathbf{g}$ and computed by the Lanczos method ($M_h^{-1}A_h$ is symmetric and positive definite) as proposed in [94]. The maximum number of Krylov basis elements is 2000. The error in mean between the



(a) Error in mean for the standard and modified elliptic methods.



(b) Two realization of the field.



(c) Mean square error.

Figure 7.1 – Logit-normal random field with Matèrn covariance function of order $\nu = 3/2$.

approximate and the exact homogenized coefficient eq. (7.16) is plotted in Figure 7.1a. Since the expected value of $a^{0,R,L,T}$ cannot be computed exactly, we approximate it by the empirical average eq. (7.30) with $N = 1000$ i.i.d. samples. The red line shows the error decay for the standard auxiliary problem with periodic BCs and no oversampling. In this case, the only source of error is due to the BCs. The error for the modified elliptic approach is represented by the blue line. In this other case, the global error is the contribution of the boundary and systematic error. As one can see, the red curve in Figure 7.1a decays at a slow rate of $R^{-\frac{1}{2}}$. The blue curve in Figure 7.1a follows a faster convergence trend of, approximately, $\mathcal{O}(R^{-1})$, thanks

to the scaling of T. Figure 7.1a also displays the error in mean when $N = 10^5$ Monte Carlo samples are chosen. The plot does not show any difference with respect to the other cases, so we conclude that the error decay does not depend on the number of samples.

In Figure 7.1c we report the mean square resonance error. It decays as $R^{-d/2}$, following the Central Limit Theorem trend. This means that the statistical error dominates the other errors and that it satisfies the Conjecture 1.

7.3.4 One dimensional lognormal random coefficients

We test the convergence of the approximate homogenized coefficient for a random diffusion coefficient distributed according to the lognormal law. A lognormal random field is an isotropic random field $a(\cdot) \in \Omega$ for which there exist $b, c > 0$ such that

$$a(x) = ce^{bZ(x)},$$

where Z is a Gaussian random field of zero average and long range covariance function given in eq. (7.32). For this test, we have chosen $b = c = 1$; an example of the field in shown in Figure 7.2a. The model of lognormally distributed random coefficients is widely used in the environmental engineering community, see e.g. [52, 57, 115]. However, such a coefficient does not belong to $\mathscr{M}(\alpha, \beta)$, so it is not guaranteed that it follows the theoretical estimates that we derived in the previous sections. In the one dimensional case, the homogenized coefficient can be computed by the harmonic mean: $a^0 = \mathbb{E}\left[a(\cdot)^{-1}\right]^{-1}$. Hence, in the lognormal case,

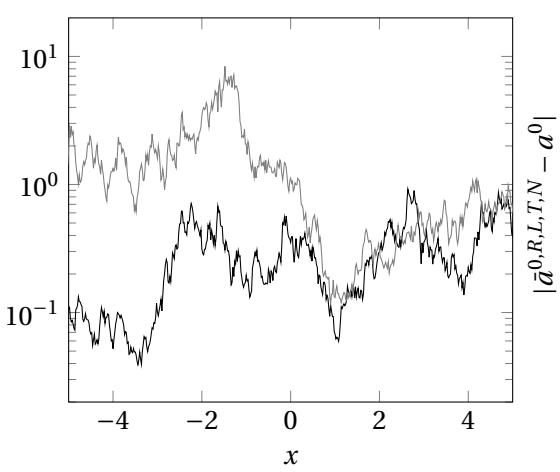
$$a^0 = \left(\int_{\mathbb{R}} a(y)^{-1} f_Z(y)\, dy\right)^{-1} = ce^{-b^2/2},$$

where f_Z is the Gaussian probability density function. We computed the approximation to the homogenized coefficients by FE discretization on a grid with mesh size $h = 2^{-8}$. The modified auxiliary problem eq. (7.1) is solved over the domain is $K_R := (-R/2, R/2)$ with periodic boundary conditions. The values of the R parameter vary from 5 to 500. The other parameters are $L = 2R/3$, for the size of the averaging domain K_L, and T, for the modified forcing term. The approximation to the homogenized coefficients are computed as in eq. (7.4). Since the lognormal random field does not belong to the class $\mathscr{M}(\alpha, \beta)$, it is not possible to choose the value of the T parameter according to the optimal scaling derived in eq. (7.33). So, in our computations, we choose the value of T as

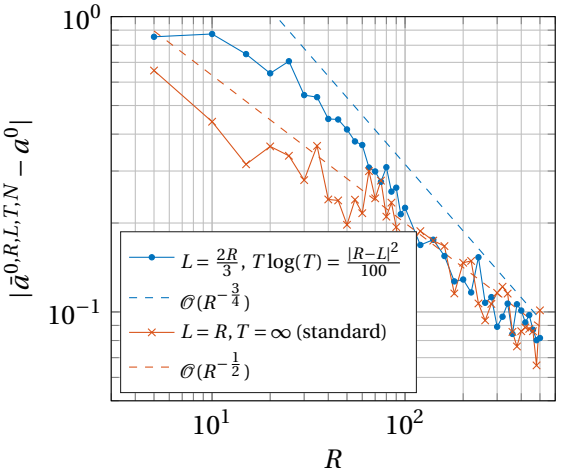
$$T \log(T) = \frac{|R - L|^2}{100}.$$

The right-hand side of eq. (7.1) is approximated in the FE space as in the previous example.

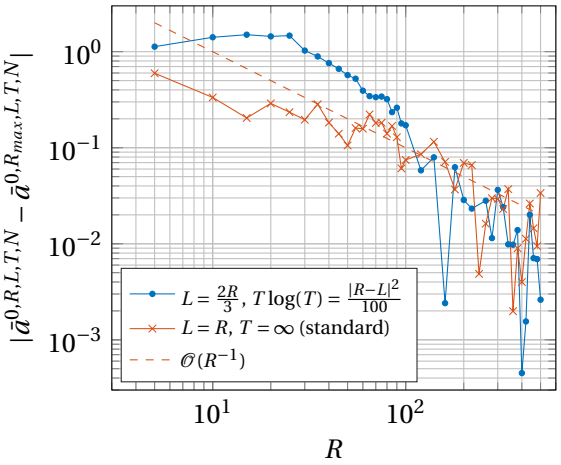
The error in mean between the approximate and the exact homogenized coefficient eq. (7.16) is plotted in Figure 7.2b. The expected value of $a^{0,R,L,T}$ is approximated by independently drawing $N = 1000$ samples of the lognormal random field. The red line of Figure 7.2b shows



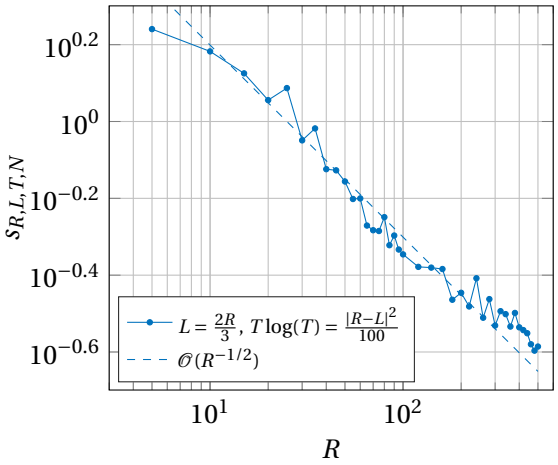
(a) Two realizations of the field.



(b) Error in mean for the standard and modified elliptic methods.



(c) Boundary error for the standard and modified elliptic methods.



(d) Mean square error.

Figure 7.2 – Lognormal random field with covariance function $r(t) = (1+|t|)^{-1/2}$.

the error decay for the standard auxiliary problem with periodic BCs and no oversampling, while the blue line displays the decay of the error for the modified elliptic method. In the first case, the only source of error are the BCs, while in the second case the error is made up of two contributions: the boundary error which converges exponentially and is more visible in the range of small domains ($R < 100$), and the the systematic error dominating for larger values of R.

Next, we show the decay of the boundary error. The boundary error is defined as the difference $\mathbb{E}\left[a^{0,R,L,T} - a^{0,L,T}\right]$ and it is supposed to be controlled by an exponentially decaying deterministic upper bound. The values of $a^{0,L,T}$ are not directly accessible (they involve the solution of the corrector problem over the infinite domain), so we approximate them

by $a^{0,R_{max},L,T} \approx a^{0,L,T}$, with $R_{max} = 500$. Additionally, we average $a^{0,R,L,T}$ and $a^{0,R_{max},L,T}$ over $N = 1000$ i.i.d. samples. The exponentially decaying (in R) difference between the empirical averages $\bar{a}^{0,R,L,T,N}$ and $\bar{a}^{0,R_{max},L,T,N}$ is depicted in Figure 7.2c.

We conclude by underlying that the statistical error decays with the expected rate of $R^{-1/2}$ also in this case, Figure 7.2d.

7.3.5 Two dimensional lognormal field with exponential covariance

As a last numerical test, we study the convergence for a two dimensional lognormal random field with exponential covariance function, such as the one depicted in Figure 7.3a. The field is sampled by generating a Gaussian random field over the uniform grid

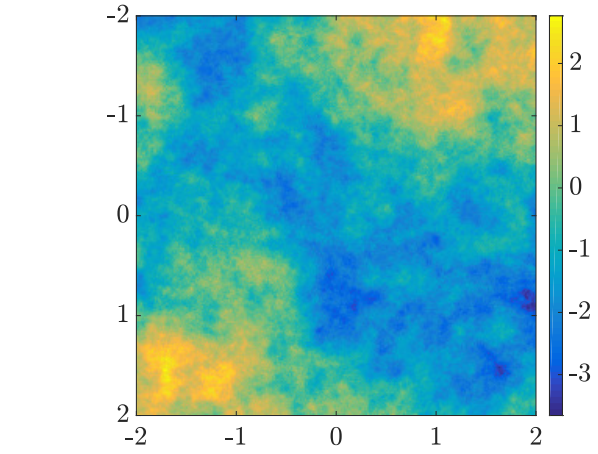
$$\left\{(x_i, y_j) \in K_R : x_i = ih - \frac{R}{2}, y_j = jh - \frac{R}{2}\right\},$$

coinciding with the set of vertices of the structured mesh of stepsize $h = 2^{-5}$ on K_R. Also in this case we have chosen the parameters $b = c = 1$, as in Section 7.3.4. The covariance function of the Gaussian field is exponential. Two representations of the field are depicted in Figures 7.3a and 7.3b.

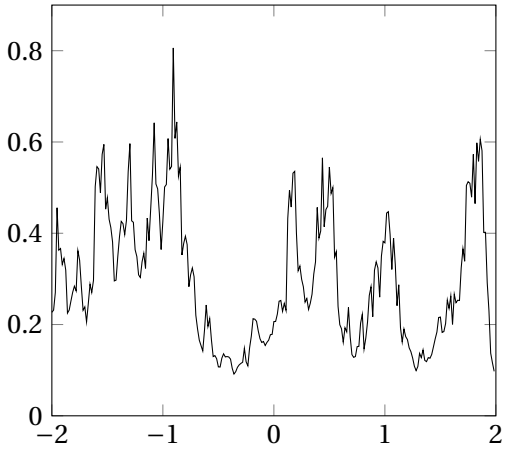
The correctors are computed by the finite element method with $\mathbb{P}_1$-elements, and the right-hand side is calculated by the Krylov subspace method with up to 2000 basis elements. The average is approximated by drawing $N = 200$ i.i.d. samples of the lognormal field. The convergence behaviour of the mean error is pictured in Figure 7.3c for both the method discussed in this work and the truncated domain approach of eq. (6.14). The choice of the modelling parameters R, L, T is reported in Figure 7.3a, and R ranges from 5 to 32. In this case we notice that the convergence rates improve to 1 and 3/2 for, respectively, the standard and the modified elliptic approaches. Figure 7.3d proves that the statistical error decays as $R^{-d/2}$.

7.4 Conclusion

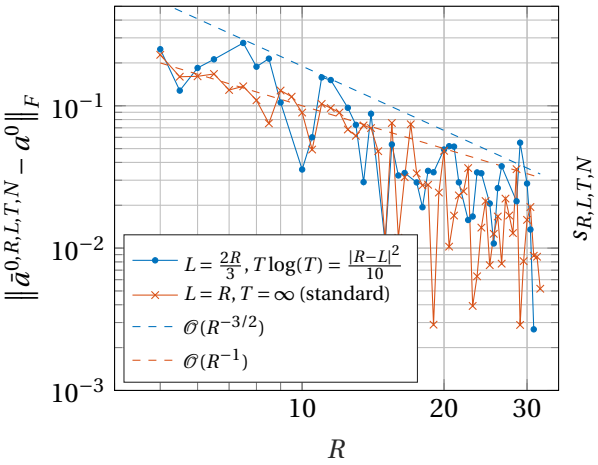
In this chapter we addressed the problem of estimating the systematic error for the modified elliptic model, defined as the difference between the true effective coefficient a^0 and its approximation by the corrector problem eq. (7.6) over the infinite domain $\mathbb{R}^d$ and we discussed an attempt to bound the statistical error. By exploiting the time decay properties of solutions of linear parabolic equations, we found that the systematic error scales as $T^{-d/2}$, where T is the final time. The parabolic solution, evaluated at time T, enters as a source term into eq. (7.2). Through the same time decay properties, it is possible to prove the existence of a corrector $\nabla\chi_T \in \mathscr{L}^2_{pot}$. In comparison to the standard approach, the numerical experiments show that the modified elliptic method slightly improves the error convergence rate with



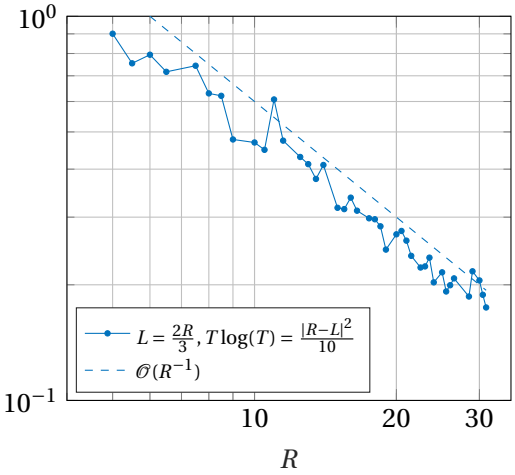
(a) Logarithmic colour scale.



(b) Field along the line $y = 0$.



(c) Error in mean for the standard and modified elliptic methods.



(d) Mean square error.

Figure 7.3 – Lognormal random field with exponential covariance function.

respect to the cell size R and upon choosing the final time T such that

$$c\,|R-L| \le T \le C\,|R-L|^2\,,$$

for some $c, C > 0$. Moreover, the theoretical bound of Theorem 7.9 is verified.

8 Conclusion and outlook

8.1 Conclusion

The present study aimed to develop novel upscaling techniques to reduce the resonance error which affects numerical homogenization methods and limits the accuracy of multiscale simulations. If the standard correctors are employed, the resonance error decays as R^{-1}, where R is the size of the domain of the corrector problem, see Chapter 2. Improving the convergence rate is thus crucial for achieving accurate solutions with a reasonable computational cost.

In this thesis, we presented, developed and analysed two novel micro-corrector problems, defined by eqs. (4.2) and (5.1), to improve the convergence rate of the resonance error for second order linear elliptic equations with fast oscillating coefficients. The two methods are based on the relations derived in Theorem 3.2 that interprets the solution and the energy of an elliptic equation as time integrals of the solution and energy of a parabolic problem. The convergence rates of the resonance error associated to the two models were derived under the assumptions of both periodic and stationary random coefficients in Theorems 4.1, 5.1 and 7.9.

The first approach consists in approximating a^0 by solving the parabolic corrector equation eq. (4.2) and employing eq. (4.1). The approximation of a^0 and the *a priori* error bound on the resonance error depend on the modelling parameters: the size of the corrector cell, denoted by R, the size of the support of the filtering function, L, and the final time, T. Our analysis holds under the assumption of periodicity of the coefficients, but the method that we propose can be applied also to non-periodic structures. The use of q-fold differentiable filtering functions, together with the weak dependence of the solution on the boundary conditions, allows to prove *a priori* upper bounds on the resonance error with arbitrary convergence rate with respect to L, T and $|R-L|^2/T$. By the optimal scaling $L, T = \mathcal{O}(R)$, it is possible to achieve arbitrary convergence rates in R, as we proved in Theorem 4.1. The proof is based on three results:

i) the approximation property of filtering functions stated in Lemma 4.3;

ii) the exponential decay in time of the L^2-norm of the periodic correctors;

iii) the exponential decay in space of the Green's function for parabolic problems, as stated by the Nash-Aronson inequality.

The parabolic approach inspired the development of a modified elliptic approach where, in addition to the standard elliptic corrector equation, the right-hand side $e^{-TA}g^i$ is present. *A priori* resonance error bounds for this homogenization approach were derived using the results of Lemma 4.15 on the exponential decay of the boundary layer functions θ^i's. This approach overcomes the use of a time integration scheme to solve the parabolic equation, but the right-hand side $e^{-TA}g^i$ must be pre-computed. This is the main computational issue of the modified elliptic approach and we addressed it by computing the exponential of a properly chosen matrix, A_h. Many techniques are available to approximate e^{-TA_h}, among which we chose the Krylov subspace method. The modified elliptic approach is endowed with the same convergence results as the parabolic scheme, as proved in Theorem 5.1. Many numerical experiments supported our theoretical results, even for more general class of coefficients. Both the parabolic and the modified elliptic methods significantly improve the convergence rates of the resonance error, in the periodic and quasi-periodic case. Moreover, they can also be used to alleviate the boundary error in problems with stationary random coefficients. Additionally, these novel approaches show a more favourable cost-accuracy ratio, in comparison to the standard method: our theoretical analysis revealed that the growth of the computational cost for increasing desired accuracy is slower for the two approaches than for the standard one.

As a last step, we focused on a more challenging question: estimating the resonance error for the modified elliptic approach in the framework of stochastic homogenization. Finding efficient ways to compute the effective coefficients for random micro-structures still represents a great challenge from the mathematical, computational and practical points of view, and it could have tremendous applications in material sciences and engineering. In this setting, specific micro-models and numerical techniques have to be developed, since the tools that are employed in the periodic context, such as filtering functions, cease to be useful in the random case. The recent monograph [19] could provide a deeper understanding of the mathematical structure of stochastic homogenization and could be of help in the analysis of novel numerical upscaling methods.

In Chapter 7, the resonance error is decomposed into a systematic and a statistical error (variance) and the two terms are studied separately. The error analysis that we have carried out reveals that the truncation/systematic error decays as $T^{-d/2}$, while it shows the exponential convergence $e^{-\alpha T}$ in the periodic case. Therefore, the decay of the resonance error in the random case is severely hampered, independently from the use of filtering functions. On the other hand, the statistical error can be reduced by Monte Carlo iterations or by variance reduction techniques, as described in Chapter 6. The ergodic assumption allows to reduce the variance of the (random) upscaled coefficients by taking larger cells for the corrector problem. Numerical tests show that the statistical error decays as $L^{-d/2}$ allowing us to conclude that

less Monte Carlo iterations are needed when larger cells are considered.

8.2 Outlook

Within the scope of the present study, many objectives are achieved and the accomplishments are presented above. However, several challenges remain to be addressed and we suggest that further research could focus on the aspects which are detailed below.

- In this work, we have focused on the derivation of homogenized coefficients for scalar equations with symmetric coefficients. Further research may involve the extension of the two approaches to systems of differential equations, such as linear elasticity problems, or to equations with non-symmetric coefficients. In both cases the theory is well-developed, e.g. [45], but it would be necessary to employ the heat kernel bounds for parabolic systems, which could be found, e.g., in [43, 75] and references therein.
- Multiscale non-linear differential equations are of great interest from the point of view of mechanical applications [73] and of mathematical theory [122]. Further research could address the extension of the parabolic and modified elliptic methods to quasi-linear and non-linear problems, in order to evaluate their error convergence properties. This task could be quite challenging, as the homogenized equation is not explicitly constructed but is defined as the G-limit of a sequence of (non-linear) operators.
- In Chapters 4 and 5, the cost-accuracy ratios are derived by theoretical speculations on the computational cost of the time integration scheme (for the parabolic case) or the solution of the linear system (for the modified elliptic method). This approach explains how the cost scales with the desired accuracy, but it cannot predict the wall- and CPU-times. In practical situations, if a rough accuracy is accepted, using the standard upscaling scheme over a small cell could reveal more efficient than other approaches. Computational experiments could address the question of evaluating for which values of the cell size it is more convenient to use the standard approach or the ones that we propose. Optimized implementation in interpreted programming language (such as C/C++) may be necessary to conduct this study.
- In the stochastic analysis of Chapter 7, we proved numerically that the statistical error decays as $L^{-d/2}$, but a rigorous proof is still missing, though we proposed a way to address it. Future research could focus on proving this bound.
- The resonance error in the numerical homogenization of random coefficients cannot convergence faster than $L^{-d/2}$ (see [119]), due to the statistical error. In order to make sure that this optimal convergence behaviour is attained, it is crucial to design novel corrector equations with high decay rates of the systematic and boundary errors, as the one proposed in Chapter 7. However, the use of alternative corrector equation with a faster convergence of the systematic error can be beneficial and would allow to achieve the optimal convergence rate.

- In connection to the point above, although the convergence rate of the statistical error cannot be improved, several works attempted either to mitigate the computational effort or to reduce the variance of the approximation [59, 108]. The convergence rate of the Quasi-Monte Carlo method for scalar diffusion equations with lognormal random diffusion fields has been analysed in [91]. This technique could be successfully applied to improve the convergence rate with respect to the number of samples in the context of modified elliptic corrector problems for numerical homogenization, but research works in this direction are still missing.

Bibliography

[1] Abdulle, A. (2002). Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23(6):2041–2054.

[2] Abdulle, A. (2005). On a priori error analysis of fully discrete heterogeneous multiscale FEM. *Multiscale Model. Simul.*, 4(2):447–459.

[3] Abdulle, A. (2006). Analysis of a heterogeneous multiscale FEM for problems in elasticity. *Math. Models Methods Appl. Sci.*, 16(4):615–635.

[4] Abdulle, A. (2009). The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. In *Multiple scales problems in biomathematics, mechanics, physics and numerics*, volume 31 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 133–181. Gakkōtosho, Tokyo.

[5] Abdulle, A. (2011). A priori and a posteriori error analysis for numerical homogenization: a unified framework. *Ser. Contemp. Appl. Math. CAM*, 16:280–305.

[6] Abdulle, A., Arjmand, D., and Paganoni, E. (2019). Exponential decay of the resonance error in numerical homogenization via parabolic and elliptic cell problems. *C. R. Math. Acad. Sci. Paris*, 357:545–551.

[7] Abdulle, A., Arjmand, D., and Paganoni, E. (2020a). Analytical and numerical study of a modified cell problem for the numerical homogenization of multiscale random fields. *Submitted, preprint available at https://arxiv.org/abs/2007.10828*.

[8] Abdulle, A., Arjmand, D., and Paganoni, E. (2020b). An elliptic local problem with exponential decay of the resonance error for numerical homogenization. *Submitted, preprint available at https://arxiv.org/abs/2001.06315*.

[9] Abdulle, A., Arjmand, D., and Paganoni, E. (2020c). A parabolic local problem with exponential decay of the resonance error for numerical homogenization. *Submitted, preprint available at https://arxiv.org/abs/2001.05543*.

[10] Abdulle, A. and Bai, Y. (2012). Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems. *J. Comput. Phys.*, 231(21):7014–7036.

[11] Abdulle, A. and E, W. (2003). Finite difference heterogeneous multi-scale method for homogenization problems. *J. Comput. Phys.*, 191(1):18–39.

[12] Abdulle, A. and Grote, M. J. (2011). Finite element heterogeneous multiscale method for the wave equation. *Multiscale Model. Simul.*, 9(2):766–792.

[13] Abdulle, A. and Medovikov, A. A. (2001). Second order Chebyshev methods based on orthogonal polynomials. *Numer. Math.*, 90(1):1–18.

[14] Abdulle, A. and Schwab, C. (2005). Heterogeneous multiscale FEM for diffusion problems on rough surfaces. *Multiscale Model. Simul.*, 3(1):195–220.

[15] Abdulle, A. and Vilmart, G. (2012). A priori error estimates for finite element methods with numerical quadrature for nonmonotone nonlinear elliptic problems. *Numer. Math.*, 121(3):397–431.

[16] Arjmand, D. and Runborg, O. (2016). A time dependent approach for removing the cell boundary error in elliptic homogenization problems. *J. Comput. Phys.*, 314(Supplement C):206–227.

[17] Arjmand, D. and Runborg, O. (2017). Estimates for the upscaling error in heterogeneous multiscale methods for wave propagation problems in locally periodic media. *Multiscale Model. Simul.*, 15(2):948–976.

[18] Arjmand, D. and Stohrer, C. (2016). A finite element heterogeneous multiscale method with improved control over the modeling error. *Communications in Mathematical Sciences*, 14(2):463–487.

[19] Armstrong, S., Kuusi, T., and Mourrat, J.-C. (2019). *Quantitative Stochastic Homogenization and Large-Scale Regularity*. Springer.

[20] Aronson, D. G. (1963). On the Green's function for second order parabolic differential equations with discontinuous coefficients. *Bull. Amer. Math. Soc.*, 69(6):841–847.

[21] Aronson, D. G. (1967). Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73(6):890–896.

[22] Aronson, D. G. (1968). Non-negative solutions of linear parabolic equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, Ser. 3, 22(4):607–694.

[23] Babuška, I., Banerjee, U., and Osborn, J. E. (2003). Survey of meshless and generalised finite elements methods: A unified approach. *Acta Numerica*, pages 1–125.

[24] Babuška, I., Banerjee, U., and Osborn, J. E. (2004). Generalized finite element methods — main ideas, results and perspective. *International Journal of Computational Methods*, 01(01):67–103.

[25] Babuška, I., Caloz, G., and Osborn, J. E. (1994). special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, 31:945–981.

[26] Babuška, I. and Lipton, R. (2011). Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. *Multiscale Model. Simul.*, 9:373–406.

[27] Babuška, I. and Melenk, J. M. (1997). The partition of unity method. *Internat. J. Numer. Methods Engrg.*, 40:727–758.

[28] Babuška, I. and Osborn, J. E. (1983). Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.*, 20:510–536.

[29] Bai, Y. (2013). *Reduced order modeling techniques for numerical homogenization methods applied to linear and nonlinear multiscale problems*. PhD thesis, EPFL, Lausanne.

[30] Barrault, M., Maday, Y., Nguyen, N.-C., and Patera, A. T. (2004). An 'empirical interpolation method': Application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris*, Ser.I 339:667–672.

[31] Bensoussan, A., Lions, J.-L., and Papanicolaou, G. (1978). *Asymptotic analysis for periodic structures*. North-Holland Publishing Co., Amsterdam.

[32] Blanc, X., Costaouec, R., Le Bris, C., and Legoll, F. (2012). Variance reduction in stochastic homogenization: The technique of antithetic variables. In Engquist, B., Runborg, O., and Tsai, Y.-H. R., editors, *Numerical Analysis of Multiscale Computations*, pages 47–70, Berlin, Heidelberg. Springer Berlin Heidelberg.

[33] Blanc, X. and Le Bris, C. (2010). Improving on computation of homogenized coefficients in the periodic and quasi-periodic settings. *Netw. Heterog. Media*, 5(1):1–29.

[34] Blanc, X., Le Bris, C., and Legoll, F. (2016). Some variance reduction methods for numerical stochastic homogenization. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2066):20150168.

[35] Bourgeat, A. and Piatnitski, A. (2004). Approximations of effective coefficients in stochastic homogenization. *Ann. Inst. Henri Poincaré Probab. Stat.*, 40(2):153–165.

[36] Brezis, H. (2010). *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media.

[37] Brezzi, F., Franca, L. P., Hughes, T. J. R., and Russo, A. (1997). $b = \int g$. *Comput. Methods Appl. Mech. Engrg.*, 145(3-4):329–339.

[38] Budáč, O. (2016). *Multiscale methods for Stokes flow in heterogeneous media*. PhD thesis, École Polytechnique Fédérale de Lausanne, Lausanne.

[39] Cancès, E., Castella, F., Chartier, P., Faou, E., Le Bris, C., Legoll, F., and Turinici, G. (2004). High-order averaging schemes with error bounds for thermodynamical properties calculations by molecular dynamics simulations. *The Journal of Chemical Physics*, 121(21):10346–10355.

[40] Cancès, E., Castella, F., Chartier, P., Faou, E., Le Bris, C., Legoll, F., and Turinici, G. (2005). Long-time averaging for integrable hamiltonian dynamics. *Numer. Math.*, 100:21–232.

[41] Cances, E., Ehrlacher, V., Legoll, F., and Stamm, B. (2015). An embedded corrector problem to approximate the homogenized coefficients of an elliptic equation. *Comptes Rendus Mathematique*, 353(9):801–806.

[42] Cancès, E., Ehrlacher, V., Legoll, F., Stamm, B., and Xiang, S. (2020). An embedded corrector problem for homogenization. part ii: Algorithms and discretization. *Journal of Computational Physics*, 407:109254.

[43] Choi, J. and Kim, S. (2014). Green's functions for elliptic and parabolic systems with robin-type boundary conditions. *Journal of Functional Analysis*, 267(9):3205–3261.

[44] Ciarlet, P. G. (1978). *The finite element method for elliptic problems*, volume 4 of *Studies in Mathematics and its Applications*. North-Holland.

[45] Cioranescu, D. and Donato, P. (1999). *An introduction to homogenization*, volume 17 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, New York.

[46] Dal Maso, G. and Modica, L. (1986a). Nonlinear stochastic homogenization. *Annali di Matematica Pura ed Applicata*, 144(1):347–389.

[47] Dal Maso, G. and Modica, L. (1986b). Nonlinear stochastic homogenization and ergodic theory. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1986(368):28–42.

[48] Dietrich, C. R. and Newsam, G. N. (1997). Fast and exact simulation of stationary gaussian processes through circulant embedding of the covariance matrix. *SIAM Journal on Scientific Computing*, 18(4):1088–1107.

[49] Druskin, V. and Knizhnerman, L. (1998). Extended krylov subspaces: Approximation of the matrix square root and related functions. *SIAM Journal on Matrix Analysis and Applications*, 19(3):755–771.

[50] Druskin, V., Knizhnerman, L., and Zaslavsky, M. (2009). Solution of large scale evolutionary problems using rational krylov subspaces with optimized shifts. *SIAM Journal on Scientific Computing*, 31(5):3760–3780.

[51] Dunford, N. and Schwartz, J. T. (1988). *Linear operators*. Wiley.

[52] Durlofsky, L. J. (1991). Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. *Water Resour. Res.*, 27(5):699–708.

[53] E, W. (2011). *Principles of multiscale modeling*. Cambridge University Press, Cambridge.

[54] E, W. and Engquist, B. (2003). The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132.

[55] E, W., Ming, P., and Zhang, P. (2005). Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.*, 18(1):121–156.

[56] E, W. and Vanden-Eijden, E. (2008). Some critical issues for the "equation-free" approach to multiscale modeling. arXiv:0806.1621 [math.NA].

[57] Eberhard, J., Attinger, S., and Wittum, G. (2004). Coarse graining for upscaling of flow in heterogeneous porous media. *Multiscale Modeling & Simulation*, 2(2):269–301.

[58] Efendiev, Y. and Hou, T. Y. (2009). *Multiscale finite element methods. Theory and applications*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York.

[59] Efendiev, Y., Kronsbein, C., and Legoll, F. (2015). Multilevel monte carlo approaches for numerical homogenization. *SIAM Multiscale Model. Simul.*, 13(4):1107–1135.

[60] Efendiev, Y. R., Hou, T. Y., and Wu, X.-H. (2000). Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, 37(3):888–910.

[61] Egloffe, A.-C., Gloria, A., Mourrat, J.-C., and Nguyen, T. N. (2014). Random walk in random environment, corrector equation and homogenized coefficients: from theory to numerics, back and forth. *IMA Journal of Numerical Analysis*, 35(2):499–545.

[62] Engquist, B., Holst, H., and Runborg, O. (2011). Multi-scale methods for wave propagation in heterogeneous media. *Commun. Math. Sci.*, 9(1).

[63] Engquist, B., Holst, H., and Runborg, O. (2012). Multiscale methods for wave propagation in heterogeneous media over long time. In *Numerical analysis of multiscale computations*, pages 167–186. Springer.

[64] Engquist, B., Lötstedt, P., and Runborg, O. (2005). *Multiscale Methods in Science and Engineering*, volume 44. Springer.

[65] Engquist, B., Lötstedt, P., and Runborg, O. (2009). *Multiscale Modeling and Simulation in Science*, volume 66. Springer.

[66] Engwer, C., Henning, P., Målqvist, A., and Peterseim, D. (2019). Efficient implementation of the localized orthogonal decomposition method. *Computer Methods in Applied Mechanics and Engineering*, 350:123 – 153.

[67] Evans, L. C. (2010). *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[68] Fan, J., Kim, K., Nagayasu, S., and Nakamura, G. (2013). A gradient estimate for solutions to parabolic equations with discontinuous coefficients. *Electronic Journal of Differential Equations*, 2013(93):1–24.

[69] Fernandes, P. R., Rodrigues, H. C., Guedes, J. M., and Coelho, P. G. (2012). Multiscale modelling on bone mechanics – application to tissue engineering and bone quality analysis. *IFAC Proceedings Volumes*, 45(2):1013–1017.

[70] Fischer, J. (2019). The choice of representative volumes in the approximation of effective properties of random materials. *Archive for Rational Mechanics and Analysis*, 234(2):635–726.

[71] Friedman, A. (1964). *Partial Differential Equations of Parabolic Type*. Englewood Cliffs, NJ.

[72] Göckler, T. and Grimm, V. (2013). Convergence analysis of an extended krylov subspace method for the approximation of operator functions in exponential integrators. *SIAM Journal on Numerical Analysis*, 51(4):2189–2213.

[73] Geers, M. G. D., Kouznetsova, V. G., Matouš, K., and Yvonnet, J. (2017). *Homogenization Methods and Multiscale Modeling: Nonlinear Problems*, pages 1–34. American Cancer Society.

[74] George, A. and Ng, E. (1988). On the complexity of sparse QR and LU factorization of finite-element matrices. *SIAM J. Sci. and Stat. Comput.*, 9(5):849–861.

[75] Gesztesy, F., Mitrea, M., and Nichols, R. (2014). Heat kernel bounds for elliptic partial differential operators in divergence form with robin-type boundary conditions. *Journal d'Analyse Mathématique*, 122(1):229–287.

[76] Gilles, M. A. and Townsend, A. (2019). Continuous analogues of krylov subspace methods for differential operators. *SIAM Journal on Numerical Analysis*, 57(2):899–924.

[77] Gloria, A. (2011a). Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(1):1–38.

[78] Gloria, A. (2011b). Reduction of the resonance error. Part 1: Approximation of homogenized coefficients. *Math. Models Methods Appl. Sci.*, 21(8):1601–1630.

[79] Gloria, A. and Habibi, Z. (2016). Reduction in the resonance error in numerical homogenization ii: Correctors and extrapolation. *Found. Comput. Math.*, 16(1):217–296.

[80] Gloria, A., Neukamm, S., and Otto, F. (2015). Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Inventiones mathematicae*, 199(2):455–515.

[81] Gloria, A. and Otto, F. (2011). An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856.

[82] Gloria, A. and Otto, F. (2012). An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28.

[83] Gloria, A. and Otto, F. (2017). Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc.*, 19:3489–3548.

[84] Gloria, Antoine (2012). Numerical homogenization: survey, new results, and perspectives. *ESAIM: Proc.*, 37:50–116.

[85] Golub, G. H. and Van Loan, C. F. (2013). *Matrix Computations*. J. Hopkins Uni. Press.

[86] Güttel, S. (2013). Rational krylov approximation of matrix functions: Numerical methods and optimal pole selection. *GAMM-Mitteilungen*, 36(1):8–31.

[87] Hannukainen, A., Mourrat, J.-C., and Stoppels (2019). Computing homogenized coefficients via a multiscale representation and hierarchical hybrid grids. available at https://arxiv.org/abs/1905.06751.

[88] Henning, P. and Målqvist, A. (2014). Localized orthogonal decomposition techniques for boundary value problems. *SIAM J. Sci. Comput.*, 36(4):A1609–A1634.

[89] Henning, P., Målqvist, A., and Peterseim, D. (2014). A localized orthogonal decomposition method for semi-linear elliptic problems. *M2AN Math. Model. Numer. Anal.*, 48(5):1331–1349.

[90] Henning, P. and Peterseim, D. (2013). Oversampling for the Multiscale Finite Element Method. *SIAM Multiscale Model. Simul.*, 11(4):1149–1175.

[91] Herrmann, L. and Schwab, C. (2018). QMC integration for lognormal-parametric, elliptic PDEs: local supports and product weights. *Numerische Mathematik*, 141(1):63–102.

[92] Higham, N. (2008). *Functions of matrices : theory and computation*. Society for Industrial and Applied Mathematics, Philadelphia, Pa.

[93] Hill, R. (1963). Elastic properties of reinforced solids: some theoretical principles. *Journal of the Mechanics and Physics of Solids*, 11(5):357–372.

[94] Hochbruck, M. and Lubich, C. (1997). On Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 34:1911–1925.

[95] Hochbruck, M. and Ostermann, A. (2010). Exponential integrators. *Acta Numerica*, 19:209–286.

[96] Hou, T. Y. and Wu, X.-H. (1997). A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189.

[97] Hou, T. Y., Wu, X.-H., and Cai, Z. (1999). Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.*, 68(227):913–943.

[98] Hou, T. Y., Wu, X.-H., and Zhang, Y. (2004). Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation. *Commun. Math. Sci.*, 2(2):185–205.

[99] Hughes, T. J., Feijóo, G. R., Mazzei, L., and Quincy, J.-B. (1998). The variational multiscale method – a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166(1):3 – 24. Advances in Stabilized Methods in Computational Mechanics.

[100] Hughes, T. J. R. (1995). Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Comput. Methods Appl. Mech. Engrg.*, 127(1-4):387–401.

[101] Jikov, V. V., Kozlov, S. M., and Oleinik, O. A. (1994). *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, Heidelberg.

[102] Kanit, T., Forest, S., Galliet, I., Mounoury, V., and Jeulin, D. (2003). Determination of the size of the representative volume element for random composites: statistical and numerical approach. *International Journal of Solids and Structures*, 40(13):3647 – 3679.

[103] Kevrekidis, I. G., Gear, C. W., Hyman, J. M., Kevrekidis, P. G., Runborg, O., and Theodoropoulos, C. (2003). Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.*, 1(4):715–762.

[104] Kozlov, S. M. (1979). The averaging of random operators. *Mat. Sb. (N.S.)*, 109(151)(2):188–202, 327.

[105] Ladyzhenskaya, O. A., Solonnikov, V. A., and Ural'tseva, N. N. (1968). *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I.

[106] Le Bris, C. and Legoll, F. (2017). Examples of computational approaches for elliptic, possibly multiscale PDEs with random inputs. *Journal of Computational Physics*, 328:455–473.

[107] Le Bris, C., Legoll, F., and Minvielle, W. (2016). Special quasirandom structures: A selection approach for stochastic homogenization. *Monte Carlo Methods and Applications*, 22(1).

[108] Legoll, F. and Minvielle, W. (2015a). A control variate approach based on a defect-type theory for variance reduction in stochastic homogenization. *Multiscale Modeling & Simulation*, 13(2):519–550.

[109] Legoll, F. and Minvielle, W. (2015b). Variance reduction using antithetic variables for a nonlinear convex stochastic homogenization problem. *Discrete & Continuous Dynamical Systems - S*, 8(1):1–27.

[110] Li, P. and Yau, S. T. a. (1983). On the Schrödinger equation and the eigenvalue problem. *Comm. Math. Phys.*, 88(3):309–318.

[111] Lions, J.-L. and Magenes, E. (1968). *Problèmes aux limites non homogènes et applications*, volume 1 of *Travaux et recherches mathématiques*. Dunod, Paris.

[112] Lundgren, T. S. (1993). A small-scale turbulence model. *Physics of Fluids A: Fluid Dynamics*, 5(6):1472–1483.

[113] Maday, Y., Nguyen, N.-C., Patera, A. T., and Pau, G. S. (2009). A general multipurpose interpolation procedure: the magic points. *Commun. Pure Appl. Anal.*, 8(1):383–404.

[114] Målqvist, A. and Peterseim, D. (2014). Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603.

[115] McCarthy, J. F. (1995). Comparison of fast algorithms for estimating large-scale permeabilities of heterogeneous media. *Transport in Porous Media*, 19:123–137. 10.1007/BF00626662.

[116] Ming, P. and Zhang, P. (2007). Analysis of the heterogeneous multiscale method for parabolic homogenization problems. *Math. Comp.*, 76(257):153–177.

[117] Moler, C. and Van Loan, C. (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, 20(4):801–836.

[118] Moler, C. and Van Loan, C. F. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45(1):3–49.

[119] Mourrat, J.-C. (2019). Efficient methods for the estimation of homogenized coefficients. *Found. Comput. Math.*, 19(2):435–483.

[120] Murat, F. and Tartar, L. (1997). H-convergence. In *Topics in the mathematical modelling of composite materials*, volume 31 of *Progr. Nonlinear Differential Equations Appl.*, pages 21–43. Birkhäuser Boston, Boston, MA.

[121] Nash, J. (1958). Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80(4):931–954.

[122] Pankov, A. (1997). *G-convergence and homogenization of nonlinear partial differential operators*, volume 422 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht.

[123] Papanicolaou, G. C. and Varadhan, S. R. S. (1979). Boundary value problems with rapidly oscillating random coefficients. *Random fields*, 1:835–873.

[124] Payne, L. E. and Weinberger, H. F. (1960). An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.*, 5:286–292.

[125] Saad, Y. (2003). *Iterative methods for sparse linear systems*, volume 82. SIAM.

[126] Safarov, Y. and Vassiliev, D. (1997). *The asymptotic distribution of eigenvalues of partial differential operators*, volume 155 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI. Translated from the Russian manuscript by the authors.

[127] Salciccioli, M., Stamatakis, M., Caratzoulas, S., and Vlachos, D. G. (2011). A review of multiscale modeling of metal-catalyzed reactions: Mechanism development for complexity and emergent behavior. *Chemical Engineering Science*, 66(19):4319–4355.

[128] Shiari, B. and Miller, R. E. (2016). Multiscale modeling of crack initiation and propagation at the nanoscale. *Journal of the Mechanics and Physics of Solids*, 88:35–49.

[129] Spagnolo, S. (1968). Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 22(4):571–597.

[130] Strouboulis, T., Babuška, I., and Copps, K. (2000). The design and analysis of the generalized finite element method. *Comput. Methods Appl. Mech. Engrg.*, 181:43–69.

[131] van der Houwen, P. and Sommeijer, B. P. (1980). On the internal stage Runge-Kutta methods for large m-values. *Z. Angew. Math. Mech.*, 60:479–485.

[132] van Genuchten, M. T. (1980). A closed–form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44:892–898.

[133] Wu, X. H., Efendiev, Y., and Hou, T. Y. (2002). Analysis of upscaling absolute permeability. *Discrete Contin. Dyn. Syst. Ser. B*, 2(2):185–204.

[134] Yue, X. and E, W. (2007). The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size. *J. Comput. Phys.*, 222(2):556–572.

[135] Yurinskiĭ, V. V. (1986). Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27(4):167–180, 215.

[136] Zheng, X., Smith, W., Jackson, J., Moran, B., Cui, H., Chen, D., Ye, J., Fang, N., Rodriguez, N., Weisgraber, T., and Spadaccini, C. M. (2016). Multiscale metallic metamaterials. *Nature Materials*, 15(10):1100–1106.

Edoardo Paganoni

+41 (0)77 9644661
+39 328 3358190
edoardo.paga@gmail.com
edoardo-paganoni

Current position

2016–2020 **Doctoral Research Assistant**, *Swiss Federal Institute of Technology (EPFL)*, Lausanne (CH), Laboratory of Numerical Analysis and Computational Mathematics, thesis advisor: A. Abdulle.

Education

2016–2020 **PhD in Mathematics**, *Swiss Federal Institute of Technology (EPFL)*, Lausanne (CH).

2014–2015 **MSc in Industrial and Applied Mathematics**, *Eindhoven University of Technology (TU/e), Politecnico di Torino*, Eindhoven (NL), Torino (IT), "Erasmus Plus" exchange program for double degree. Final grade: 110/110 cum Laude.

2010–2013 **BSc in Mathematics for Engineering**, *Politecnico di Torino*, Torino (IT), *Final grade: 110/110 cum Laude*.

Professional experience

Nov 2015–Feb 2016 **Fidia SpA**, *Junior Software Developer*, Torino (IT).

Nov 2014–Jul 2015 **Philips Research**, *Internship in the Population Health Department*, Eindhoven (NL).

Ago 2012–Oct 2012 **EDISU (Regional Board for Access to Higher Education)**, *Internship*, Torino.

Conferences, workshops and schools

Jul 2020 **14th WCCM & ECCOMAS**, Paris (F).
Cancelled because of COVID-19

Oct 2019 **Workshop on New Trends in Asymptotic Methods for Multiscale PDEs**, *Karlstads Universitet*, Karlstad (S), Scientific presentation.

Oct 2019 **Workshop on New Trends in Asymptotic Methods for Multiscale PDEs**, *Karlstads Universitet*, Karlstad (S), Scientific presentation.

Oct 2019 **Fall School on Mathematical and Computational Aspects of Machine Learning**, *Mathematical Research Center "Ennio de Giorgi", Scuola Normale Superiore*, Pisa (IT).

May 2019 **Swiss Numerics Day**, *Universitá della Svizzera Italiana (USI)*, Lugano (CH), Poster presentation.

Feb 2019 **SIAM CSE conference**, *SIAM (Society for Industrial and Applied Mathematics)*, Spokane, WA, Scientific presentation.

Jun 2018 **ECCOMAS ECCM-ECFD conference**, *ECCOMAS (European Community on Computational Methods in Applied Sciences)*, Glasgow (UK), Scientific presentation.

Jan 2017 **Workshop on Multiscale Methods for Stochastic Problems**, *Université de Genève*, Geneva (CH).

Jan 2017 **Winter School on Numerical Analysis of Multiscale Problems**, *Universität Bonn*, Bonn (DE).

Scholarships & Awards

2019 Tutoring award - EPFL

2014 ALSP scholarship - TU/e

2012 "Alfaclass Update" scholarship - Fondazione CRT

2010 INdAM scholarship - Istituto Nazionale di Alta Matematica "Francesco Severi"

2010 Entrance tests and exams results scholarship - Politecnico di Torino

Publications

[1] A. Abdulle, D. Arjmand, and E. Paganoni. Analytical and numerical study of a modified cell problem for the numerical homogenization of multiscale random fields. *Submitted, preprint available at https://arxiv.org/abs/2007.10828*, 2020.

[2] Assyr Abdulle, Doghonay Arjmand, and Edoardo Paganoni. Exponential decay of the resonance error in numerical homogenization via parabolic and elliptic cell problems. *C. R. Math. Acad. Sci. Paris*, 357:545–551, 2019.

[3] Assyr Abdulle, Doghonay Arjmand, and Edoardo Paganoni. An elliptic local problem with exponential decay of the resonance error for numerical homogenization. *Submitted, preprint available at https://arxiv.org/abs/2001.06315*, 2020.

[4] Assyr Abdulle, Doghonay Arjmand, and Edoardo Paganoni. A parabolic local problem with exponential decay of the resonance error for numerical homogenization. *Submitted, preprint available at https://arxiv.org/abs/2001.05543*, 2020.

Technical skills

Scientific computing	MATLAB®, FreeFEM++, R;
Programming Languages	C/C++, Fortran, Python;
Simulation software	ANSYS-FLUENT®, COMSOL®;

Languages

Italian	Mother tongue
English	Excellent
French	Fluent

Interests

Street and B/W photography, mid-long distance running, mountain hiking.

Referees

Prof. A. Abdulle, EPFL, assyr.abdulle@epfl.ch

Dr. D. Arjmand, Mälardalen University, doghonay.arjmand@mdh.se

Dr. B. Bakker, Philips Research, bart.bakker@philips.com