Freie Universität Berlin

Adaptive Multilevel Monte Carlo Methods for Random Elliptic Problems

Vorgelegt von
Evgenia Youett

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

Berlin, 2018

Betreuer: **Prof. Dr. Ralf Kornhuber**

Erstgutachter: **Prof. Dr. Ralf Kornhuber**

Zweitgutachter: **Prof. Dr. Robert Scheichl**

Tag der Disputation: 26.11.2018

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe. Ich erkläre weiterhin, dass ich die vorliegende Arbeit oder deren Inhalt nicht in einem früheren Promotionsverfahren eingereicht habe.

Evgenia Youett

Berlin, den 28.06.2018

Acknowledgements

I wish to thank many people who have helped me and have supported me in many different ways in completing this thesis. My deepest appreciation goes to my supervisor Ralf Kornhuber, not only for his valuable scientific advice and patient guidance but also for his optimism, encouragement and great moral support. I would also like to express my gratitude to former and current members of the AG Kornhuber for the wonderful environment in which I had the pleasure to work. Special thanks go to Carsten Gräser for his willingness to help and his valuable pieces of advise during my work. I thank Ana Djurdjevac for her help with probability theory, with hundreds of little mathematical questions I had over the last years, for the feedback on some parts of this thesis and for her support. I thank Max Kahnt for his help with DUNE and his constant support. I thank Ansgar Burchardt for sharing the code for the wear tests simulations. I thank Han Cheng Lie for his feedback and many valuable comments on this thesis. I thank the DAAD, MPIMG, BMBF and CRC 1114 for the funding I received from them during my PhD studies. I thank the IMPRS and BMS for their support and for the funding they provided for attending scientific conferences. I would especially like to express my gratitude to Kirsten Kelleher for the great support with all formalities at FU Berlin. I thank Max Gunzburger for inviting me to the conference on UQ in Trieste and to the INI programme on UQ in Cambridge. I am grateful to Robert Scheichl for encouraging me to investigate more theoretical issues than I planned, that in my opinion improved my thesis. I am grateful to Christof Schütte and Martin Weiser for giving me a chance to start working at FU Berlin. I very much appreciate the support that I received from Tim Conrad in the beginning of my PhD studies. I would also like to thank the late Elias Pipping for sharing his PhD experience and making me believe that everything is possible. I thank former and current members of the Biocomputing and the Computational Molecular Biology groups of FU Berlin for helping me to keep a good balance between work and social life, completion of this thesis would have been much harder without many of them. I thank my former advisor Nikolay Bessonov for all knowledge and experience he has shared with me and for opening the world of science to me. I would like to thank my family for their support and doubtlessness in my success. I thank my friends for their support. Last but not least, I would like to thank my husband Jonathan Youett for supporting me in all possible aspects, for his help with numerous mathematical issues and DUNE, for believing in me and making my life wonderful.

Contents

Chapter 1

Introduction

Mathematical modelling and computational simulations are important tools in many scientific and practical engineering applications. However, simulation results may differ from real observations, due to various uncertainties that may arise from model inputs representing physical parameters and material properties, external loads, initial or boundary conditions. Therefore, the problem of identification, quantification and interpretation of different uncertainties arising from mathematical models and raw data is becoming increasingly important in computational and engineering problems.

One example of a problem involving uncertainties that is often considered in the literature (e.g. [29, 85, 86]) is the leakage of radioactive waste from a geological repository to the groundwater. The movement of groundwater through soil and rocks can be described by Darcy's model for the flow of fluids through a porous media

$$\mathbf{v} = -\frac{\mathbf{A}}{\eta}\big(\nabla u - \rho \mathbf{g}\big) \quad \text{in } D, \tag{1.1}$$

where $\mathbf{v}$ denotes the filtration velocity, $\mathbf{A}$ is the permeability tensor, η is the fluid viscosity, u is the pressure, ρ is the fluid density, $\mathbf{g}$ is the field of external forces and $D \subset \mathbb{R}^d$, $d = 1, 2, 3$. The complete system of equations for incompressible fluids also includes the incompressibility condition

$$\nabla \cdot \mathbf{v} = 0 \quad \text{in } D \tag{1.2}$$

and appropriate boundary conditions. Equations (1.1) and (1.2) lead to the problem

$$-\nabla \cdot (\mathbf{A}\nabla u) = f \quad \text{in } D, \tag{1.3}$$

with $f := -\rho\nabla \cdot (\mathbf{A}\mathbf{g})$, complemented with appropriate boundary conditions as before.

The permeability tensor $\mathbf{A}$ plays the role of a source of uncertainties in this problem. Usually, the measurements for $\mathbf{A}$ are only available at a few spatial points and are often inaccurate, which makes approximation of the tensor field quite uncertain. One of the attempts to address this uncertainty problem is to model $\mathbf{A}$ as a random field over a probability space, and to estimate averaged properties of the random field, such as the mean and correlation function, from measurements. In this work we only consider scalar-valued permeability coefficients, i.e. $\mathbf{A}$ can be represented as $\mathbf{A} = \alpha\mathbf{I}$, where α is a scalar-valued random field and $\mathbf{I}$ is the unit tensor. In practically relevant problems it is often the case that the correlation length of the field α is significantly smaller than the size of the considered domain D, which leads to a large number of random variables needed for an accurate parametrization of the random field.

A quantity of interest for problems with random input data might be an approximation to the expected solution, or to the expectation or higher moments of an output of interest defined by a functional of the random solution.

Various approaches have been developed in order to tackle the problem of uncertainty quantification in different mathematical models that include random input data, and in particular problems of the form (1.3). An overview of existing methods can be found in e.g. [67]. Polynomial chaos (e.g. [94, 95]) is a spectral method utilizing orthogonal decompositions of the input data. This method exhibits a polynomial dependence of its computational cost on the dimension of the probability space where the uncertain input data is defined. Stochastic Galerkin methods (e.g. [10, 70]) are an intrusive approach that utilizes ideas of the well known Galerkin methods, originally developed for deterministic equations, for approximating solutions in the probability space as well. Stochastic Galerkin methods appear to be a powerful tool, however, they place certain restrictions on the random input data, e.g. linear dependence of random fields on a number of independent random variables with bounded support. Moreover, the computational cost of the methods depends exponentially on the number of random variables parametrizing the random fields. Stochastic collocation methods (e.g. [9]) are another approach that assumes dependence of the random inputs on a number of random variables, however, it allows this dependence to be nonlinear. Further, the random variables may be correlated and have unbounded support. Stochastic collocation methods, as well as polynomial chaos and stochastic Galerkin methods, suffer from the so-called "curse of dimensionality", i.e. their performance depends on the dimension of the probability space where the uncertain input data is defined. This restricts the application of all three types of methods to problems where the input data can be parametrized by few random variables.

A class of methods that are insensitive to the dimensionality of the probability spaces are Monte Carlo (MC) methods. While the classical MC method is very robust and extremely simple, its convergence is rather slow compared to the methods mentioned above. Sampling of stochastic data entails numerical solution of numerous deterministic problems which makes performance the main weakness of this method. Several approaches have been developed in order to improve efficiency of the standard MC method. Utilizing deterministically chosen integration points in MC simulations leads to quasi-Monte Carlo methods (e.g. [42, 63]) that exhibit an improved order of convergence. A variance reduction technique known as multilevel Monte Carlo (MLMC) method, first introduced by Heinrich for approximating high-dimensional integrals and solving integral equations in [53] and extended further by Giles to integration related to stochastic differential equations (e.g. [40, 41]), combines the MC method with multigrid ideas by introducing a suitable hierarchy of subproblems associated with a corresponding mesh hierarchy. MLMC methods allow a considerable improvement of efficiency compared to the classical MC method and have become a powerful tool in a variety of applications. We refer to the works on MLMC methods applied to elliptic problems with random coefficients [14, 28, 29, 87], random elliptic problems with multiple scales [2], random parabolic problems [13] and random elliptic variational inequalities [16, 60].

The efficiency of MLMC applied to random partial differential equations (PDEs) was further enhanced by a number of scientific groups working on different aspects of the approach. Optimization of MLMC parameters and selection of meshes from a given uniform mesh hierarchy reducing the computational cost of the MLMC method were performed by Collier et al. [30] and Haji-Ali et al. [51]. The advantages of MLMC and quasi-Monte Carlo methods were combined by Kuo et al. [64, 65]. Haji-Ali et al. [50] proposed a multi-index Monte Carlo method that can be viewed as a generalization of MLMC, working with higher order mixed differences instead of first order differences as in MLMC and reducing the variance of the hierarchical differences. In some cases the multi-index Monte Carlo method allows an improvement of the computational cost compared to the classical MLMC method, see [50]. Further, the ideas of multi-index Monte Carlo were combined with the ideas of quasi-Monte Carlo methods by Robbe et al. in [77]. Finally, multilevel ideas were combined with stochastic collocation methods by Teckentrup et al. in [88].

Another approach to reduce the computational cost of MLMC methods is to use adaptive mesh refinement techniques. Time discretization of an Itô stochastic differential equation by an adaptively chosen hierarchy of time steps has been suggested by Hoel et al. [54, 55] and a similar approach was presented by Gerstner and Heinz [38], including applications in computational finance. Several works were recently done in this direction for MLMC methods with application to random partial differential equations. Eigel et al. [35] suggested an algorithm for constructing an adaptively refined hierarchy of meshes based on expectations of pathwise local error indicators and illustrated its properties by numerical experiments. Elfverson et al. [36] suggested a sample-adaptive MLMC method for approximating failure probability functionals and Detommaso et al. [33] introduced continuous level Monte Carlo treating the level as a continuous variable as a general framework for sample-adaptive level hierarchies. Adaptive refinement techniques were also combined with a multilevel collocation method by Lang et al. in [66].

In this thesis we introduce a novel framework that utilizes the multilevel ideas and allows for exploiting the advantages of adaptive finite element techniques. In contrast to the standard MLMC method, where levels are characterized by a hierarchy of uniform meshes, we associate the MLMC levels with a chosen sequence of tolerances Tol_l, $l \in \mathbb{N} \cup \{0\}$. Each deterministic problem corresponding to a MC sample on level l is then approximated up to accuracy Tol_l. This can be done, for example, using pathwise a posteriori error estimation and adaptive mesh refinement techniques. We emphasize the pathwise manner of mesh refinement in this case, which results in different adaptively constructed hierarchies of meshes for different samples. We further introduce an adaptive MLMC finite element method for random linear elliptic problems based on a residual-based a posteriori error estimation technique. We provide a careful analysis of the novel method based on a generalization of existing results, for deterministic residual-based error estimation, to the random setting. We complement our theoretical results by numerical simulations illustrating the advantages of our approach compared to the standard MLMC finite element method when applied to problems with random singularities. Parts of the material in this thesis have been published in [61].

This thesis is organized as follows. Chapter 2 contains the formulation of random elliptic variational equalities and inequalities, the two model problems that we consider throughout the thesis. We postulate a set of assumptions for these problems that guarantee the well-posedness of these problems and the regularity of solutions. We further specify the quantities of interest derived from the random solutions to the introduced problems and give some examples of random fields appearing in the problem formulations.

In Chapter 3 we present an abstract framework for adaptive MLMC methods together with error estimates formulated in terms of the desired accuracy and upper bounds for the expected computational cost of the methods.

The general theory presented in Chapter 3 is applied to MLMC finite element methods for random linear elliptic problems in Chapter 4. In the case of uniform refinement we recover the existing classical MLMC convergence results (e.g. [28, 87]) together with the estimates for the expected computational cost of the method. Then we introduce an adaptive MLMC finite element method and formulate assumptions for an adaptive finite element algorithm, which provide MLMC convergence and desired bounds for the computational cost. Further, we introduce an adaptive algorithm based on residual a posteriori error estimation and demonstrate that it fulfills the assumptions stated for the adaptive MLMC finite element method.

In Chapter 5 we compare the performance of the classical uniform and the novel adaptive MLMC finite element methods by presenting results of numerical simulations for two random linear elliptic problems. Our numerical results show that for problems with highly localized random input data, the adaptive MLMC method achieves a significant reduction of computational cost compared to the uniform one.

Finally, in Chapter 6 we present a practically relevant problem of uncertainty quantification in wear tests of knee implants. Since the mathematical problem that describes the wear tests is formulated on a 3-dimensional domain with a complex geometry, construction of an uniform mesh hierarchy required for the standard MLMC finite element method appears to be problematic. The coarsest mesh in the hierarchy should resolve the geometry of the domain, whereas the finest mesh should be computationally feasible in terms of size. The framework of adaptive MLMC methods is a natural choice in this situation. We apply an adaptive MLMC finite element method utilizing hierarchical error estimation techniques to the pathwise problems and estimate the expected mass loss of the implants appearing in the wear tests.

Chapter 2

Random Variational Problems

In this chapter we introduce two problems that are considered throughout the thesis. The random linear elliptic problem described in Section 2.2 is the primary problem for the theoretical analysis presented in Chapter 4. The random variational inequality introduced in Section 2.3 is important for the practical application presented in Chapter 6. Some comments on the existing analytical results for this type of problem are stated in Chapter 4.

2.1 Notation

In what follows let $D \subset \mathbb{R}^d$ be an open, bounded, convex, polyhedral domain. The space dimension d can take values $1, 2, 3$.

Let $L^p(D)$ denote the space of Lebesgue-measurable, p-integrable, real-valued functions on D with the norm defined as

$$\|v\|_{L^p(D)} := \begin{cases} \left(\int_D |v(x)|^p \, dx\right)^{\frac{1}{p}}, & \text{if} \quad 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{x \in D} |v(x)|, & \text{if} \quad p = \infty. \end{cases}$$

The Sobolev space $H^k(D) := \{v \in L^2(D) : D^\alpha v \in L^2(D) \text{ for } |\alpha| \leq k\} \subset L^2(D)$, $k \in \mathbb{N} \cup \{0\}$ consists of functions with square integrable weak derivatives of order $|\alpha| \leq k$, where α denotes the d-dimensional multi-index, and is a Hilbert space. The inner product on $H^k(D)$ is defined as

$$(v, w)_{H^k(D)} := \int_D \sum_{|\alpha| \leq k} D^\alpha v(x) \cdot D^\alpha w(x) \, dx,$$

and induces the norm

$$\|v\|_{H^k(D)} := (v, v)^{\frac{1}{2}}_{H^k(D)}.$$

Note that $L^2(D) = H^0(D)$.

With $C_0^\infty(D)$ denoting the space of infinitely smooth functions in D with compact support, we denote its closure under the norm $\| \cdot \|_{H^1(D)}$ by $H^1_0(D)$.

We will also make use of the Sobolev spaces $W^{k,\infty}(D)$ defined by $W^{k,\infty}(D) := \{v \in L^\infty(D) : D^\alpha v \in L^\infty(D) \text{ for } |\alpha| \leq k\},\ k \in \mathbb{N}$ with the norm

$$\|v\|_{W^{k,\infty}(D)} := \max_{0\leq|\alpha|\leq k} \operatorname*{ess\,sup}_{x\in D} |D^\alpha v(x)|.$$

Let $C^k(\overline{D})$, $k \in \mathbb{N} \cup \{0\}$ denote the space of continuous functions which are k times continuously differentiable with the norm

$$\|v\|_{C^k(\overline{D})} := \sum_{0\leq|\alpha|\leq k} \sup_{x\in\overline{D}} |D^\alpha v(x)|.$$

For a real $0 < r \leq 1$ and $k \in \mathbb{N} \cup \{0\}$ we introduce the Hölder space $C^{k,r}(\overline{D})$ equipped with the norm

$$\|v\|_{C^{k,r}(\overline{D})} := \|v\|_{C^k(\overline{D})} + \max_{|\alpha|=k} \sup_{\substack{x,y\in\overline{D}\\ x\neq y}} \frac{|D^\alpha v(x) - D^\alpha v(y)|}{|x-y|^r}.$$

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, where Ω denotes a non-empty set of events, $\mathcal{A} \subset 2^\Omega$ is a σ-algebra on Ω and $\mathbb{P} : \mathcal{A} \to [0,1]$ is a probability measure. The space $(\Omega, \mathcal{A}, \mathbb{P})$ is called *separable* if there exists a countable family $(A_n)_{n=1}^\infty$ of subsets of $\mathcal{A}$ such that the σ-algebra generated by the family $(A_n)_{n=1}^\infty$ coincides with $\mathcal{A}$.

The expected value of a measurable random variable $\xi : \Omega \to \mathbb{R}$ is defined as

$$\mathbb{E}[\xi] := \int_\Omega \xi(\omega) d\mathbb{P}(\omega),$$

and its variance is defined as

$$\mathbb{V}[\xi] := \mathbb{E}[(\xi - \mathbb{E}[\xi])^2].$$

Let B be a Banach space of real-valued functions on the domain D with norm $\|\cdot\|_B$. We endow the space B with the Borel sigma algebra to render it a measurable space.

A random field $f : \Omega \to B$ is called *simple* if it is of the form $f = \sum_{n=1}^N \psi_{A_n} v_n$, $N \in \mathbb{N}$, where $A_n \in \mathcal{A}$ and $v_n \in B$ for all $n = 1, \ldots, N$. Here ψ_A denotes the indicator function of the set A, i.e. $\psi_A(\omega) = 1$, if $\omega \in A$, and $\psi_A(\omega) = 0$, if $\omega \notin A$. A random field f is called *strongly measurable* if there exists a sequence of simple functions f_n, such that $\lim_{n\to\infty} f_n = f$ pointwise in Ω.

For the given Banach space B, we introduce the Bochner-type space $L^p(\Omega, \mathcal{A}, \mathbb{P}; B)$ of strongly measurable, p-integrable mappings $f : \Omega \to B$ with the norm

$$\|f\|_{L^p(\Omega,\mathcal{A},\mathbb{P};B)} := \begin{cases} \left(\int_\Omega \|f(\cdot,\omega)\|_B^p \, d\mathbb{P}(\omega) \right)^{1/p}, & \text{if} \quad 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{\omega\in\Omega} \|f(\cdot,\omega)\|_B, & \text{if} \quad p = \infty. \end{cases}$$

In order to keep the notation short we use the abbreviation $L^p(\Omega; B) :=$ $L^p(\Omega, \mathcal{A}, \mathbb{P}; B)$ and $L^p(\Omega) := L^p(\Omega; \mathbb{R})$. We use the convention $\frac{1}{\infty} = 0$ when talking about the orders of L^p-spaces.

We will often restrict ourselves to a Hilbert space H and the Bochner space $L^2(\Omega; H)$. It is easy to see that $L^2(\Omega; H)$ is also a Hilbert space with the scalar product

$$(v, w)_{L^2(\Omega;H)} := \int_\Omega (v(\cdot, \omega), w(\cdot, \omega))_H \, d\mathbb{P}(\omega), \qquad v, w \in L^2(\Omega; H).$$

The expected value of an H-valued random variable f is defined as

$$\mathbb{E}[f] := \int_\Omega f(\cdot, \omega) \, d\mathbb{P}(\omega) \in H.$$

For positive quantities a and b we write $a \lesssim b$ if the ratio a/b is uniformly bounded by a constant independent of $\omega \in \Omega$. We write $a \simeq b$ if $a \lesssim b$ and $b \lesssim a$.

2.2 Random elliptic variational equalities

For a given $\omega \in \Omega$ we consider the following random elliptic partial differential equation subject to homogeneous Dirichlet boundary conditions

$$\begin{aligned} -\nabla \cdot (\alpha(x, \omega) \nabla u(x, \omega)) &= f(x, \omega) \quad \text{for } x \in D, \\ u(x, \omega) &= 0 \qquad\quad\;\; \text{for } x \in \partial D. \end{aligned} \tag{2.1}$$

Here α and f are a random coefficient and a source function respectively. We restrict ourselves to homogeneous Dirichlet boundary conditions for ease of presentation.

We let real-valued random variables $\alpha_{\min}$ and $\alpha_{\max}$ be such that

$$\alpha_{\min}(\omega) \le \alpha(x, \omega) \le \alpha_{\max}(\omega) \quad \text{a.e. in } D \times \Omega. \tag{2.2}$$

If realizations of the random coefficient α are continuous, the random variables $\alpha_{\min}$ and $\alpha_{\max}$ can be defined as

$$\alpha_{\min}(\omega) := \min_{x \in \overline{D}} \alpha(x, \omega), \quad \alpha_{\max}(\omega) := \max_{x \in \overline{D}} \alpha(x, \omega), \tag{2.3}$$

and are finite a.e. in Ω.

We impose the following assumptions on the random coefficient α and on the random right hand side f. This set of assumptions is rather usual in the context of random elliptic PDEs, similar ones were made e.g. in [28, 87].

Assumption 2.2.1. (i) $\alpha_{\min} > 0$ a.s. and $\frac{1}{\alpha_{\min}} \in L^p(\Omega)$ for all $p \in [1, \infty)$,

(ii) $\alpha \in L^p(\Omega; W^{1,\infty}(D))$ for all $p \in [1, \infty)$,

(iii) $f \in L^{p_f}(\Omega; L^2(D))$ for some $p_f \in (2, \infty]$.

Remark 2.2.1. We notice that in order for Assumption (2.2.1) (ii) to make sense the function $\alpha : \Omega \to W^{1,\infty}(D)$ must be Bochner-integrable. Therefore, it includes the assumption that the mapping $\Omega \ni \omega \mapsto \alpha(\cdot, \omega) \in W^{1,\infty}(D)$ is strongly measurable. Note that strong measurability of α implies its measurability, see [84, Lemma 9.12]. As for Assumption (2.2.1) (iii), it includes the assumption that the mapping $\Omega \ni \omega \mapsto f(\cdot, \omega) \in L^2(D)$ is measurable, which implies strong measurability, since $L^2(D)$ is separable [84, Theorem 9.3].

Let us notice that Assumption 2.2.1 (ii) implies that $\alpha_{\max} \in L^p(\Omega)$ for all $p \in [1, \infty)$. It also implies that realisations of the random coefficient α are in $W^{1,\infty}(D)$ a.s., which together with convexity of the spatial domain D yields that they are a.s. Lipschitz continuous [37, Theorem 4.5], [52, Theorem 4.1]. Assumption 2.2.1 (iii) implies that realisations of the random right hand side f are a.s. in $L^2(D)$.

We will work with the weak formulation of problem (2.1). For a fixed $\omega \in \Omega$, it takes the form of the random variational equality

$$u(\cdot, \omega) \in H_0^1(D): \qquad a(\omega; u(\cdot, \omega), v) = \ell(\omega; v), \qquad \forall v \in H_0^1(D), \tag{2.4}$$

where the bilinear and linear forms are given by

$$a(\omega; u, v) := \int_D \alpha(x, \omega) \nabla u(x) \cdot \nabla v(x)\, dx, \quad \ell(\omega; v) := \int_D f(x, \omega) v(x)\, dx. \tag{2.5}$$

Assumption 2.2.1 (i-ii), the Poincaré and Cauchy-Schwarz inequalities ensure that for all $v, w \in H_0^1(D)$ and almost all $\omega \in \Omega$ there holds

$$c_a(\omega) \|v\|_{H^1(D)}^2 \le a(\omega; v, v), \quad a(\omega; v, w) \le \alpha_{\max}(\omega) \|v\|_{H^1(D)} \|w\|_{H^1(D)},$$

where $0 < c_a(\omega)$ and $\alpha_{\max}(\omega) < \infty$ a.s., i.e. the bilinear form a is a.s. pathwise coercive and continuous with coercivity constant $c_a(\omega) = \alpha_{\min}(\omega) c_D$, where c_D depends only on D, and continuity constant $\alpha_{\max}(\omega)$.

Proposition 2.2.1. *Let Assumption 2.2.1 hold, then the pathwise problem* (2.4) *admits a unique solution $u(\cdot, \omega)$ for almost all $\omega \in \Omega$ and $u \in L^p(\Omega; H_0^1(D))$ for all $p \in [1, p_f)$.*

Proof. Assumption 2.2.1 in combination with the Lax-Milgram theorem [49, Theorem 7.2.8] yields existence and uniqueness of the solution to (2.4) for almost all $\omega \in \Omega$ and it holds

$$\|u(\cdot, \omega)\|_{H^1(D)} \le \frac{1}{c_a(\omega)} \|f(\cdot, \omega)\|_{L^2(D)}. \tag{2.6}$$

Measurability of the mapping $\Omega \ni \omega \mapsto u(\cdot, \omega) \in H_0^1(D)$ is provided by [73, Theorem 1]. Assumption 2.2.1 (i, iii) together with (2.6) and Hölder's inequality imply $\|u\|_{L^p(\Omega; H_0^1(D))} < \infty$ for all $p \in [1, p_f)$, see cf. [28]. □

Regularity results similar to the following can be found e.g. in [42, 65].

Proposition 2.2.2. *Let Assumption 2.2.1 hold and $u(\cdot,\omega)$ denote the solution to* (2.4) *for some $\omega \in \Omega$. Then $u(\cdot,\omega) \in H^2(D)$ almost surely and there holds*

$$\|u(\cdot,\omega)\|_{H^2(D)} \lesssim \left(\frac{1}{\alpha_{\min}(\omega)} + \frac{\|\alpha(\cdot,\omega)\|_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right) \|f(\cdot,\omega)\|_{L^2(D)},$$

where the hidden constant depends only on D.

Proof. Since $\alpha(\cdot,\omega) \in W^{1,\infty}(D)$ a.s. and the domain D is convex, the solution $u(\cdot,\omega)$ belongs to $H^2(D)$ for almost all $\omega \in \Omega$ according to standard well known results (see [20, 49]). Moreover, we have

$$\|\Delta u(\cdot,\omega)\|_{L^2} \leq \frac{1}{\alpha_{\min}(\omega)} \Big(\|f(\cdot,\omega)\|_{L^2(D)} + \|\nabla\alpha(\cdot,\omega)\|_{L^\infty(D)} \|\nabla u(\cdot,\omega)\|_{L^2(D)} \Big).$$

Equivalence of the norms $\|v\|_{H^2(D)}$ and $\|\Delta v\|_{L^2(D)}$ for all $v \in H^2(D) \cap H^1_0(D)$ and (2.6) imply the claim. □

Remark 2.2.2. In the case when there exist real values $\alpha_{\min}$ and $\alpha_{\max}$ such that

$$0 < \alpha_{\min} \leq \alpha(x,\omega) \leq \alpha_{\max} < \infty \quad \text{a.e. in } D \times \Omega,$$

Assumption 2.2.1 (i) also holds for $p = \infty$. The arguments of Proposition 2.2.1 lead in this case to $u \in L^p(\Omega; H^1_0(D))$ for all $p \in [1, p_f]$ and Proposition 2.2.2 holds as well. We complement the set of Assumptions in this case by letting $p = \infty$ in Assumption 2.2.1 (ii).

2.3 Random elliptic variational inequalities

For a given $\omega \in \Omega$ let us consider the random obstacle problem subject to homogeneous Dirichlet boundary conditions

$$\begin{aligned} -\nabla \cdot (\alpha(x,\omega)\nabla u(x,\omega)) &\geq f(x,\omega) && \text{for } x \in D, \\ u(x,\omega) &\geq \chi(x,\omega) && \text{for } x \in D, \\ (\nabla \cdot (\alpha(x,\omega)\nabla u(x,\omega)) + f(x,\omega))\,(u(x,\omega) - \chi(x,\omega)) &= 0 && \text{for } x \in D, \\ u(x,\omega) &= 0 && \text{for } x \in \partial D. \end{aligned} \tag{2.7}$$

Here α is a random coefficient, f is a random source function and χ is a random obstacle function.

In addition to Assumption 2.2.1 we make the following assumption on the obstacle function.

Assumption 2.3.1. (i) $\chi(x,\omega) \leq 0 \quad$ a.e. in $D \times \Omega$,

(ii) $\chi \in L^{p_f}(\Omega; H^2(D))$, where p_f is as in Assumption 2.2.1 (iii).

For a fixed $\omega \in \Omega$ the weak form of problem (2.7) takes the form

$$u(\cdot,\omega) \in K(\omega): \qquad a(\omega; u(\cdot,\omega), v - u(\cdot,\omega)) \geq \ell(\omega; v - u(\cdot,\omega)), \qquad \forall v \in K(\omega), \tag{2.8}$$

where

$$K(\omega) := \{v \in H_0^1(D) : \; v \geq \chi(\cdot,\omega) \quad \text{a. e. in } D\},$$

and the bilinear and linear forms are defined as in (2.5).

Note that problem (2.8) can be reformulated such that $\chi = 0$ by introducing the new variable $w = u - \chi$.

Similar to the linear case, well-posedness of the random variational inequality (2.8) can be shown under Assumptions 2.2.1 and 2.3.1, see [60, Theorem 3.3, Remark 3.5].

Proposition 2.3.1. *Let Assumptions 2.2.1 and 2.3.1 hold. Then the pathwise problem* (2.8) *admits a unique solution $u(\cdot,\omega)$ for almost all $\omega \in \Omega$ and $u \in L^p(\Omega; H_0^1(D))$ for all $p \in [1, p_f)$.*

Further, we have H^2-regularity of the solutions to (2.8), see [60, Theorem 3.7, Remark 3.9] for the proof.

Proposition 2.3.2. *Let Assumptions 2.2.1 and 2.3.1 hold and $u(\cdot,\omega)$ denote the solution to* (2.8) *for some $\omega \in \Omega$. Then $u(\cdot,\omega) \in H^2(D)$ almost surely and there holds*

$$\|u(\cdot,\omega)\|_{H^2(D)} \lesssim \left(\frac{1}{\alpha_{\min}(\omega)} + \frac{\|\alpha(\cdot,\omega)\|_{W^{1,\infty}(D)}}{\alpha_{\min}^2(\omega)} \right) \|f(\cdot,\omega)\|_{L^2(D)},$$

where the hidden constant depends only on D.

Remark 2.3.1. If Assumption 2.2.1 is modified according to Remark 2.2.2, it is possible to show that $u \in L^p(\Omega; H_0^1(D))$ for all $p \in [1, p_f]$ and Proposition 2.3.2 holds.

Problem (2.4) can be viewed as a special case of problem (2.8) with $\chi = -\infty$ and $K = H_0^1(D)$, but we keep them separate for easier referencing.

2.4 Quantities of interest

We consider two types of quantities of interest for both problems (2.4) and (2.8). First we seek the expectation of the solution

$$\mathbb{E}[u] = \int_\Omega u(\cdot,\omega) \, d\mathbb{P}(\omega). \tag{2.9}$$

Secondly, we might be interested in some property of the solution derived by a measurable functional $Q(\omega; u) : \Omega \times H_0^1(D) \to \mathbb{R}$ that defines an output of interest. We assume that for almost all $\omega \in \Omega$ the functional $Q(\omega; \cdot)$ is globally Lipschitz continuous.

Assumption 2.4.1. (i) There exists a real-valued positive random variable C_Q, such that for almost all $\omega \in \Omega$ there holds

$$|Q(\omega; v) - Q(\omega; w)| \leq C_Q(\omega) \|v - w\|_{H^1(D)}, \quad \forall v, w \in H_0^1(D).$$

(ii) $C_Q \in L^{p_Q}(\Omega)$ for some $p_Q \in (2, \infty]$.

We then seek the expected value

$$\mathbb{E}[Q(\cdot; u)] = \int_\Omega Q(\omega; u(\cdot, \omega)) \, d\mathbb{P}(\omega), \tag{2.10}$$

where $u(\cdot, \omega)$ denotes the solutions to (2.4) or to (2.8) for $\omega \in \Omega$.

2.5 Log-normal random fields

Before we introduce log-normal random fields we give a brief introduction to uniformly bounded random fields.

Uniform random fields

Let $\alpha \in L^2(\Omega; L^2(D))$ be an uniformly bounded coefficient with the mean $\mathbb{E}[\alpha]$, i.e. there exist real values $\alpha_{\min}, \alpha_{\max}$ such that

$$0 < \alpha_{\min} \leq \alpha(x, \omega) \leq \alpha_{\max} < \infty \quad \text{a.e. in } D \times \Omega,$$

Then the field α fulfills the stronger version of Assumption 2.2.1 (i), described in Remark 2.2.2.

We assume that the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is separable, which implies separability of $L^2(\Omega)$ [23, Theorem 4.13] and $L^2(\Omega; L^2(D)) \cong L^2(\Omega) \otimes L^2(D)$ [85, Section 3.5]. We then assume that for each $\omega \in \Omega$ the random field $\alpha(\cdot, \omega)$ can be represented as the series

$$\alpha(\cdot, \omega) = \mathbb{E}[\alpha] + \sum_{m=1}^{\infty} \xi_m(\omega) \varphi_m, \tag{2.11}$$

where $(\varphi_m)_{m \in \mathbb{N}}$ is an orthogonal functions system in $L^2(D)$ and $(\xi_m)_{m \in \mathbb{N}}$ are mutually independent random variables, distributed uniformly on the interval $[-\frac{1}{2}, \frac{1}{2}]$.

Let us assume that $\mathbb{E}[\alpha] \in W^{1,\infty}(D)$ and that $(\varphi_m)_{m \in \mathbb{N}} \subset W^{1,\infty}(D)$. If we have

$$\sum_{m=1}^{\infty} \|\varphi_m\|_{W^{1,\infty}(D)} < \infty,$$

then the expansion defined in (2.11) converges uniformly in $W^{1,\infty}(D)$ for all $\omega \in \Omega$ (see [80, Section 2.3]). If we in addition assume that $\|\nabla \alpha\|_{L^\infty(D)} \in L^\infty(\Omega)$, then the uniform random field α also satisfies the stronger version of Assumption 2.2.1 (ii), described in Remark 2.2.2.

Log-normal random fields

The so-called *log-normal* fields are of the form $\alpha(x,\omega) = \exp(g(x,\omega))$ for some Gaussian field $g : \overline{D} \times \Omega \to \mathbb{R}$. Note that $g \in L^2(\Omega; L^2(D))$. We again assume that $(\Omega, \mathcal{A}, \mathbb{P})$ is separable, which implies $L^2(\Omega; L^2(D)) \cong L^2(\Omega) \otimes L^2(D)$. Without loss of generality we restrict ourselves to Gaussian random fields with zero mean, i.e. $\mathbb{E}[g] = 0$. We also assume that g fulfills

$$\mathbb{E}[\big(g(x,\cdot) - g(x',\cdot)\big)^2] \le C_g |x - x'|^{2\beta_g}, \quad \forall x, x' \in \overline{D}, \tag{2.12}$$

with some positive constants β_g and C_g.

The two-point covariance function of the Gaussian field g is defined as

$$r_g(x,x') := \mathbb{E}[g(x,\cdot)g(x',\cdot)], \quad x, x' \in \overline{D}.$$

It can be shown (cf. [67, Section 2.1.1]) that under the above mentioned assumptions on the random field g the function r_g is continuous on $\overline{D} \times \overline{D}$ and

$$\int_D \int_D r_g^2(x,x')\, dx\, dx' < \infty.$$

We further restrict ourselves to random fields with isotropic covariance functions, i.e.

$$r_g(x,x') = k_g(|x - x'|), \quad x, x' \in \overline{D}, \tag{2.13}$$

for some $k_g \in C(\mathbb{R}^+)$.

The following result is known as the Kolmogorov continuity theorem and provides regularity of path realizations of the fields g and α, see [42, Proposition 1].

Proposition 2.5.1. *Assume that a Gaussian random field g fulfills* (2.12) *with constants $\beta_g \in (0,1]$ and $C_g > 0$, and its covariance function fulfills* (2.13). *Then there exists a version of g denoted by $\tilde{g}$ (i.e. $\tilde{g}(x,\cdot) = g(x,\cdot)$ a.s. for all $x \in \overline{D}$), such that $\tilde{g}(\cdot,\omega) \in C^{0,t}(\overline{D})$ for almost all $\omega \in \Omega$ and for any $0 \le t < \beta_g \le 1$. Moreover, $\tilde{\alpha}(\cdot,\omega) := \exp(\tilde{g}(\cdot,\omega)) \in C^{0,t}(\overline{D})$ for almost all $\omega \in \Omega$.*

We will identify α and $\tilde{\alpha}$ with each other. Proposition 2.5.1 provides Hölder continuity of realizations of the log-normal field α. Therefore, the definition (2.3) makes sense for this type of random field. According to [27, Proposition 2.2], $\frac{1}{\alpha_{\min}}, \alpha_{\max} \in L^p(\Omega)$ for all $p \in [1,\infty)$. Since $\alpha_{\min} > 0$ a.s. by definition of log-normal fields, it provides the validity Assumption 2.2.1 (i).

We introduce now the Karhunen-Loève (KL) expansion, which serves the purpose of parametrizing a log-normal field by a set of Gaussian random variables, and give some of its properties. We cite [4, 27, 39, 67, 80] for details.

Since the covariance function r_g is positive definite and symmetric by definition, and continuous and hence square-integrable under the above mentioned assumptions, the linear covariance operator with the kernel r_g, defined as

$$\mathcal{C} : L^2(D) \to L^2(D), \quad (\mathcal{C}v)(x) := \int_D r_g(x, x')v(x')\, dx',$$

is symmetric, positive semi-definite and compact on $L^2(D)$. According to standard results from the theory of integral operators (see e.g. [32, Chapter 3]), its eigenvalues $(\lambda_m)_{m\in\mathbb{N}}$ are real, non-negative and can be ordered in decreasing order

$$\lambda_1 \geq \lambda_2 \geq \cdots \to 0,$$

and there holds

$$\sum_{m=1}^{\infty} \lambda_m^2 < \infty.$$

The eigenfunctions $(\varphi_m)_{m\in\mathbb{N}}$ form an orthonormal basis of $L^2(D)$.

According to Mercer's theorem [4, Theorem 3.3.1], the covariance function (2.13) admits the spectral decomposition

$$r_g(x, x') = \sum_{m=1}^{\infty} \lambda_m \varphi_m(x)\varphi_m(x'), \quad x, x' \in \overline{D},$$

which converges uniformly in $\overline{D} \times \overline{D}$.

The field g can be then expanded into the KL series (see [39, Section 2.3.1] for details), i.e. for all $x \in \overline{D}$

$$g(x, \omega) = \sum_{m=1}^{\infty} \sqrt{\lambda_m}\xi_m(\omega)\varphi_m(x) \quad \text{in } L^2(\Omega),$$

and the series converges uniformly almost surely [85, Theorem 11.4]. The sequence $(\xi_m)_{m\in\mathbb{N}}$ consists of random variables defined as

$$\xi_m := \frac{1}{\sqrt{\lambda_m}}(g, \varphi_m)_{L^2(D)}, \quad m \in \mathbb{N}.$$

The variables $(\xi_m)_{m\in\mathbb{N}}$ have zero mean, unit variance and are mutually uncorrelated. Since the field g is Gaussian, the variables $(\xi_m)_{m\in\mathbb{N}}$ are also Gaussian and mutually independent.

For each $\omega \in \Omega$ the KL expansion of the realization $\alpha(\cdot, \omega)$ of a log-normal field then takes the form

$$\alpha(\cdot, \omega) = \exp\Big(\sum_{m=1}^{\infty} \sqrt{\lambda_m}\xi_m(\omega)\varphi_m\Big). \tag{2.14}$$

If we assume that $(\varphi_m)_{m\in\mathbb{N}} \subset W^{1,\infty}(D)$ and

$$\sum_{m=1}^{\infty} \sqrt{\lambda_m}\|\varphi_m\|_{W^{1,\infty}(D)} < \infty,$$

then the expansion defined in (2.14) converges in $W^{1,\infty}(D)$ a.s. (see [80, Section 2.4]). If we additionally assume that $\|\nabla \alpha\|_{L^\infty(D)} \in L^p(\Omega)$ for all $p \in [1, \infty)$, Assumption 2.2.1 (ii) also holds.

Remark 2.5.1. Pathwise evaluations of log-normal random fields with the Matèrn covariance function, given by

$$r_g(|x-x'|) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \big(\sqrt{2\nu}\frac{|x-x'|}{\lambda_C}\big)^\nu K_\nu\big(2\sqrt{2\nu}\frac{|x-x'|}{\lambda_C}\big),$$

where $\Gamma(\cdot)$ is the gamma function, $K_\nu(\cdot)$ is the modified Bessel function of the second kind, ν is the smoothness parameter, σ^2 is the variance and λ_C is the correlation length, belong to $C^1(\overline{D})$ a.s. (see [42, Remark 4]) when $\nu > 1$.

The KL expansion, truncated after a finite number of terms $N_{\mathrm{KL}} \in \mathbb{N}$, is known to provide the optimal finite representation of the random field g in the sense of the following proposition, see [92, Theorem 1.2.2].

Proposition 2.5.2. *Among all truncated expansions with $N_{\mathrm{KL}} \in \mathbb{N}$ terms approximating the random field g that take the form*

$$g_{N_{\mathrm{KL}}}(x,\omega) = \sum_{m=1}^{N_{\mathrm{KL}}} \sqrt{\lambda_m}\xi_m(\omega)\phi_m(x),$$

where $(\phi_m)_{m\in\mathbb{N}}$ is a system of mutually orthogonal functions in $L^2(D)$, the KL expansion minimizes the integrated mean square error

$$\int_D E\Big[\Big(\sum_{m=N_{\mathrm{KL}}+1}^{\infty} \sqrt{\lambda_m}\xi_m\phi_m(x)\Big)^2\Big]\, dx.$$

Chapter 3

Abstract Adaptive Multilevel Monte Carlo Methods

In this chapter we introduce abstract multilevel Monte Carlo methods for approximating the expected solution to variational equalities of the form (2.4) or variational inequalities of the form (2.8), and for approximating expected outputs of interest for these problems. We present a set of assumptions for the spatial pathwise approximations to the solutions, together with convergence results and bounds for the expected computational cost of the MLMC methods provided by these assumptions. The framework presented in this chapter extends the range of methods that can be utilized for constructing the pathwise approximations to adaptive finite element methods, as will be shown in Chapter 4.

We let Assumption 2.2.1 hold for both types of problem. In addition we let Assumption 2.3.1 hold for problem (2.8). Then, according to Propositions 2.2.1 and 2.3.1, the solutions u to problems (2.4) and (2.8) belong to $L^p(\Omega; H_0^1(D))$ for all $p \in [1, p_f)$, where p_f is defined in Assumption 2.2.1 (iii). Particularly, since $p_f \in (2, \infty]$, we have $u \in L^2(\Omega; H_0^1(D))$.

We omit the dependence of u on the spatial variable and use the notation $u(\omega)$ for $u(\cdot, \omega)$ in this chapter.

3.1 Approximation of the expected solution

In this section we construct an approximation to the expected value $\mathbb{E}[u]$ introduced in Section 2.4. We concentrate here on the discretization in the stochastic domain and only make some assumptions on the approximation in the spatial domain; the latter will be discussed in detail in Chapter 4. We use multilevel Monte Carlo methods for integration in the stochastic domain.

In order to introduce the multilevel Monte Carlo method we first define the Monte Carlo estimator

$$E_M[u] := \frac{1}{M} \sum_{i=1}^{M} u_i \in L^2(\Omega; H_0^1(D)), \tag{3.1}$$

where u_i, $i = 1, \dots, M$, denote $M \in \mathbb{N}$ independent, identically distributed (i.i.d.) copies of u.

The following lemma provides a representation for the Monte Carlo approximation error, see [14, 16, 61].

Lemma 3.1.1. *The Monte Carlo approximation $E_M[u]$ defined in* (3.1) *of the expectation $\mathbb{E}[u]$ satisfies*

$$\|\mathbb{E}[u] - E_M[u]\|^2_{L^2(\Omega;H^1_0(D))} = M^{-1} V[u],$$

where

$$V[u] := \mathbb{E}[\|u\|^2_{H^1(D)}] - \|\mathbb{E}[u]\|^2_{H^1(D)}.$$

Proof. Since the $(u_i)_{i=1}^M$ are i.i.d., we have

$$\begin{aligned}
\|\mathbb{E}[u] - E_M[u]\|^2_{L^2(\Omega;H^1_0(D))} &= \mathbb{E}\left[\left\|\mathbb{E}[u] - \frac{1}{M}\sum_{i=1}^M u_i\right\|^2_{H^1(D)}\right] \\
&= \mathbb{E}\left[\left\|\frac{1}{M}\sum_{i=1}^M (\mathbb{E}[u] - u_i)\right\|^2_{H^1(D)}\right] \\
&= \tfrac{1}{M^2}\textstyle\sum_{i=1}^M \mathbb{E}\left[\|\mathbb{E}[u] - u_i\|^2_{H^1(D)}\right] \\
&= \frac{1}{M}\mathbb{E}\left[\|\mathbb{E}[u] - u\|^2_{H^1(D)}\right] \\
&= \tfrac{1}{M}\left(\mathbb{E}[\|u\|^2_{H^1(D)}] - \|\mathbb{E}[u]\|^2_{H^1(D)}\right).
\end{aligned}$$

$\square$

Remark 3.1.1. It is easy to see that

$$V[u] = \mathbb{E}[\|u\|^2_{H^1(D)}] - \|\mathbb{E}[u]\|^2_{H^1(D)} \leq \mathbb{E}[\|u\|^2_{H^1(D)}] = \|u\|^2_{L^2(\Omega;H^1_0(D))}.$$

We will use this relation later for proving convergence of multilevel Monte Carlo methods.

Now, we introduce spatial approximations for $u(\omega)$, $\omega \in \Omega$. We do not specify how these approximations can be constructed and only state a set of assumptions for them.

For given initial tolerance $Tol_0 > 0$ and reduction factor $0 < q < 1$ we define a sequence of tolerances by

$$Tol_l := q Tol_{l-1}, \quad l \in \mathbb{N}. \tag{3.2}$$

For each fixed $\omega \in \Omega$ we introduce a sequence of approximations $\tilde{u}_l(\omega)$, $l \in \mathbb{N} \cup \{0\}$ to the solution $u(\omega)$ and assume the following properties.

Assumption 3.1.1. For all $l \in \mathbb{N} \cup \{0\}$ the mappings $\Omega \ni \omega \mapsto \tilde{u}_l(\omega) \in H_0^1(D)$ are measurable, and there exists $C_{\text{dis}} \in L^2(\Omega)$ such that for almost all $\omega \in \Omega$ the approximate solutions $\tilde{u}_l(\omega)$ satisfy the error estimate

$$\|u(\omega) - \tilde{u}_l(\omega)\|_{H^1(D)} \leq C_{\text{dis}}(\omega) Tol_l.$$

Note that Assumption 3.1.1 implies

$$\|u - \tilde{u}_l\|_{L^2(\Omega;H_0^1(D))} \leq \|C_{\text{dis}}\|_{L^2(\Omega)} Tol_l, \tag{3.3}$$

for all $l \in \mathbb{N} \cup \{0\}$.

Given a random variable ζ, we shall denote the computational cost for evaluating ζ by $Cost(\zeta)$.

Assumption 3.1.2. For all $l \in \mathbb{N} \cup \{0\}$ the approximations $\tilde{u}_l(\omega)$ to the solutions $u(\omega)$, $\omega \in \Omega$ can be evaluated at expected computational cost

$$\mathbb{E}[Cost(\tilde{u}_l)] \leq C_{\text{cost}} Tol_l^{-\gamma}$$

with constants $\gamma, C_{\text{cost}} > 0$ independent of l and ω.

Now, we are ready to introduce the inexact multilevel Monte Carlo approximation to $\mathbb{E}[u]$. For an $L \in \mathbb{N}$ we define

$$E^L[\tilde{u}_L] := \sum_{l=0}^{L} E_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}], \tag{3.4}$$

where $\tilde{u}_{-1} := 0$ and $M_l \in \mathbb{N}$, $l = 0, \ldots, L$. Note that the numbers of samples M_l may be different for different l, and that the samples on different levels are independent.

The following proposition presents a basic identity for the approximation error of the MLMC estimator, see [16].

Proposition 3.1.1. *The multilevel Monte Carlo approximation* (3.4) *of the expected solution* $\mathbb{E}[u]$ *with* $M_l \in \mathbb{N}$, $l = 0, \ldots, L$, *satisfies*

$$\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|^2_{L^2(\Omega;H_0^1(D))} = \|\mathbb{E}[u - \tilde{u}_L]\|^2_{H^1(D)} + \sum_{l=0}^{L} M_l^{-1} V[\tilde{u}_l - \tilde{u}_{l-1}]. \tag{3.5}$$

Proof. Since the expectations on different levels are estimated independently, we have

$$
\begin{aligned}
\left\|\mathbb{E}[u] - E^L[\tilde{u}_L]\right\|^2_{L^2(\Omega;H_0^1(D))} &= \|\mathbb{E}[u] - \mathbb{E}[\tilde{u}_L]\|^2_{L^2(\Omega;H_0^1(D))} + \|\mathbb{E}[\tilde{u}_L] - E^L[\tilde{u}_L]\|^2_{L^2(\Omega;H_0^1(D))} \\
&= \|\mathbb{E}[u-\tilde{u}_L]\|^2_{H^1(D)} + \left\|\sum_{l=0}^{L}(\mathbb{E} - E_{M_l})[\tilde{u}_l - \tilde{u}_{l-1}]\right\|^2_{L^2(\Omega;H_0^1(D))} \\
&= \|\mathbb{E}[u-\tilde{u}_L]\|^2_{H^1(D)} + \sum_{l=0}^{L}\|(\mathbb{E} - E_{M_l})[\tilde{u}_l - \tilde{u}_{l-1}]\|^2_{L^2(\Omega;H_0^1(D))} \\
&= \|\mathbb{E}[u-\tilde{u}_L]\|^2_{H^1(D)} + \sum_{l=0}^{L} M_l^{-1} V[\tilde{u}_l - \tilde{u}_{l-1}].
\end{aligned}
$$

□

Proposition 3.1.1 shows that the error of the multilevel Monte Carlo estimator has two components, one of which depends on the discretization of the random solutions and the second one represents the sampling error and includes a variance-like operator.

We finally state a convergence theorem for the inexact multilevel Monte Carlo methods described in this section. The proof follows the same steps as the proofs of similar results [29, 41], see also [61].

Theorem 3.1.1. *Let Assumptions 3.1.1-3.1.2 hold. Then for any $Tol > 0$ there exists an $L \in \mathbb{N}$ and a sequence of integers $\{M_l\}_{l=0}^L$ providing*

$$\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|_{L^2(\Omega;H_0^1(D))} \leq Tol,$$

and the estimator $E^L[\tilde{u}_L]$ can be evaluated at expected computational cost

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] = \begin{cases} \mathcal{O}(Tol^{-2}), & \gamma < 2, \\ \mathcal{O}(L^2 Tol^{-2}), & \gamma = 2, \\ \mathcal{O}(Tol^{-\gamma}), & \gamma > 2, \end{cases}$$

where the constants depend only on q, γ, C_{cost}, Tol_0, $\|u\|_{L^2(\Omega;H_0^1(D))}$ and $\|C_{\text{dis}}\|_{L^2(\Omega)}$.

Proof. We set L to be the smallest integer such that $L \geq log_q(Tol_0^{-1} 2^{-1/2}\|C_{\text{dis}}\|^{-1}_{L^2(\Omega)} Tol)$, which ensures $Tol_L \leq 2^{-1/2}\|C_{\text{dis}}\|^{-1}_{L^2(\Omega)} Tol < Tol_{L-1}$. Proposition 3.1.1, together with Jensen's inequality, Remark 3.1.1, the Cauchy-Schwarz and triangle inequalities, the choice of L and (3.3), provides

$$\begin{aligned}
&\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|^2_{L^2(\Omega;H^1_0(D))} = \|\mathbb{E}[u - \tilde{u}_L]\|^2_{H^1(D)} + \sum_{l=0}^{L} M_l^{-1} V[\tilde{u}_l - \tilde{u}_{l-1}] \\
&\leq (\mathbb{E}[\|u - \tilde{u}_L\|_{H^1(D)}])^2 + \sum_{l=0}^{L} M_l^{-1} \|\tilde{u}_l - \tilde{u}_{l-1}\|^2_{L^2(\Omega;H^1_0(D))} \\
&\leq \|u - \tilde{u}_L\|^2_{L^2(\Omega;H^1_0(D))} + \sum_{l=0}^{L} M_l^{-1} (\|\tilde{u}_l - u\|_{L^2(\Omega;H^1_0(D))} + \|u - \tilde{u}_{l-1}\|_{L^2(\Omega;H^1_0(D))})^2 \\
&\leq \tfrac{1}{2} Tol^2 + M_0^{-1} (\|C_{\text{dis}}\|_{L^2(\Omega)} Tol_0 + \|u\|_{L^2(\Omega;H^1_0(D))})^2 \\
&\qquad\qquad + (1 + q^{-1})^2 \|C_{\text{dis}}\|^2_{L^2(\Omega)} \textstyle\sum_{l=1}^{L} M_l^{-1} Tol_l^2.
\end{aligned}$$

Now, we choose M_0 to be the smallest integer such that

$$M_0 \geq C_0 Tol^{-2},$$

where $C_0 := 4(\|C_{\text{dis}}\|_{L^2(\Omega)} Tol_0 + \|u\|_{L^2(\Omega;H^1_0(D))})^2$. We choose the values for M_l, $l = 1, \dots L$ differently for different values of γ.

For $\gamma < 2$ we set M_l to be the smallest integer such that

$$M_l \geq C_1 q^{\frac{\gamma+2}{2} l} Tol^{-2}, \;\; l = 1, \dots, L,$$

with $C_1 := 4\|C_{\text{dis}}\|^2_{L^2(\Omega)} ((1 - q^{\frac{2-\gamma}{2}})^{-1} - 1)(1 + q^{-1})^2 Tol_0^2$. Then we have

$$\begin{aligned}
&\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|^2_{L^2(\Omega;H^1_0(D))} \leq \\
&\qquad \tfrac{1}{2} Tol^2 + \tfrac{1}{4} Tol^2 + Tol^2 (1 + q^{-1})^2 \|C_{\text{dis}}\|^2_{L^2(\Omega)} C_1^{-1} Tol_0^2 \sum_{l=1}^{L} q^{\frac{2-\gamma}{2} l} \leq Tol^2.
\end{aligned}$$

For $\gamma = 2$ we set M_l to be the smallest integer such that

$$M_l \geq C_2 L q^{2l} Tol^{-2}, \;\; l = 1, \dots, L,$$

with $C_2 := 4\|C_{\text{dis}}\|^2_{L^2(\Omega)} (1 + q^{-1})^2 Tol_0^2$. Then the error can be bounded as follows

$$\begin{aligned}
&\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|^2_{L^2(\Omega;H^1_0(D))} \leq \\
&\qquad \tfrac{1}{2} Tol^2 + \tfrac{1}{4} Tol^2 + Tol^2 (1 + q^{-1})^2 \|C_{\text{dis}}\|^2_{L^2(\Omega)} C_2^{-1} Tol_0^2 \leq Tol^2.
\end{aligned}$$

Finally, for $\gamma > 2$ we choose M_l to be the smallest integer such that

$$M_l \geq C_3 q^{\frac{2-\gamma}{2} L} q^{\frac{\gamma+2}{2} l} Tol^{-2}, \;\; l = 1, \dots, L,$$

with $C_3 := 4\|C_{\text{dis}}\|^2_{L^2(\Omega)}(1-q^{\frac{\gamma-2}{2}})^{-1}(1+q^{-1})^2 Tol_0^2$. Then

$$\|\mathbb{E}[u]-\mathbb{E}^L[\tilde{u}_L]\|^2_{L^2(\Omega;H_0^1(D))} \leq$$
$$\tfrac{1}{2}Tol^2+\tfrac{1}{4}Tol^2+Tol^2(1+q^{-1})^2\|C_{\text{dis}}\|^2_{L^2(\Omega)}C_3^{-1}Tol_0^2\sum_{l=0}^{L}q^{\frac{\gamma-2}{2}l} \leq Tol^2.$$

We set $Tol_{-1} := 0$ and utilize Assumption 3.1.2, then the expected computational cost of the MLMC estimator is bounded by

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] \leq C_{\text{cost}}\sum_{l=0}^{L}M_l(Tol_l^{-\gamma}+Tol_{l-1}^{-\gamma}) \leq C_{\text{cost}}(1+q^{\gamma})\sum_{l=0}^{L}M_l Tol_l^{-\gamma}.$$

Now, we denote $\tilde{C}_i := \max\{C_0, C_i\},\ i=1,2,3$ and consider different values of γ.

For $\gamma < 2$ we have

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] \leq C_{\text{cost}}(1+q^{\gamma})(\tilde{C}_1 Tol_0^{-\gamma}Tol^{-2}\sum_{l=0}^{L}q^{\frac{2-\gamma}{2}l}+\sum_{l=0}^{L}Tol_l^{-\gamma}) \leq$$
$$C_{\text{cost}}(1+q^{\gamma})(\tilde{C}_1 Tol_0^{-\gamma}(1-q^{\frac{2-\gamma}{2}})^{-1}Tol^{-2}+2^{\frac{\gamma}{2}}q^{-\gamma}\|C_{\text{dis}}\|^{\gamma}_{L^2(\Omega)}(1-q^{\gamma})^{-1}Tol^{-\gamma}),$$

for $\gamma = 2$

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] \leq C_{\text{cost}}(1+q^2)(\tilde{C}_2 Tol_0^{-2}L(L+1)+2q^{-2}\|C_{\text{dis}}\|^2_{L^2(\Omega)}(1-q^2)^{-1})Tol^{-2},$$

and for $\gamma > 2$

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] \leq C_{\text{cost}}(1+q^{\gamma})(\tilde{C}_3 Tol_0^{-2}Tol^{-\gamma}2^{\frac{\gamma-2}{2}}q^{2-\gamma}\|C_{\text{dis}}\|^{\gamma-2}_{L^2(\Omega)}\sum_{l=0}^{L}q^{\frac{\gamma-2}{2}l}$$
$$+\sum_{l=0}^{L}Tol_l^{-\gamma}) \leq$$
$$C_{\text{cost}}(1+q^{\gamma})(\tilde{C}_3 Tol_0^{-2}2^{\frac{\gamma-2}{2}}q^{2-\gamma}\|C_{\text{dis}}\|^{\gamma-2}_{L^2(\Omega)}(1-q^{\frac{\gamma-2}{2}})^{-1}$$
$$+2^{\frac{\gamma}{2}}q^{-\gamma}\|C_{\text{dis}}\|^{\gamma}_{L^2(\Omega)}(1-q^{\gamma})^{-1})Tol^{-\gamma}.$$

□

Remark 3.1.2. If we let the stronger set of assumptions described in Remark 2.2.2 hold, it is enough to assume $p_f \in [2,\infty]$ in order for Theorem 3.1.1 to hold.

3.2 Approximation of the expected output of interest

In this section we approximate the expected output of interest $\mathbb{E}[Q(u)]$ defined in (2.10). As in the previous section, we concentrate on the approximation in the stochastic domain using the multilevel Monte Carlo method.

The Monte Carlo approximation to $\mathbb{E}[Q(u)]$ is defined by

$$E_M[Q(u)] := \frac{1}{M}\sum_{i=1}^{M} Q(u_i), \tag{3.6}$$

where, again, $M \in \mathbb{N}$ and u_i, $i = 1, \dots, M$ are i.i.d. copies of u.

It is well known and easy to verify that the Monte Carlo approximation (3.6) has the properties

$$\mathbb{E}[E_M[Q(u)]] = \mathbb{E}[Q(u)], \quad \mathbb{V}[E_M[Q(u)]] = M^{-1}\mathbb{V}[Q(u)]. \tag{3.7}$$

We follow the previous section and assume that for each $\omega \in \Omega$ we possess sequences of approximations $\tilde{u}_l(\omega)$, $l \in \mathbb{N} \cup \{0\}$ which fulfill the following assumption.

Assumption 3.2.1. For all $l \in \mathbb{N} \cup \{0\}$ the mappings $\Omega \ni \omega \mapsto \tilde{u}_l(\omega) \in H_0^1(D)$ are measurable, and there exists $C_{\text{dis}} \in L^{p_{\text{dis}}}(\Omega)$ for some $p_{\text{dis}} \in [2, p_f)$ such that for almost all $\omega \in \Omega$ the approximate solutions $\tilde{u}_l(\omega)$ satisfy the error estimate

$$\|u(\omega) - \tilde{u}_l(\omega)\|_{H^1(D)} \leq C_{\text{dis}}(\omega) Tol_l.$$

Note that Assumption 3.2.1 implies

$$\|u - \tilde{u}_l\|_{L^p(\Omega; H_0^1(D))} \leq \|C_{\text{dis}}\|_{L^p(\Omega)} Tol_l, \tag{3.8}$$

for all $p \in [1, p_{\text{dis}}]$ and for all $l \in \mathbb{N} \cup \{0\}$.

We also assume that the approximations $\tilde{u}_l(\omega)$ satisfy Assumption 3.1.2 and that the cost of computing the quantity of interest $Q(\tilde{u}_l(\omega))$ is negligible compared to the cost of computing the function $\tilde{u}_l(\omega)$ for all $l \in \mathbb{N} \cup \{0\}$ and $\omega \in \Omega$.

For a given $L \in \mathbb{N}$ we introduce the multilevel Monte Carlo approximation to $\mathbb{E}[Q(u)]$ as

$$E^L[Q(\tilde{u}_L)] := \sum_{l=0}^{L} E_{M_l}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})], \tag{3.9}$$

where we set $Q(\tilde{u}_{-1}) := 0$ and $M_l \in \mathbb{N}$, $l = 0, \dots, L$. Again, the numbers of samples M_l may be different for different levels l.

We are interested in the so-called *mean square error* (MSE) defined as follows

$$e^2(E^L[Q(\tilde{u}_L)]) := \mathbb{E}\left[\left(\mathbb{E}[Q(u)] - E^L[Q(\tilde{u}_L)]\right)^2\right].$$

The following proposition states a basic representation for the MSE of the MLMC estimator, see e.g. [29].

Proposition 3.2.1. *The multilevel Monte Carlo approximation* (3.9) *of the expected output of interest* $\mathbb{E}[Q(u)]$ *with* $M_l \in \mathbb{N}$, $l = 0, \ldots, L$, *satisfies*

$$e^2(E^L[Q(\tilde{u}_L)]) = (\mathbb{E}[Q(u) - Q(\tilde{u}_L)])^2 + \sum_{l=0}^{L} M_l^{-1} \mathbb{V}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})]. \tag{3.10}$$

Proof. Using (3.7), (3.9) and independence of samples on different levels we have

$$\begin{aligned} e^2(E^L[Q(\tilde{u}_L)]) &= (\mathbb{E}[Q(u)] - \mathbb{E}[Q(\tilde{u}_L)])^2 + \mathbb{V}[E^L[Q(\tilde{u}_L)]] \\ &= (\mathbb{E}[Q(u) - Q(\tilde{u}_L)])^2 + \sum_{l=0}^{L} \mathbb{V}[E_{M_l}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})]] \\ &= (\mathbb{E}[Q(u) - Q(\tilde{u}_L)])^2 + \sum_{l=0}^{L} M_l^{-1} \mathbb{V}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})]. \end{aligned}$$

□

We are now ready to state a convergence theorem for the inexact multilevel Monte Carlo methods, for approximating the expected output of interest introduced in Section 2.4.

Theorem 3.2.1. *Let Assumption 2.4.1 hold and* $p_f \in (2\frac{p_Q}{p_Q-2}, \infty]$. *Further, let Assumption 3.2.1 hold with* $p_{\text{dis}} = 2\frac{p_Q}{p_Q-2}$ *and let Assumption 3.1.2 hold. Then for any* $Tol > 0$ *there exists an* $L \in \mathbb{N}$ *and a sequence of integers* $\{M_l\}_{l=0}^{L}$, *such that*

$$e(E^L[Q(\tilde{u}_L)]) \leq Tol,$$

and the estimator $E^L[Q(\tilde{u}_L)]$ *can be evaluated at expected computational cost*

$$\mathbb{E}[Cost(E^L[Q(\tilde{u}_L)])] = \begin{cases} \mathcal{O}(Tol^{-2}), & \gamma < 2, \\ \mathcal{O}(L^2 Tol^{-2}), & \gamma = 2, \\ \mathcal{O}(Tol^{-\gamma}), & \gamma > 2, \end{cases}$$

where the constants depend only on q, γ, C_{cost}, Tol_0, $\|u\|_{L^{p_{\text{dis}}}(\Omega; H_0^1(D))}$, $\|C_{\text{dis}}\|_{L^{p_{\text{dis}}}(\Omega)}$, $\|C_Q\|_{L^{p_Q}(\Omega)}$.

Proof. We set L to be the smallest integer such that $L \geq log_q(Tol_0^{-1} 2^{-1/2} \|C_Q\|^{-1}_{L^{p_Q}(\Omega)} \|C_{\text{dis}}\|^{-1}_{L^{p_{\text{dis}}}(\Omega)} Tol)$, which ensures $Tol_L \leq 2^{-1/2} \|C_Q\|^{-1}_{L^{p_Q}(\Omega)} \|C_{\text{dis}}\|^{-1}_{L^{p_{\text{dis}}}(\Omega)} Tol < Tol_{L-1}$. Proposition 3.2.1, together with the Cauchy-Schwarz, triangle and Hölder's inequalities, (3.8), Assumption 2.4.1 and the choice of L, yields

$$\begin{aligned}
e^2(E^L[Q(\tilde{u}_L)]) &= (\mathbb{E}[Q(u) - Q(\tilde{u}_L)])^2 + \sum_{l=0}^{L} M_l^{-1}\mathbb{V}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})] \\
&\leq \|Q(u) - Q(\tilde{u}_L)\|^2_{L^2(\Omega)} + \sum_{l=0}^{L} M_l^{-1}\|Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})\|^2_{L^2(\Omega)} \\
&\leq \|Q(u) - Q(\tilde{u}_L)\|^2_{L^2(\Omega)} + \sum_{l=0}^{L} M_l^{-1}\Big(\|Q(\tilde{u}_l) - Q(u)\|^2_{L^2(\Omega)} \\
&\qquad\qquad + \|Q(u) - Q(\tilde{u}_{l-1})\|^2_{L^2(\Omega)}\Big)^2 \\
&\leq \|C_Q\|^2_{L^{p_Q}(\Omega)}\|C_{\text{dis}}\|^2_{L^{p_{\text{dis}}}(\Omega)}Tol_L^2 \\
&\qquad + (1+q^{-1})^2\|C_Q\|^2_{L^{p_Q}(\Omega)}\|C_{\text{dis}}\|^2_{L^{p_{\text{dis}}}(\Omega)}\sum_{l=0}^{L} M_l^{-1}Tol_l^2 \\
&\leq \frac{1}{2}Tol^2 + M_0^{-1}\|C_Q\|^2_{L^{p_Q}(\Omega)}(\|C_{\text{dis}}\|_{L^{p_{\text{dis}}}(\Omega)}Tol_0 + \|u\|_{L^{p_{\text{dis}}}(\Omega;H_0^1(D))})^2 \\
&\qquad + (1+q^{-1})^2\|C_Q\|^2_{L^{p_Q}(\Omega)}\|C_{\text{dis}}\|^2_{L^{p_{\text{dis}}}(\Omega)}\sum_{l=1}^{L} M_l^{-1}Tol_l^2.
\end{aligned}$$

Finally, we choose M_0 to be the smallest integer such that

$$M_0 \geq C_0 Tol^{-2},$$

where $C_0 := 4\|C_Q\|^2_{L^{p_Q}(\Omega)}(\|C_{\text{dis}}\|_{L^{p_{\text{dis}}}(\Omega)}Tol_0 + \|u\|_{L^{p_{\text{dis}}}(\Omega;H_0^1(D))})^2$. We also set M_l to be the smallest integers such that

$$M_l \geq \begin{cases} C_1 q^{\frac{\gamma+2}{2}l} Tol^{-2}, & \gamma < 2, \\ C_2 L q^{2l} Tol^{-2}, & \gamma = 2, \\ C_3 q^{\frac{2-\gamma}{2}L} q^{\frac{\gamma+2}{2}l} Tol^{-2}, & \gamma > 2, \end{cases}$$

for $l = 1, \dots, L$, where

$$C_1 := 4\|C_Q\|^2_{L^{p_Q}(\Omega)}\|C_{\text{dis}}\|^2_{L^{p_{\text{dis}}}(\Omega)}((1 - q^{\frac{2-\gamma}{2}})^{-1} - 1)(1+q^{-1})^2 Tol_0^2,$$

$$C_2 := 4\|C_Q\|^2_{L^{p_Q}(\Omega)}\|C_{\text{dis}}\|^2_{L^{p_{\text{dis}}}(\Omega)}(1+q^{-1})^2 Tol_0^2,$$

$$C_3 := 4\|C_Q\|^2_{L^{p_Q}(\Omega)}\|C_{\text{dis}}\|^2_{L^{p_{\text{dis}}}(\Omega)}(1 - q^{\frac{\gamma-2}{2}})^{-1}(1+q^{-1})^2 Tol_0^2.$$

The rest of the proof is then almost identical to the proof of Theorem 3.1.1. □

Remark 3.2.1. If we let the stronger set of assumptions described in Remark 2.2.2 hold, it is enough to assume $p_f \in [2\frac{p_Q}{p_Q-2}, \infty]$ in order for Theorem 3.1.1 to hold.

Chapter 4

Multilevel Monte Carlo Finite Element Methods

In this chapter we introduce and investigate two multilevel Monte Carlo methods that fit into the framework described in Chapter 3. We first overview well established multilevel Monte Carlo finite element methods [2, 14, 16, 28, 60, 87] and then introduce a novel method that combines the ideas of MLMC and adaptive finite element methods and present a theoretical analysis for this method. We concentrate on the elliptic variational equalities introduced in Section 2.2, i.e.

$$u(\cdot,\omega) \in H_0^1(D): \qquad a(\omega; u(\cdot,\omega), v) = \ell(\omega; v), \qquad \forall v \in H_0^1(D), \tag{4.1}$$

where the bilinear and linear forms are defined in (2.5). We let Assumption 2.2.1 hold throughout the chapter. The analysis presented in this chapter simplifies in a transparent way when elliptic problems that fulfill the set of stronger assumptions discussed in Remark 2.2.2 are considered.

4.1 Notation

We consider partitions of the spatial domain $\overline{D} \subset \mathbb{R}^d$ into non-overlapping subdomains. We require these subdomains to be simplices, i.e. line segments if $d = 1$, triangles if $d = 2$, or tetrahedra if $d = 3$. The union of such subdomains is labelled $\mathcal{T}$. We denote a simplex by T and call it an *element*.

We call a partition $\mathcal{T}$ *admissible* if any two elements in $\mathcal{T}$ are either disjoint or their intersection is a vertex or a complete edge ($d = 2, 3$) or a complete face ($d = 3$). Admissibility of a mesh means that this mesh does not contain hanging nodes, i.e. nodes that not only exist in element corners, but also on element edges or faces, see Figure 4.1. In the case $d = 1$ any partition is admissible.

Let $\mathcal{E}$ denote the set of interior nodes in the case $d = 1$, the set of interior edges in the case $d = 2$ and the set of interior faces in the case $d = 3$. For any $E \in \mathcal{E}$, $d > 1$ we set $h_E := |E|^{\frac{1}{d-1}}$, where $|\cdot|$ denotes the $(d-1)$-dimensional Lebesgue measure. In what follows we call an $E \in \mathcal{E}$ a *face* for all values of d.

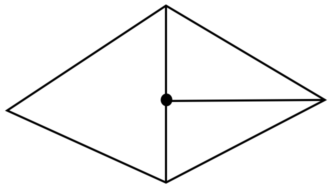

FIGURE 4.1: An example of a hanging node (•).

We define the shape parameter of a mesh $\mathcal{T}$ as

$$C_{\mathcal{T}} := \begin{cases} \max\limits_{\substack{T_1,T_2\in\mathcal{T}\\ T_1\cap T_2\neq\emptyset}} \dfrac{h_{T_1}}{h_{T_2}}, & d=1,\\ \max\limits_{T\in\mathcal{T}} \dfrac{\operatorname{diam}(T)}{h_T}, & d=2,3, \end{cases}$$

where $\operatorname{diam}(\cdot)$ denotes the Euclidean diameter of an element, $h_T := |T|^{\frac{1}{d}}$ and $|\cdot|$ denotes the d-dimensional Lebesgue measure. We call a partition *shape-regular* if its shape parameter is bounded from above. In the case $d=2$, shape-regularity of a partition means that the smallest angles of all elements in this partition stay bounded away from zero.

Note that if a 2- or 3-dimensional partition $\mathcal{T}$ is shape-regular then for any $T_1, T_2 \in \mathcal{T}$ such that $T_1 \cap T_2 \neq \emptyset$, and for any $E_1, E_2 \in \mathcal{E}$ such that $E_1 \cap E_2 \neq \emptyset$, the ratios $\frac{h_{T_1}}{h_{T_2}}$, $\frac{h_{E_1}}{h_{E_2}}$ and $\frac{h_{E_i}}{h_{T_j}}$, where $i, j \in \{1,2\}$, are bounded from above and below by constants dependent only on $C_{\mathcal{T}}$.

We call a sequence of nested partitions $\mathcal{T}_0 \subset \mathcal{T}_1 \subset \ldots$ *shape-regular* if $C_{\mathcal{T}_k} \leq \Gamma < \infty$, $k \in \mathbb{N} \cup \{0\}$ for some Γ only dependent on $\mathcal{T}_0$.

We further denote the union of elements sharing a face $E \in \mathcal{E}$ and sharing at least one vertex with $T \in \mathcal{T}$ by ϕ_E and ϕ_T respectively.

Finally, let $|\mathcal{T}|$ denote the number of elements in a given partition $\mathcal{T}$.

4.2 Uniform multilevel Monte Carlo finite element methods

In this section we introduce well established multilevel Monte Carlo finite element methods, based on a hierarchy of uniform meshes [2, 14, 16, 28, 60, 87], and trace how they fit into the framework described in Chapter 3.

We start by deriving spatial approximations of solutions to the pathwise weak problem (4.1). We consider a sequence of partitions $(\mathcal{T}_{(k)})_{k\in\mathbb{N}\cup\{0\}}$ of the spatial domain D into simplices. We assume that $\mathcal{T}_{(0)}$ is admissible and shape-regular and $(\mathcal{T}_{(k)})_{k\in\mathbb{N}}$ is constructed by consecutive refinement by bisection [68, 76] preserving shape-regularity, i.e. there exist $\Gamma > 0$, such that $C_{\mathcal{T}_{(k)}} \leq \Gamma$ for all $k \in \mathbb{N} \cup \{0\}$. We also have

$$h_k = 2^{-k\frac{b}{d}} h_0, \qquad k \in \mathbb{N}, \tag{4.2}$$

where $h_k := \max_{T \in \mathcal{T}_{(k)}} h_T$, $k \in \mathbb{N} \cup \{0\}$ and $b \geq 1$ denotes the number of bisections applied to each element.

Remark 4.2.1. "Red" refinement [12, 15] can also be utilized for uniform mesh refinement, which leads to (4.2) with $b = d$.

We now define a sequence of nested first order finite element spaces

$$S_{(0)} \subset S_{(1)} \subset \cdots \subset S_{(k)} \subset \ldots,$$

where

$$S_{(k)} := \{v \in H_0^1(D) \cap C(\overline{D}) : \ v|_T \in \mathcal{P}_1(T),\ \forall T \in \mathcal{T}_{(k)}\}, \tag{4.3}$$

where $\mathcal{P}_1(T)$ denotes the space of linear polynomials on T. Since the meshes $(\mathcal{T}_{(k)})_{k \in \mathbb{N} \cup \{0\}}$ are shape-regular, and the dimensions of the spaces $S_{(k)}$ coincide with the numbers of interior vertices in the corresponding meshes, the relation $\dim S_{(k)} \simeq |\mathcal{T}_{(k)}|$ holds for all $k \in \mathbb{N} \cup \{0\}$, where the constants depend only on the shape-regularity parameter Γ.

We consider the pathwise approximations $u_{(k)}(\cdot, \omega) \in S_{(k)}$, $k \in \mathbb{N} \cup \{0\}$, characterized by

$$u_{(k)}(\cdot, \omega) \in S_{(k)} : \qquad a(\omega; u_{(k)}(\cdot, \omega), v) = \ell(\omega; v), \qquad \forall v \in S_{(k)},\ \omega \in \Omega, \tag{4.4}$$

where the bilinear form a and the linear form ℓ are as in (2.5).

Analogously to the infinite dimensional case, for almost all $\omega \in \Omega$ the variational equation (4.4) admits a unique solution $u_{(k)}(\cdot, \omega)$ according to the Lax-Milgram theorem and there holds

$$\|u_{(k)}(\cdot, \omega)\|_{H^1(D)} \leq \frac{1}{c_a(\omega)} \|f(\cdot, \omega)\|_{L^2(D)}. \tag{4.5}$$

Furthermore, we have $u_{(k)} \in L^p(\Omega; S_{(k)})$ for all $p \in [1, p_f)$, where p_f is defined in Assumption 2.2.1 (iii), and in particular the solution mapping $\Omega \ni \omega \mapsto u_{(k)}(\omega) \in S_{(k)} \subset H_0^1(D)$ is measurable.

The finite element solutions admit the following a priori error estimate.

Lemma 4.2.1. *The error estimate*

$$\|u(\cdot, \omega) - u_{(k)}(\cdot, \omega)\|_{H^1(D)} \leq C_{\mathrm{un}}(\omega) h_k \tag{4.6}$$

holds for almost all $\omega \in \Omega$, where

$$C_{\mathrm{un}}(\omega) := c_{\mathrm{un}} \left(\frac{\alpha_{\max}(\omega)}{\alpha_{\min}(\omega)} \right)^{\frac{1}{2}} \left(\frac{1}{\alpha_{\min}(\omega)} + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha_{\min}^2(\omega)} \right) \|f(\cdot, \omega)\|_{L^2(D)} > 0,$$

and $c_{\mathrm{un}} > 0$ depends only on D and the mesh shape-regularity parameter Γ.

Proof. Assumption 2.2.1 provides a.s. H^2-regularity of pathwise solutions $u_{(k)}(\cdot,\omega)$, see Proposition 2.2.2. According to well known results ([20, 49], see also [28, Lemma 3.7, Lemma 3.8]), we have

$$\|u(\cdot,\omega) - u_{(k)}(\cdot,\omega)\|_{H^1(D)} \le h_k c_{\text{un}} \left(\frac{\alpha_{\max}(\omega)}{\alpha_{\min}(\omega)}\right)^{\frac{1}{2}} \|u(\cdot,\omega)\|_{H^2(D)},$$

where $c_{\text{un}} > 0$ depends only on D and the mesh shape-regularity parameter Γ. Incorporating the estimate from Proposition 2.2.2 yields the claim of the lemma. □

Finiteness of $\|C_{\text{un}}\|_{L^p(\Omega)}$ for all $p \in [1, p_f)$ follows from Assumption 2.2.1 and several applications of Hölder's inequality, see Proposition C.0.6 for details.

We can choose $\mathcal{T}_{(0)}$ such that $h_0 = Tol_0$ and define an uniform MLMC hierarchy according to

$$\begin{aligned} S_0 &:= S_{(0)}, \quad u_0 := u_{(0)}, \\ S_l &:= S_{(k_l)}, \quad u_l := u_{(k_l)}, \quad l \in \mathbb{N}, \end{aligned} \tag{4.7}$$

where k_l is the smallest integer such that

$$2^{-\frac{b}{d}k_l} \le q^l, \tag{4.8}$$

and $0 < q < 1$ is the reduction factor from (3.2). The choice (4.7) and Lemma 4.2.1 provide the a priori estimate

$$\|u(\cdot,\omega) - u_l(\cdot,\omega)\|_{H^1(D)} \le C_{\text{un}}(\omega) Tol_l, \tag{4.9}$$

for the finite element solutions $u_l(\cdot,\omega),\ \omega \in \Omega$.

Remark 4.2.2. It is possible to weaken Assumption 2.2.1 (ii) and extend the MLMC method to the case where the realizations of the random coefficient α are only Hölder continuous, as it was done in [28, 87]. However, we adhere to this stronger assumption and the resulting estimates in order to be able to compare the results of this and the following sections.

Usually, the solutions $u_l(\cdot,\omega)$ are computed only approximately. We denote the algebraic approximations of $u_l(\cdot,\omega)$ by $\tilde{u}_l(\cdot,\omega)$ and require that for almost all $\omega \in \Omega$ and all $l \in \mathbb{N} \cup \{0\}$ the approximations $\tilde{u}_l(\omega)$ fulfill

$$\|u_l(\cdot,\omega) - \tilde{u}_l(\cdot,\omega)\|_{H^1(D)} \le Tol_l. \tag{4.10}$$

This can be achieved by performing sufficiently many steps of any iterative solver (e.g. [20, 22]) for elliptic problems that converges for almost all $\omega \in \Omega$.

Finally, measurability of $\Omega \ni \omega \mapsto \tilde{u}_l(\cdot,\omega) \in S_l$ is inherited from u_l for all $l \in \mathbb{N} \cup \{0\}$. Assumptions 3.1.1 and 3.2.1 then follow from (4.9), (4.10), the triangle inequality and Proposition C.0.6 with $C_{\text{dis}} = C_{\text{un}} + 1$ and any $p_{\text{dis}} \in [2, p_f)$.

Remark 4.2.3. In practical applications the coefficient α is often represented as a series (e.g. (2.11) or (2.14)) that needs to be truncated in order to be feasible for numerical simulations; this leads to an additional truncation error. Furthermore, in order to solve problem (4.4) one typically needs to approximate the integrals appearing in the bilinear and the linear forms by some quadrature, because the exact evaluation is generally impossible. This again leads to an additional error. In this and the following sections we do not take into account the quadrature and the truncation errors. Instead, we cite [28, 86], where some analysis is done for these types of error.

Multigrid methods [20, 22] have linear complexity in the dimension of the finite element spaces for computing a approximate solution at a mesh level k. Thus, they have linear complexity in $|\mathcal{T}_{(k)}|$, i.e.

$$Cost(\tilde{u}_{(k)}(\omega)) \leq C_{\mathrm{MG}}(\omega)|\mathcal{T}_{(k)}|$$

for all $k \in \mathbb{N} \cup \{0\}$ and almost all ω, where C_{MG} depends only on Γ, $\alpha_{\min}$, $\alpha_{\max}$. We make the following assumption.

Assumption 4.2.1. $\mathbb{E}[C_{\mathrm{MG}}] < \infty$.

For uniform meshes, $|\mathcal{T}_{(k_l)}|$ is bounded by $h_{k_l}^{-d}$ and thus by Tol_l^{-d}, up to a constant that depends only on D. Thus, the multigrid methods satisfy Assumption 3.1.2 with $\gamma = d$ and a constant C_{cost} that depends only on D, Γ and $\mathbb{E}[C_{\mathrm{MG}}]$.

We are finally ready to state convergence theorems for the uniform MLMC-FE methods. The following theorem follows from Theorem 3.1.1 and the results presented in this section.

Theorem 4.2.1. *Let Assumption 2.2.1 hold, $u(\cdot,\omega)$, $\omega \in \Omega$ denote the solutions to problem* (4.1) *and $\tilde{u}_l(\cdot,\omega)$, $\omega \in \Omega$ denote the algebraic approximations to the finite element solutions defined in* (4.7), *that fulfill* (4.10). *Finally, let a multigrid method be used for the iterative solution of the pathwise discretized problems of the form* (4.4) *and let Assumption 4.2.1 hold. Then for any $Tol > 0$ there exists an $L \in \mathbb{N}$ and a sequence of integers $\{M_l\}_{l=0}^{L}$ providing*

$$\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|_{L^2(\Omega;H_0^1(D))} \leq Tol,$$

and the estimator $E^L[\tilde{u}_L]$ can be evaluated at expected computational cost

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] = \begin{cases} \mathcal{O}(Tol^{-2}), & d < 2, \\ \mathcal{O}(L^2 Tol^{-2}), & d = 2, \\ \mathcal{O}(Tol^{-\gamma}), & d > 2, \end{cases}$$

where the constants depend only on q, d, Γ, D, Tol_0, $\mathbb{E}[C_{\mathrm{MG}}]$, $\|u\|_{L^2(\Omega;H_0^1(D))}$ and $\|C_{\mathrm{un}}\|_{L^2(\Omega)}$.

Finally, the following theorem follows from Theorem 3.2.1 and the results presented in this section.

Theorem 4.2.2. *Let Assumption 2.2.1 hold, $u(\cdot,\omega)$, $\omega \in \Omega$ denote the solutions to problem* (4.1) *and $\tilde{u}_l(\cdot,\omega)$, $\omega \in \Omega$ denote the algebraic approximations to the finite element solutions defined in* (4.7), *that fulfill* (4.10). *Let Assumption 2.4.1 hold and $p_f \in (2\frac{p_Q}{p_Q-2}, \infty]$.*

Finally, let a multigrid method be used for the iterative solution of the pathwise discretized problems of the form (4.4) *and let Assumption 4.2.1 hold. Then for any* $Tol > 0$ *there exists an* $L \in \mathbb{N}$ *and a sequence of integers* $\{M_l\}_{l=0}^{L}$*, such that*

$$e(E^L[Q(\tilde{u}_L)]) \leq Tol,$$

and the estimator $E^L[Q(\tilde{u}_L)]$ *can be evaluated at expected computational cost*

$$\mathbb{E}[Cost(E^L[Q(\tilde{u}_L)])] = \begin{cases} \mathcal{O}(Tol^{-2}), & d < 2, \\ \mathcal{O}(L^2 Tol^{-2}), & d = 2, \\ \mathcal{O}(Tol^{-\gamma}), & d > 2, \end{cases}$$

where the constants depend only on $q, d, \Gamma, D, Tol_0, \mathbb{E}[C_{\mathrm{MG}}], \|C_Q\|_{L^{p_Q}(\Omega)}, \|u\|_{L^{p_{\mathrm{un}}}(\Omega;H_0^1(D))}$, $\|C_{\mathrm{un}}\|_{L^{p_{\mathrm{un}}}(\Omega)}$ *and* $p_{\mathrm{un}} = 2\frac{p_Q}{p_Q-2}$.

Remark 4.2.4. Note that using duality arguments (see e.g. [28, 87]) it is possible to obtain pathwise estimates of the kind

$$|Q(\omega; u(\cdot,\omega)) - Q(\omega; u_l(\cdot,\omega))| \leq C_Q(\omega) C_{\mathrm{dis}}(\omega) Tol_l^{\alpha},$$

where $\alpha \geq 1$, for almost all $\omega \in \Omega$, when uniform MLMC hierarchies are used. Therefore, it is possible to obtain a sharper error bound for the uniform MLMC-FE methods for approximation of expected outputs of interest (see e.g. [28, Theorem 4.1]).

Remark 4.2.5. Similar arguments can be used for analysing the uniform MLMC-FE methods for the variational inequality (2.8). Assumptions 2.2.1 and 2.3.1, together with application of an iterative solver [31, 44, 59, 69, 89] that converges for almost all $\omega \in \Omega$ for computing $\tilde{u}_l(\omega)$, provide Assumptions 3.1.1 and 3.2.1 for problem (2.8), see [60, 61] for details. Assumption 3.1.2 does not hold for non-linear problems in general. For problem (2.8) the expected computational cost for computing $\tilde{u}_l$ for all $l \in \mathbb{N} \cup \{0\}$ can be bounded by

$$\mathbb{E}[Cost(\tilde{u}_l)] \leq C_{\mathrm{cost}}(1 + \log(Tol_l^{-d}))^{\mu} Tol_l^{-d}. \tag{4.11}$$

Standard monotone multigrid (STDMMG) methods [59, 69] provide (4.11) under certain assumptions with $\mu = 4$ in $d = 1$ space dimension, with $\mu = 5$ in $d = 2$ space dimensions, and a suitable constant C_{cost} independent of l, see [60, Section 4.5], [11, Corollary 4.1]. In spite of computational evidence, no theoretical justification of mesh-independent convergence rates seem to be available for $d = 3$. The logarithmic term in the cost in (4.11) results in a logarithmic term in the bound for the cost of MLMC method, see [61, Theorem 3.2].

4.3 Adaptive multilevel Monte Carlo finite element methods

In this section we introduce an adaptive version of multilevel Monte Carlo methods and show how they fit into the framework described in Section 3. We consider a sequence of nested finite element spaces $(S_{(k)}(\omega))_{k\in\mathbb{N}\cup\{0\}}$, defined as in (4.3), associated with a corresponding sequence of partitions $(\mathcal{T}_{(k)}(\omega))_{k\in\mathbb{N}\cup\{0\}}$, which, for each fixed $\omega \in \Omega$, is obtained by successive adaptive refinement of a given admissible and shape-regular partition $\mathcal{T}_{(0)}(\omega) = \mathcal{T}_{(0)}$ with corresponding mesh size h_0. We set

$$S_{(0)}(\omega) = S_{(0)}, \qquad \omega \in \Omega.$$

For each fixed $\omega \in \Omega$ we apply an adaptive algorithm providing a hierarchy of subspaces $S_{(k)}(\omega)$ and corresponding approximations $u_{(k)}(\cdot,\omega)$, solutions of the problem

$$u_{(k)}(\cdot,\omega) \in S_{(k)}(\omega): \qquad a(\omega; u_{(k)}(\cdot,\omega), v) = \ell(\omega; v), \qquad \forall v \in S_{(k)}(\omega), \tag{4.12}$$

where the bilinear form a and the linear form ℓ are defined in (2.5).

As shown before, problem (4.12) admits a unique solution $u_{(k)}(\cdot,\omega)$ according to the Lax-Milgram theorem for almost all $\omega \in \Omega$ and there holds

$$\|u_{(k)}(\cdot,\omega)\|_{H^1(D)} \leq \frac{1}{c_a(\omega)} \|f(\cdot,\omega)\|_{L^2(D)}. \tag{4.13}$$

In what follows we assume convergence of the pathwise adaptive algorithm controlled by an a posteriori error estimator.

Assumption 4.3.1. For all $k \in \mathbb{N} \cup \{0\}$ and almost all $\omega \in \Omega$ we have

$$\|u(\cdot,\omega) - u_{(k)}(\cdot,\omega)\|_{H^1(D)} \leq C_{\mathrm{ad}}(\omega)\eta_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)), \tag{4.14}$$

where $C_{\mathrm{ad}} \in L^{p_{\mathrm{ad}}}(\Omega)$ for some $p_{\mathrm{ad}} \in [2, p_f)$ and $\eta_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega))$ is an a posteriori error estimator that satisfies

$$\eta_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)) \xrightarrow{k\to\infty} 0.$$

Note that this assumption suits the requirements for both MLMC methods presented in Chapter 3, although for the MLMC methods for approximation of the expected solutions we only need $p_{\mathrm{ad}} = 2$.

We can choose $\mathcal{T}_{(0)}$ and corresponding $S_{(0)}$ such that

$$h_0 = \frac{\|C_{\mathrm{ad}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}}{\|C_{\mathrm{un}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}} Tol_0, \tag{4.15}$$

which, according to Lemma 4.2.1, yields

$$\|u(\cdot,\omega) - u_{(0)}(\cdot,\omega)\|_{H^1(D)} \leq C_{\mathrm{un}}(\omega) \frac{\|C_{\mathrm{ad}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}}{\|C_{\mathrm{un}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}} Tol_0. \tag{4.16}$$

Then, having an algorithm that fulfills Assumption 4.3.1 at hand, we can define a hierarchy of subspaces and corresponding solutions according to

$$\begin{aligned} S_0(\omega) &:= S_{(0)}, & u_0(\cdot,\omega) &:= u_{(0)}(\cdot,\omega), \\ S_l(\omega) &:= S_{(k_l(\omega))}(\omega), & u_l(\cdot,\omega) &:= u_{(k_l(\omega))}(\cdot,\omega), \quad l \in \mathbb{N}, \end{aligned} \tag{4.17}$$

for almost all $\omega \in \Omega$, where $k_l(\omega)$ is the smallest natural number such that

$$\eta_{\mathcal{T}_{(k_l(\omega))}(\omega)}(u_{(k_l(\omega))}(\cdot,\omega)) \leq Tol_l, \tag{4.18}$$

and Tol_l is chosen according to (3.2). Condition (4.18) together with (4.14) provides

$$\|u(\cdot,\omega) - u_l(\cdot,\omega)\|_{H^1(D)} \leq C_{\mathrm{ad}}(\omega) Tol_l. \tag{4.19}$$

In contrast to the uniform case, Assumption 4.3.2 is not provided by [73, Theorem 1] for $l \in \mathbb{N}$ as in Proposition 2.2.1 in the adaptive case, because solutions $u_l(\cdot,\omega)$ are sought in random spaces $S_l(\omega)$. Therefore, we shall assume that the solutions corresponding to the levels $l \in \mathbb{N}$ are measurable.

Assumption 4.3.2. For all $l \in \mathbb{N}$ the solution map $\Omega \ni \omega \mapsto u_l(\cdot,\omega) \in H^1_0(D)$ is measurable.

Using the arguments from Section 4.2, we also require that for almost all $\omega \in \Omega$ the approximations $\tilde{u}_l(\cdot,\omega)$ to $u_l(\cdot,\omega)$ fulfill (4.10), which again can be achieved by iterative solvers [20, 22] converging for almost all $\omega \in \Omega$.

As in the uniform case, measurability of $\Omega \ni \omega \mapsto \tilde{u}_l(\cdot,\omega) \in H^1_0(D)$ is inherited from u_l. Assumptions 3.1.1 and 3.2.1 follow then from (4.10), (4.16), (4.19) and the triangle inequality with $C_{\mathrm{dis}}(\omega) = \max\{C_{\mathrm{un}}(\omega)\frac{\|C_{\mathrm{ad}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}}{\|C_{\mathrm{un}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}}, C_{\mathrm{ad}}(\omega)\} + 1$.

We will leave Assumption 3.1.2 open for now, because, although multigrid methods can be applied for solving pathwise discretized problems with computational cost bounded by $C_{\mathrm{MG}}(\omega)|\mathcal{T}_{(k_l(\omega))}(\omega)|$, there is no obvious relation between Tol_l and $|\mathcal{T}_{(k_l(\omega))}(\omega)|$ and it has to be investigated further.

We now state convergence theorems for the adaptive MLMC-FE methods. The following theorem follows from Theorem 3.1.1 and the results presented in this section.

Theorem 4.3.1. *Let Assumption 2.2.1 hold and $u(\cdot,\omega)$, $\omega \in \Omega$ denote the solutions to problem* (4.1)*. Let Assumption 4.3.1 hold and $\tilde{u}_l(\cdot,\omega)$, $\omega \in \Omega$ denote the algebraic approximations to the finite element solutions defined in* (4.17)*, that fulfill* (4.10)*. Finally, let Assumptions 4.3.2 and 3.1.2 hold. Then for any $Tol > 0$ there exists an $L \in \mathbb{N}$ and a sequence of integers $\{M_l\}_{l=0}^L$ providing*

$$\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|_{L^2(\Omega;H^1_0(D))} \leq Tol,$$

and the estimator $E^L[\tilde{u}_L]$ *can be evaluated at expected computational cost*

$$\mathbb{E}[Cost(E^L[\tilde{u}_L])] = \begin{cases} \mathcal{O}(Tol^{-2}), & \gamma < 2, \\ \mathcal{O}(L^2 Tol^{-2}), & \gamma = 2, \\ \mathcal{O}(Tol^{-\gamma}), & \gamma > 2, \end{cases}$$

where the constants depend only on q, γ, C_{cost}, Tol_0, $\mathbb{E}[C_{\text{MG}}]$, $\|u\|_{L^2(\Omega;H_0^1(D))}$ *and* $\|C_{\text{ad}}\|_{L^2(\Omega)}$.

The following theorem follows from Theorem 3.2.1 and the results presented in this section.

Theorem 4.3.2. *Let Assumptions 2.2.1 hold and* $u(\cdot,\omega)$, $\omega \in \Omega$ *denote the solutions to problem* (4.1). *Let Assumption 2.4.1 hold,* $p_f \in (2\frac{p_Q}{p_Q-2}, \infty]$ *and Assumption 4.3.1 hold with* $p_{\text{ad}} = 2\frac{p_Q}{p_Q-2}$. *Further, let* $\tilde{u}_l(\cdot,\omega)$, $\omega \in \Omega$ *denote the algebraic approximations to the finite element solutions defined in* (4.17), *that fulfill* (4.10). *Finally, let Assumptions 4.3.2 and 3.1.2 hold. Then for any* $Tol > 0$ *there exists an* $L \in \mathbb{N}$ *and a sequence of integers* $\{M_l\}_{l=0}^L$, *such that*

$$e(E^L[Q(\tilde{u}_L)]) \leq Tol,$$

and the estimator $E^L[Q(\tilde{u}_L)]$ *can be evaluated at expected computational cost*

$$\mathbb{E}[Cost(E^L[Q(\tilde{u}_L)])] = \begin{cases} \mathcal{O}(Tol^{-2}), & \gamma < 2, \\ \mathcal{O}(L^2 Tol^{-2}), & \gamma = 2, \\ \mathcal{O}(Tol^{-\gamma}), & \gamma > 2, \end{cases}$$

where the constants depend only on q, γ, C_{cost}, Tol_0, $\mathbb{E}[C_{\text{MG}}]$, $\|u\|_{L^{p_{\text{ad}}}(\Omega;H_0^1(D))}$, $\|C_{\text{ad}}\|_{L^{p_{\text{ad}}}(\Omega)}$, $\|C_Q\|_{L^{p_Q}(\Omega)}$.

In the following subsections we describe adaptive algorithms based on a posteriori error estimation that can be used for adaptive multilevel Monte Carlo methods. We provide an example of such algorithms that in combination with multigrid methods ensures the validity of Assumptions 4.3.1, 4.3.2 and 3.1.2.

4.3.1 Adaptive algorithms based on pathwise a posteriori error estimation

In this section we present the concept of adaptive algorithms based on a posteriori error estimation. We closely follow [25] where a general framework for adaptive algorithms based on a set of axioms for the utilized error estimators is introduced. The following standard loop is the basis for a general adaptive algorithm.

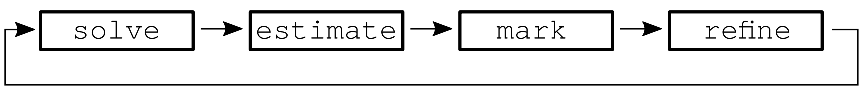

FIGURE 4.2: Standard loop for a general adaptive algorithm.

We consider an a posteriori error estimator that can be represented as

$$\eta_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)) = \Big(\sum_{T\in\mathcal{T}_{(k)}(\omega)} \eta_T(u_{(k)}(\cdot,\omega))^2 \Big)^{\frac{1}{2}}, \tag{4.20}$$

where

$$\eta_T(\cdot) : S_{(k)}(\omega) \to [0,\infty) \quad \text{for all} \quad T \in \mathcal{T}_{(k)}(\omega)$$

are local error contributions that can be used as refinement indicators.

If Assumption 4.3.1 holds, Algorithm 1 almost surely provides adaptive solutions $u_{(k)}(\cdot,\omega)$, such that $\|u(\cdot,\omega) - u_{(k)}(\cdot,\omega)\|_{H^1(D)} \leq C_{\text{ad}}(\omega)\varepsilon$ for any $\varepsilon > 0$.

Algorithm 1 Pathwise Adaptive Finite Element Algorithm

Input: $\mathcal{T}_{(0)}, \omega, \theta \in (0,1], \varepsilon$
Initialize $\mathcal{T}_{(0)}(\omega) := \mathcal{T}_{(0)}$
for $k = 0,1,\ldots$ **do**
  Compute finite element solution $u_{(k)}(\cdot,\omega)$
  Compute $\eta_T(u_{(k)}(\cdot,\omega))$ for all $T \in \mathcal{T}_{(k)}(\omega)$
  if $\eta_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)) \leq \varepsilon$ **then**
    stop algorithm
  end if
  Determine set $\mathcal{M}_{(k)}(\omega) \subset \mathcal{T}_{(k)}(\omega)$ of minimal cardinality such that

$$\theta\eta^2_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)) \leq \sum_{T\in\mathcal{M}_{(k)}(\omega)} \eta^2_T(u_{(k)}(\cdot,\omega)) \tag{4.21}$$

  Refine at least the elements $T \in \mathcal{M}_{(k)}(\omega)$ to generate $\mathcal{T}_{(k+1)}(\omega)$
end for
Output: $u_{(k)}(\cdot,\omega)$ or $Q(u_{(k)}(\cdot,\omega)), k = 0,1,\ldots,k_{\text{end}}$

Condition (4.21) is known as the Dörfler marking strategy [34] with parameter θ and it is the essential component of the part `mark` in the loop presented in Figure 4.2. The parameter θ may depend on ω if desired. We will discuss the parts `estimate` and `refine` in more detail in the next sections.

Remark 4.3.1. In practical realizations of Algorithm 1 the subsets $\mathcal{M}_{(k)}(\omega)$, $k = 0,1,\ldots$ may be required to be only of *almost optimal* cardinality, i.e. they should fulfill $|\mathcal{M}_{(k)}(\omega)| \leq C_{\min}|\mathcal{S}_{(k)}(\omega)|$, where $\mathcal{S}_{(k)}(\omega)$, $k = 0,1,\ldots$ are the sets of minimal cardinality. In this case we assume that $C_{\min} > 0$ is independent of ω and k for all $\omega \in \Omega$ and $k \in \mathbb{N}\cup\{0\}$. Determining a set of true minimal cardinality in Algorithm 1 in general requires sorting the local error contributions which results in log-linear complexity, whereas a set of almost minimal cardinality can be determined in linear complexity, see [82, Section 5].

4.3.2 Mesh refinement

For the step `refine` of the refinement loop we need to provide a procedure for constructing the mesh $\mathcal{T}_{(k+1)}$ from the current mesh $\mathcal{T}_{(k)}$. We omit the dependence of meshes on $\omega \in \Omega$ for brevity here. We restrict ourselves to the newest vertex bisection refinement technique [17, 26, 82, 83] for $d = 2, 3$ and the extended bisection for $d = 1$ presented in [8].

Let us now consider the set of admissible refinements of the initial partition $\mathcal{T}_{(0)}$

$$\mathbb{T} := \{\mathcal{T} : \quad \mathcal{T} \text{ is admissible refinement of } \mathcal{T}_{(0)}\},$$

obtained by the fixed types of refinement for different values of d

Each step `refine` consists of two substeps. The first substep produces the mesh $\mathcal{T}_{(k+\frac{1}{2})}$, where the subset $\mathcal{M}_{(k)} \subset \mathcal{T}_{(k)}$ is refined. The second step which is called *closure* aims at eliminating hanging nodes that appeared after the first substep, i.e. closure provides $\mathcal{T}_{(k+1)} \in \mathbb{T}$.

One of the important properties of the chosen mesh refinement types is that they allow the following closure estimate for partitions produced by Algorithm 1.

$$|\mathcal{T}_{(k)}| - |\mathcal{T}_{(0)}| \leq C_{\text{mesh}} \sum_{i=0}^{k-1} |\mathcal{M}_{(i)}|, \tag{4.22}$$

where the constant $C_{\text{mesh}} > 0$ depends only on $\mathcal{T}_{(0)}$. Since $\mathcal{T}_{(0)}$ is chosen independently of ω in Algorithm 1, the constant C_{mesh} does not depend on ω either. This property means that the cumulative number of elements added by closure is controlled by the total number of marked elements. This result has been shown for $d = 1$ in [8] and for $d = 2$ in [57]. In the case $d = 3$ this result can be found in [83] under the assumption of appropriate labelling of faces of the initial partition $\mathcal{T}_{(0)}$. It is, however, not clear whether such a labelling exists for an arbitrary initial mesh. Property (4.22) plays an important role in convergence proofs for adaptive algorithms. This result also holds when $b > 1$ number of refinements are applied to each marked element in one refinement step, see [26].

The newest vertex bisection technique in the cases $d = 2, 3$ guarantees that all elements that arise during refinement can be classified into a finite set of similarity classes that only depend on the initial partition $\mathcal{T}_{(0)}$, see [83]. This implies that the class of partitions $\mathbb{T}$ is shape-regular, i.e.

$$C_{\mathcal{T}} \leq \Gamma, \quad \forall \mathcal{T} \in \mathbb{T},$$

with a fixed $\Gamma > 0$ that depends only on $\mathcal{T}_{(0)}$. An analogous result can be found for the extended bisection technique for $d = 1$ in [8].

4.3.3 Pathwise residual-based a posteriori error estimation

In this section we introduce residual-based a posteriori error estimation [5, 17, 26, 62, 72] adapted to the random problem (4.1). Following the framework developed in [25], we show a set of properties of the residual error estimators that provide sufficient conditions for convergence of the adaptive Algorithm 1 based on this type of error estimation when the mesh refinement introduced in Section 4.3.2 is utilized. Furthermore, these properties allow one to derive the order of convergence of the algorithm; this is done in Section 4.3.5.

All results presented in this section are generalizations of existing results for residual-based error estimation for deterministic linear elliptic problems to random problems of the form (4.1).

In this section we denote the first order finite element space defined on a mesh $\mathcal{T} \in \mathbb{T}$ as in (4.3) by $S_{\mathcal{T}}$ and consider the finite element solution $u_{\mathcal{T}}(\cdot, \omega)$ of the problem

$$u_{\mathcal{T}}(\cdot,\omega) \in S_{\mathcal{T}} : \qquad a(\omega; u_{\mathcal{T}}(\cdot,\omega), v) = \ell(\omega; v), \qquad \forall v \in S_{\mathcal{T}},\ \omega \in \Omega, \tag{4.23}$$

with the bilinear form a and the linear form ℓ as in (2.5).

Let us introduce the so-called *energy norm* $\|\cdot\|_a$ induced by the bilinear form $a(\omega;\cdot,\cdot)$ defined in (2.5), given by

$$\|v\|_a := a(\omega; v, v)^{\frac{1}{2}}$$

for a function $v \in H_0^1(D)$.

The energy norm is a common norm used in the literature on finite element methods for elliptic PDEs and it is equivalent to the H^1-norm on $H_0^1(D)$. Pathwise ellipticity of the bilinear form a implies this property in the case when the energy norm satisfies upper and lower bounds with random coefficients, i.e. when

$$c_a^{1/2}(\omega)\|v\|_{H^1(D)} \le \|v\|_a \le \alpha_{\max}^{1/2}(\omega)\|v\|_{H^1(D)}, \quad u \in H_0^1(D) \tag{4.24}$$

holds for all $\omega \in \Omega$.

Recall that according to Assumptions 2.2.1 (ii, iii), for almost all $\omega \in \Omega$ the pathwise realizations $\alpha(\cdot,\omega)$ are Lipschitz continuous and $f(\cdot,\omega) \in L^2(D)$.

We start by introducing basic results that help to derive the residual error estimator. These results are well-known and we cite [5] and [90] for the proofs in the case $\alpha(x,\omega) = 1$ in $D \times \Omega$. We provide, however, the more general result here for completeness.

Proposition 4.3.1. *Let $\mathcal{T} \in \mathbb{T}$. Then almost surely, it holds that*

$$\begin{aligned}\frac{c_a^{1/2}(\omega)}{\alpha_{\max}(\omega)} &\sup_{\substack{w \in H_0^1(D) \\ \|w\|_{H^1(D)}=1}} \Big(\int_D f(x,\omega) w \, dx - \int_D \alpha(x,\omega) \nabla u_{\mathcal{T}}(x,\omega) \cdot \nabla w \, dx \Big) \\ &\leq \|u(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|_a \\ &\leq \frac{1}{c_a^{1/2}(\omega)} \sup_{\substack{w \in H_0^1(D) \\ \|w\|_{H^1(D)}=1}} \Big(\int_D f(x,\omega) w \, dx - \int_D \alpha(x,\omega) \nabla u_{\mathcal{T}}(x,\omega) \cdot \nabla w \, dx \Big).\end{aligned}$$

Proof. Let $v \in H_0^1(D)$. Then by (4.24) we have

$$\begin{aligned}\|v\|_a^2 = \int_D \alpha(x,\omega) \nabla v \cdot \nabla v \, dx &\leq \sup_{\substack{w \in H_0^1(D) \\ \|w\|_{H^1(D)}=1}} \int_D \alpha(x,\omega) \nabla v \cdot \nabla w \, dx \, \|v\|_{H^1(D)} \\ &\leq \frac{1}{c_a^{1/2}(\omega)} \sup_{\substack{w \in H_0^1(D) \\ \|w\|_{H^1(D)}=1}} \int_D \alpha(x,\omega) \nabla v \cdot \nabla w \, dx \, \|v\|_a.\end{aligned}$$

Continuity of $a(\omega; \cdot, \cdot)$ and (4.24) provide

$$\begin{aligned}\sup_{\substack{w \in H_0^1(D) \\ \|w\|_{H^1(D)}=1}} \int_D \alpha(x,\omega) \nabla v \cdot \nabla w \, dx &\leq \alpha_{\max}(\omega) \sup_{\substack{w \in H_0^1(D) \\ \|w\|_{H^1(D)}=1}} \|v\|_{H^1(D)} \|w\|_{H^1(D)} \\ &= \alpha_{\max}(\omega) \|v\|_{H^1(D)} \leq \frac{\alpha_{\max}(\omega)}{c_a^{1/2}(\omega)} \|v\|_a.\end{aligned}$$

Now, setting $v := u(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)$ and using the identity

$$\begin{aligned}&\int_D \alpha(x,\omega) \nabla (u(x,\omega) - u_{\mathcal{T}}(x,\omega)) \cdot \nabla w \, dx \\ &\qquad = \int_D f(x,\omega) w \, dx - \int_D \alpha(x,\omega) \nabla u_{\mathcal{T}}(x,\omega) \cdot \nabla w \, dx, \quad \forall w \in H_0^1(D),\end{aligned}$$

we obtain the result of the proposition. □

According to Proposition 4.3.1, the energy norm of the error is, up to ω-dependent multiplicative constants, bounded from above and below by the norm of the residual in the dual space of $H_0^1(D)$.

We denote the outer unit normal to the boundary ∂T of an element $T \in \mathcal{T}$ by $\mathbf{n}_T$. We also assign a unit normal of arbitrary orientation to every face $E \in \mathcal{E}$ and denote it by $\mathbf{n}_E$. Let us now introduce the following L^2-representation of the residual. For almost all $\omega \in \Omega$ and any $w \in H_0^1(D)$ there holds

$$\begin{aligned}
\int_D f(x,\omega)w\,dx &- \int_D \alpha(x,\omega)\nabla u_{\mathcal{T}}(x,\omega)\cdot\nabla w\,dx \\
&= \int_D f(x,\omega)w\,dx - \sum_{T\in\mathcal{T}}\int_T \alpha(x,\omega)\nabla u_{\mathcal{T}}(x,\omega)\cdot\nabla w\,dx \\
&= \int_D f(x,\omega)w\,dx - \sum_{T\in\mathcal{T}}\Big\{\int_{\partial T}\alpha(x,\omega)\nabla u_{\mathcal{T}}(x,\omega)\cdot\mathbf{n}_T w\,dx \\
&\qquad - \int_T \nabla\cdot(\alpha(x,\omega)\nabla u_{\mathcal{T}}(x,\omega))w\,dx\Big\} \qquad (4.25)\\
&= \sum_{T\in\mathcal{T}}\int_T \big(f(x,\omega)+\nabla\cdot(\alpha(x,\omega)\nabla u_{\mathcal{T}}(x,\omega))\big)w\,dx \\
&\qquad - \sum_{E\in\mathcal{E}}\int_E \mathbb{J}_E(\alpha(x,\omega)\nabla u_{\mathcal{T}}(x,\omega))\cdot\mathbf{n}_E w\,dx \\
&= \sum_{T\in\mathcal{T}}\int_T R_T(\omega;u_{\mathcal{T}}(x,\omega))w\,dx + \sum_{E\in\mathcal{E}}\int_E R_E(\omega;u_{\mathcal{T}}(x,\omega))w\,dx,
\end{aligned}$$

where

$$R_T(\omega;v) := (f(\cdot,\omega)+\nabla\cdot(\alpha(\cdot,\omega)\nabla v))|_T, \quad T\in\mathcal{T}, \tag{4.26}$$

$$R_E(\omega;v) := -(\mathbb{J}_E(\alpha(\cdot,\omega)\nabla v)\cdot\mathbf{n}_E)|_E, \quad E\in\mathcal{E}, \tag{4.27}$$

are the element and the jump residuals for a function $v \in S_{\mathcal{T}}$ respectively. Here $\mathbb{J}_E(\cdot)$ stands for the jump across the interior face E in direction of the unit normal $\mathbf{n}_E$ associated to this face, defined as

$$\mathbb{J}_E(v)(x) := \lim_{t\to 0+} v(x - t\mathbf{n}_E) - \lim_{t\to 0+} v(x + t\mathbf{n}_E), \quad x\in E.$$

Note that the expression $\mathbb{J}_E(\cdot)\cdot\mathbf{n}_E$ does not depend on the orientation of $\mathbf{n}_E$.

Now, we define a residual-based error estimator that takes the form (4.20) with local contributions given by

$$\eta_T^2(\omega;v) = h_T^2\|R_T(\omega;v)\|^2_{L^2(T)} + h_T\|R_E(\omega;v)\|^2_{L^2(\partial T\cap D)}. \tag{4.28}$$

Note that $\eta_T(\omega;\cdot)$, $T\in\mathcal{T}$ depend on ω because of the random functions α and f that enter the definition.

Monotonicity of the local mesh sizes provides monotonicity of the error estimator, i.e.

$$\eta_{\tilde{\mathcal{T}}}(\omega;v) \le \eta_{\mathcal{T}}(\omega;v), \quad \omega\in\Omega, \tag{4.29}$$

for any $v\in S_{\mathcal{T}}$, where $\tilde{\mathcal{T}}\in\mathbb{T}$ is a refinement of $\mathcal{T}$.

The following result provides *reliability* of the error estimator. We again cite [5] and [90] for proofs for the deterministic Poisson problem.

Lemma 4.3.1. *Let $\mathcal{T} \in \mathbb{T}$, then there holds*

$$\|u(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|_{H^1(D)} \leq c_{\mathrm{ad}} \frac{1}{\alpha_{\min}(\omega)} \eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)) \tag{4.30}$$

for almost all $\omega \in \Omega$, where $c_{\mathrm{ad}} > 0$ depends only on Γ, D and the dimension d.

Proof. We omit the dependence of u and $u_{\mathcal{T}}$ on their arguments here for brevity. We fix a $w \in H_0^1(D)$ and let $I_{\mathcal{T}} : H_0^1(D) \to S_{\mathcal{T}}$ be the quasi-interpolation operator defined in (A.1). Then, using Galerkin orthogonality [49, Relation (8.2.3)], the Cauchy-Schwarz inequalities for integrals and sums and the properties of the quasi-interpolation operator (A.2)-(A.3) we have:

$$\begin{aligned}
&\int_D f(x,\omega) w \, dx - \int_D \alpha(x,\omega) \nabla u_{\mathcal{T}} \cdot \nabla w \, dx \\
&\quad = \int_D f(x,\omega)(w - I_{\mathcal{T}} w) \, dx - \int_D \alpha(x,\omega) \nabla u_{\mathcal{T}} \cdot \nabla (w - I_{\mathcal{T}} w) \, dx \\
&\quad = \sum_{T \in \mathcal{T}} \int_T R_T(\omega; u_{\mathcal{T}})(w - I_{\mathcal{T}} w) \, dx + \sum_{E \in \mathcal{E}} \int_E R_E(\omega; u_{\mathcal{T}})(w - I_{\mathcal{T}} w) \, dx \\
&\quad \leq \sum_{T \in \mathcal{T}} \|R_T(\omega; u_{\mathcal{T}})\|_{L^2(T)} \|w - I_{\mathcal{T}} w\|_{L^2(T)} + \sum_{E \in \mathcal{E}} \|R_E(\omega; u_{\mathcal{T}})\|_{L^2(E)} \|w - I_{\mathcal{T}} w\|_{L^2(E)} \\
&\quad \lesssim \sum_{T \in \mathcal{T}} \|R_T(\omega; u_{\mathcal{T}})\|_{L^2(T)} h_T \|w\|_{H^1(\phi_T)} + \sum_{E \in \mathcal{E}} \|R_E(\omega; u_{\mathcal{T}})\|_{L^2(E)} h_T^{\frac{1}{2}} \|w\|_{H^1(\phi_T)} \\
&\quad \leq \Big\{ \sum_{T \in \mathcal{T}} h_T^2 \|R_T(\omega; u_{\mathcal{T}})\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}} h_T \|R_E(\omega; u_{\mathcal{T}})\|_{L^2(E)}^2 \Big\}^{\frac{1}{2}} \\
&\qquad\qquad \cdot \Big\{ \sum_{T \in \mathcal{T}} \|w\|_{H^1(\phi_T)}^2 + \sum_{E \in \mathcal{E}} \|w\|_{H^1(\phi_T)}^2 \Big\}^{\frac{1}{2}},
\end{aligned}$$

where the hidden constant depends only on Γ.

Observe that the shape-regularity of $\mathcal{T}$ implies

$$\Big\{ \sum_{T \in \mathcal{T}} \|w\|_{H^1(\phi_T)}^2 + \sum_{E \in \mathcal{E}} \|w\|_{H^1(\phi_T}^2 \Big\}^{\frac{1}{2}} \lesssim \|w\|_{H^1(D)},$$

where the hidden constant depends only on Γ and the dimension d and takes into account that every element is counted several times.

Combining these estimates with the equivalence of the error in the energy norm and the residual stated in Proposition 4.3.1, we obtain

$$\|u(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|_a \lesssim \frac{1}{\alpha_{\min}^{1/2}(\omega)} \eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)), \tag{4.31}$$

which together with (4.24) yields the claim of the lemma. □

Lemma 4.3.1 together with Assumption 2.2.1 (i) provides the first requirement of Assumption 4.3.1, since $\frac{1}{\alpha_{\min}} \in L^p(\Omega)$ for all $p \in [1, \infty)$.

We are also interested in a stronger result than provided by Lemma 4.3.1. The following lemma states the so-called *discrete reliability* of the error estimator.

Lemma 4.3.2. *For all refinements $\tilde{\mathcal{T}} \in \mathbb{T}$ of a partition $\mathcal{T} \in \mathbb{T}$ the localized upper bound*

$$\|u_{\tilde{\mathcal{T}}}(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|^2_{H^1(D)} \leq \frac{C^2_{\mathrm{rel}}}{\alpha^2_{\min}(\omega)} \sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}} \eta^2_T(\omega; u_{\mathcal{T}}(\cdot,\omega)) \tag{4.32}$$

holds for almost all $\omega \in \Omega$, where $C_{\mathrm{rel}} > 0$ depends only on Γ, D and the dimension d.

Proof. The claim of the lemma is a direct consequence of [26, Lemma 3.6]. According to the proof of this result, we have

$$a(e_{\tilde{\mathcal{T}}}(\cdot,\omega), e_{\tilde{\mathcal{T}}}(\cdot,\omega)) \leq C_{\mathrm{rel}} \left(\sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}} \eta^2_T(\omega; u_{\mathcal{T}}(\cdot,\omega)) \right)^{\frac{1}{2}} \|\nabla e_{\tilde{\mathcal{T}}}(\cdot,\omega)\|_{L^2(\mathcal{T}\setminus\tilde{\mathcal{T}})},$$

where the constant C_{rel} depends only on Γ and we denote $e_{\tilde{\mathcal{T}}}(\cdot,\omega) := u_{\tilde{\mathcal{T}}}(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)$. Together with coercivity of the bilinear form $a(\omega;\cdot,\cdot)$, it implies the claim of the lemma. □

The first and second parts of the following lemma provide *stability* of the error estimator on non-refined elements and a *reduction* property on refined elements respectively. Both of these properties are important for the convergence of the adaptive algorithm.

Lemma 4.3.3. *Let $\tilde{\mathcal{T}} \in \mathbb{T}$ be a refinement of a partition $\mathcal{T} \in \mathbb{T}$, then for almost all $\omega \in \Omega$*

(i) for all subsets $\mathcal{S} \subseteq \mathcal{T} \cap \tilde{\mathcal{T}}$ of non-refined element domains it holds

$$\begin{aligned}\sum_{T\in\mathcal{S}} \eta^2_{T\in\tilde{\mathcal{T}}}(\omega; u_{\tilde{\mathcal{T}}}(\cdot,\omega)) - (1+\delta) \sum_{T\in\mathcal{S}} \eta^2_{T\in\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega))& \\ \leq (1+\delta^{-1}) C^2_{\mathrm{stab}} \|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)} \|u_{\tilde{\mathcal{T}}}(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|^2_{H^1(D)},&\end{aligned} \tag{4.33}$$

(ii) for the refined element domain it holds

$$\begin{aligned}\sum_{T\in\tilde{\mathcal{T}}\setminus\mathcal{T}} \eta^2_T(\omega; u_{\tilde{\mathcal{T}}}(\cdot,\omega)) - (1+\delta)\rho_{\mathrm{red}} \sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}} \eta^2_T(\omega; u_{\mathcal{T}}(\cdot,\omega))& \\ \leq (1+\delta^{-1}) C^2_{\mathrm{stab}} \|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)} \|u_{\tilde{\mathcal{T}}}(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|^2_{H^1(D)},&\end{aligned} \tag{4.34}$$

where $\delta > 0$ is arbitrary, $0 < \rho_{\mathrm{red}} < 1$ and the constant $C_{\mathrm{stab}} > 0$ depends only on Γ , h_0 and the dimension d.

Proof. The proof is based on the proofs of corresponding deterministic results in [26, Proposition 3.3, Corollary 3.4].

First, we show that for any $T \in \mathcal{T}$, $\mathcal{T} \in \mathbb{T}$ and for any two discrete functions $v, w \in S_{\mathcal{T}}$ there holds

$$\eta_T^2(\omega; v) - (1+\delta)\eta_T^2(\omega; w) \le (1+\delta^{-1})\Lambda \|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)} \|v-w\|^2_{H^1(\phi_T)}, \tag{4.35}$$

where $\delta > 0$, and $\Lambda > 0$ is a constant that depends only on Γ and h_0.

Using the triangle and Young's inequalities we obtain

$$\begin{aligned} &\eta_T^2(\omega; v) \le (1+\delta)\eta_T^2(\omega; w) \\ &+ (1+\delta^{-1}) \left(h_T^2 \|\nabla \cdot \big(\alpha(\cdot,\omega)\nabla(v-w)\big)\|^2_{L^2(T)} + h_T \|R_E(\omega; v-w)\|^2_{L^2(\partial T \cap D)} \right) \end{aligned} \tag{4.36}$$

with any $\delta > 0$.

Now, we show the bounds for the second term on the right hand side in (4.36). Since $v - w \in S_{\mathcal{T}}$, we have

$$\begin{aligned} \|\nabla \cdot \big(\alpha(\cdot,\omega)\nabla(v-w)\big)\|_{L^2(T)} &= \|\nabla\alpha(\cdot,\omega) \cdot \nabla(v-w)\|_{L^2(T)} \\ &\le \|\nabla\alpha(\cdot,\omega)\|_{L^\infty(T)} \|\nabla(v-w)\|_{L^2(T)}. \end{aligned} \tag{4.37}$$

Let T' be a neighbour of T with a common face E, then using the inverse inequality (A.12), we obtain

$$\begin{aligned} \|R_E(\omega; v-w)\|_{L^2(E)} &= \|(\alpha(\cdot,\omega)\nabla(v-w)|_T - \alpha(\cdot,\omega)\nabla(v-w)|_{T'}) \cdot n_E\|_{L^2(E)} \\ &\lesssim \|\alpha(\cdot,\omega)\|_{L^\infty(T)} \left(h_T^{-\frac{1}{2}} \|\nabla(v-w)\|_{L^2(T)} + h_{T'}^{-\frac{1}{2}} \|\nabla(v-w)\|_{L^2(T')} \right) \end{aligned} \tag{4.38}$$

with the hidden constant dependent only on Γ.

Substituting (4.37) and (4.38) into (4.36) we obtain (4.35).

Now, we apply (4.35) to functions $u_{\tilde{\mathcal{T}}}$ and $u_{\mathcal{T}}$ (we again skip the arguments of the functions for brevity) over $T \in \tilde{\mathcal{T}} \cap \mathcal{T}$, sum over a subset $\mathcal{S}$ of non-refined elements and use the finite overlap property of patches ϕ_T. Then we obtain

$$\begin{aligned} &\sum_{T \in \mathcal{S}} \eta^2_{T \in \tilde{\mathcal{T}}}(\omega; u_{\tilde{\mathcal{T}}}) - (1+\delta) \sum_{T \in \mathcal{S}} \eta^2_{T \in \tilde{\mathcal{T}}}(\omega; u_{\mathcal{T}}) \\ &\qquad \le (1+\delta^{-1}) C^2_{\text{stab}} \|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)} \|u_{\tilde{\mathcal{T}}} - u_{\mathcal{T}}\|^2_{H^1(D)}, \end{aligned} \tag{4.39}$$

where the constant $C_{\text{stab}} > 0$ depends on Λ and the dimension d. Recall that

$$\eta_{T \in \tilde{\mathcal{T}}}(\omega; u_{\mathcal{T}}) = \eta_{T \in \mathcal{T}}(\omega; u_{\mathcal{T}}),$$

for any $T \in \mathcal{S}$. Incorporating this relation into (4.39) we obtain (i). Note that in a similar way one can obtain a relation symmetric to (i), i.e.

$$\sum_{T\in\mathcal{S}} \eta^2_{T\in\mathcal{T}}(\omega; u_{\mathcal{T}}) - (1+\delta)\sum_{T\in\mathcal{S}} \eta^2_{T\in\tilde{\mathcal{T}}}(\omega; u_{\tilde{\mathcal{T}}}) \\ \leq (1+\delta^{-1})C^2_{\text{stab}}\|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)}\|u_{\tilde{\mathcal{T}}} - u_{\mathcal{T}}\|^2_{H^1(D)}, \quad (4.40)$$

which holds for all $\mathcal{S} \in \mathcal{T} \cap \tilde{\mathcal{T}}$.

Analogously we obtain

$$\sum_{T\in\tilde{\mathcal{T}}\setminus\mathcal{T}} \eta^2_{T}(\omega; u_{\tilde{\mathcal{T}}}) - (1+\delta)\sum_{T\in\tilde{\mathcal{T}}\setminus\mathcal{T}} \eta^2_{T}(\omega; u_{\mathcal{T}}) \\ \lesssim (1+\delta^{-1})\|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)}\|u_{\tilde{\mathcal{T}}} - u_{\mathcal{T}}\|^2_{H^1(D)} \quad (4.41)$$

with the hidden constant that depends only on Λ and the dimension d. For an element $T \in \mathcal{T} \setminus \tilde{\mathcal{T}}$ we define the set $\tilde{\mathcal{T}}_T := \{T' \in \tilde{\mathcal{T}} : T' \subset T\}$. Since $u_{\mathcal{T}} \in S_{\mathcal{T}}$, $R_E(u_{\mathcal{T}}) = 0$ in the interior of T. Therefore,

$$\sum_{T'\in\tilde{\mathcal{T}}_T} \eta^2_{T'}(\omega; u_{\mathcal{T}}) \leq 2^{-\frac{b}{d}}\eta^2_T(\omega; u_{\mathcal{T}}), \quad (4.42)$$

because refinement by bisection provides $h_{T'} = 2^{-\frac{b}{d}}h_T$.

Finally, (4.41) and (4.42) imply (ii), where $\rho_{\text{red}} = 2^{-\frac{b}{d}}$. □

Lemma 4.3.4. *Let $u_{(k)}(\cdot,\omega)$, $k \in \mathbb{N}\cup\{0\}$, $\omega \in \Omega$ be the solutions obtained by Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28), *then there holds*

$$\sum_{i=k}^{N} \|u_{(i+1)}(\cdot,\omega) - u_{(i)}(\cdot,\omega)\|^2_{H^1(D)} \leq \frac{C_{\text{qo}}}{\alpha^2_{\min}(\omega)}\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)), \quad (4.43)$$

for any $k, N \in \mathbb{N}\cup\{0\}$, $N > k$, where the constant $C_{\text{qo}} > 0$ depends only on Γ, D and the dimension d.

Proof. For any $i \in \mathbb{N}\cup\{0\}$ and almost all $\omega \in \Omega$, Galerkin orthogonality [49, Relation (8.2.3)] provides

$$\|u(\cdot,\omega) - u_{(i+1)}(\cdot,\omega)\|^2_a = \|u(\cdot,\omega) - u_{(i)}(\cdot,\omega)\|^2_a - \|u_{(i+1)}(\cdot,\omega) - u_{(i)}(\cdot,\omega)\|^2_a.$$

Then, utilizing (4.24) and (4.31) we have

$$
\begin{aligned}
c_a(\omega)\sum_{i=k}^{N}\|u_{(i+1)}(\cdot,\omega)-u_{(i)}(\cdot,\omega)\|^2_{H^1(D)} &\le \sum_{i=k}^{N}\|u_{(i+1)}(\cdot,\omega)-u_{(i)}(\cdot,\omega)\|^2_a \\
&= \sum_{i=k}^{N}\Big(\|u(\cdot,\omega)-u_{(i)}(\cdot,\omega)\|^2_a \\
&\qquad\qquad - \|u(\cdot,\omega)-u_{(i+1)}(\cdot,\omega)\|^2_a\Big) \\
&\le \|u(\cdot,\omega)-u_{(k)}(\cdot,\omega)\|^2_a \\
&\lesssim \frac{1}{\alpha_{\min}(\omega)}\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega;u_{(k)}(\cdot,\omega)),
\end{aligned}
$$

which concludes the proof. □

Lemmas 4.3.1-4.3.4 are the analogues of the deterministic conditions for a posteriori error estimators presented as axioms in [25]. According to [25], the properties of an error estimator stated in the axioms, together with the mesh refinement described in section 4.3.2, are sufficient for convergence of the adaptive Algorithm 1. Therefore, we are finally ready to state the convergence theorem.

Theorem 4.3.3. *For all* $0<\theta\le 1$*, Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28) *and mesh refinement described in Section 4.3.2 converges almost surely in the sense that there exist* $0<\rho_{\mathrm{conv}}(\omega)<1$ *and* $C_{\mathrm{conv}}(\omega)>1$ *such that*

$$
\eta^2_{\mathcal{T}_{(k+j)}(\omega)}(\omega;u_{(k+j)}(\cdot,\omega)) \le C_{\mathrm{conv}}(\omega)\rho^j_{\mathrm{conv}}(\omega)\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega;u_{(k)}(\cdot,\omega)), \quad \forall k,j\in\mathbb{N}\cup\{0\}
$$

for almost all $\omega\in\Omega$*, where* $\rho_{\mathrm{conv}}(\omega)$ *and* $C_{\mathrm{conv}}(\omega)$ *depend only on* Γ*,* h_0*,* d*,* b*,* D*,* θ*,* $\alpha_{\min}(\omega)$ *and* $\|\alpha(\cdot,\omega)\|_{W^{1,\infty}(D)}$.

Proof. Lemmas 4.3.1, 4.3.3 and 4.3.4 provide the sufficient conditions for [25, Theorem 4.1 (i)] which states the convergence. □

Theorem 4.3.3 yields the second requirement of Assumption 4.3.1, i.e. $\eta_{\mathcal{T}_{(k)}(\omega)}(\omega;u_{(k)}(\cdot,\omega)) \xrightarrow{k\to\infty} 0$, which together with Lemma 4.3.1 provides the following result.

Theorem 4.3.4. *For all* $0<\theta\le 1$*, Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28) *and mesh refinement described in Section 4.3.2 fulfills Assumption 4.3.1 with* $C_{\mathrm{ad}}=c_{\mathrm{ad}}\frac{1}{\alpha_{\min}}$ *and all* $p_{\mathrm{ad}}\in[2,p_f)$*, where* c_{ad} *depends only on* Γ*,* D *and* d*.*

Proof. Theorem 4.3.3, Lemma 4.3.1 and Assumption 2.2.1 (i) provide the result of the theorem. □

Following [5, 17] we now introduce *oscillations* of a function $v\in S_{\mathcal{T}}$ and list some of their properties that will be required for showing the order of convergence of Algorithm 1. Let π^p_m denote the operator of L^p-best approximation onto the set of polynomials of order $\leq m$ over a domain that will be clear from the context. Then we can define oscillations of $v \in S_{\mathcal{T}}$ by

$$osc_T^2(\omega; v) := h_T^2 \|(\mathrm{id} - \pi_m^2)(R_T(\omega; v))\|_{L^2(T)}^2 + h_T \|(\mathrm{id} - \pi_n^2)(R_E(\omega; v))\|_{L^2(\partial T \cap D)}^2,$$

$$osc_{\mathcal{T}}^2(\omega; v) := \sum_{T \in \mathcal{T}} osc_T^2(\omega; v),$$

where $m, n \in \mathbb{N} \cup \{0\}$ are fixed approximation orders. As well as the error indicators $\eta_{\mathcal{T}}(\omega; \cdot)$, $\mathcal{T} \in \mathbb{T}$, the oscillation terms $osc_{\mathcal{T}}(\omega; \cdot)$, $\mathcal{T} \in \mathbb{T}$ depend on ω through the random functions α and f. Observe that

$$osc_T(\omega; v) \leq \eta_T(\omega; v) \tag{4.44}$$

for all $T \in \mathcal{T}$, $\mathcal{T} \in \mathbb{T}$, almost all $\omega \in \Omega$ and any $v \in S_{\mathcal{T}}$.

Now we are ready to introduce the lower bound, or *efficiency*, for the H^1-error. We cite [5, 90] for proofs of the bound for the deterministic Poisson problem that serve as basis for the proof of the following lemma.

Lemma 4.3.5. *For any partition $\mathcal{T} \in \mathbb{T}$ and almost all $\omega \in \Omega$ there holds*

$$\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)) \lesssim \alpha_{\max}(\omega) \|u(\cdot, \omega) - u_{\mathcal{T}}(\cdot, \omega)\|_{H^1(D)} + osc_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)), \tag{4.45}$$

where the hidden constant depends only on Γ, the dimension d and the approximation orders m, n.

Proof. In this proof we again omit the arguments of the functions u and $u_{\mathcal{T}}$ for brevity. We fix an arbitrary element $T \in \mathcal{T}$ and insert the function $w_T := \pi_m^2(R_T(\omega; u_{\mathcal{T}}))\psi_T$, where ψ_T is the element bubble function defined in (A.4) with supp $w_T \subset T$, into the L^2- representation of the residual (4.25) as a test function. Then we have

$$\int_T R_T(\omega; u_{\mathcal{T}}) w_T \, dx = \int_T \alpha(x, \omega) \nabla(u - u_{\mathcal{T}}) \cdot \nabla w_T \, dx.$$

We add $\int_T (\pi_m^2 - \mathrm{id})(R_T(\omega; u_{\mathcal{T}})) w_T \, dx$ on both sides and obtain

$$\begin{aligned}\int_T \left(\pi_m^2(R_T(\omega; u_{\mathcal{T}}))\right)^2 \psi_T \, dx = &\int_T \alpha(x, \omega) \nabla(u - u_{\mathcal{T}}) \cdot \nabla w_T \, dx \\ &- \int_T (\mathrm{id} - \pi_m^2)(R_T(\omega; u_{\mathcal{T}})) w_T \, dx.\end{aligned} \tag{4.46}$$

Properties (A.6)-(A.7) of the bubble function ψ_T and the Cauchy-Schwarz inequality imply

$$\int_T \left(\pi_m^2(R_T(\omega; u_{\mathcal{T}}))\right)^2 \psi_T \, dx \gtrsim \|\pi_m^2(R_T(\omega; u_{\mathcal{T}}))\|_{L^2(T)}^2, \tag{4.47}$$

$$\begin{aligned}\int_T \alpha(x,\omega)\nabla(u-u_{\mathcal{T}})\cdot\nabla w_T\,dx &\leq \|\alpha(\cdot,\omega)\nabla(u-u_{\mathcal{T}})\|_{L^2(T)}\|\nabla w_T\|_{L^2(T)}\\ &\lesssim \alpha_{\max}(\omega)h_T^{-1}\|u-u_{\mathcal{T}}\|_{H^1(T)}\|\pi_m^2(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)},\end{aligned} \tag{4.48}$$

$$\begin{aligned}\Big|\int_T(\mathrm{id}-\pi_m^2)(R_T(\omega;u_{\mathcal{T}}))w_T\,dx\Big| &\leq \|(\mathrm{id}-\pi_m^2)(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)}\|w_T\|_{L^2(T)}\\ &\leq \|(\mathrm{id}-\pi_m^2)(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)}\|\pi_m^2(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)},\end{aligned} \tag{4.49}$$

where the hidden constants depend only on the shape-regularity Γ and the order m.

Relations (4.46)-(4.49) provide

$$h_T\|\pi_m^2(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)} \lesssim \alpha_{\max}(\omega)\|u-u_{\mathcal{T}}\|_{H^1(T)}+h_T\|(\mathrm{id}-\pi_m^2)(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)}. \tag{4.50}$$

Now, we fix an arbitrary face $E\in\mathcal{E}$ and insert $w_E := \pi_n^2(R_E(\omega;u_{\mathcal{T}}))\psi_E$, where ψ_E is the face bubble function defined in (A.5), into the L^2- representation of the residual (4.25) as a test function, which gives

$$\begin{aligned}\int_E\big(\pi_n^2(R_E(\omega;u_{\mathcal{T}}))\big)^2\psi_E\,dx &= \int_{\phi_E}\alpha(x,\omega)\nabla(u-u_{\mathcal{T}})\cdot\nabla w_E\,dx\\ &\quad-\sum_{T\in\phi_E}\int_T\pi_m^2(R_T(\omega;u_{\mathcal{T}}))w_E\,dx\\ &\quad-\sum_{T\in\phi_E}\int_T(\mathrm{id}-\pi_m^2)(R_T(\omega;u_{\mathcal{T}}))w_E\,dx\\ &\quad-\int_E(\mathrm{id}-\pi_n^2)(R_E(\omega;u_{\mathcal{T}}))w_E\,dx.\end{aligned} \tag{4.51}$$

Properties (A.8)-(A.10) of the face bubble function ψ_E and the Cauchy-Schwarz inequality imply

$$\int_E\big(\pi_n^2(R_E(\omega;u_{\mathcal{T}}))\big)^2\psi_E\,dx \gtrsim \|\pi_n^2(R_E(\omega;u_{\mathcal{T}}))\|^2_{L^2(E)}, \tag{4.52}$$

$$\begin{aligned}&\int_{\phi_E}\alpha(x,\omega)\nabla(u-u_{\mathcal{T}})\cdot\nabla w_E\,dx\\ &\qquad\lesssim \alpha_{\max}(\omega)h_E^{-\frac{1}{2}}\|u-u_{\mathcal{T}}\|_{H^1(\phi_E)}\|\pi_n^2(R_E(\omega;u_{\mathcal{T}}))\|_{L^2(E)},\end{aligned} \tag{4.53}$$

$$\begin{aligned}&\Big|\sum_{T\in\phi_E}\int_T\pi_m^2(R_T(\omega;u_{\mathcal{T}}))w_E\,dx\Big|\\ &\qquad\lesssim h_E^{\frac{1}{2}}\|\pi_n^2(R_E(\omega;u_{\mathcal{T}}))\|_{L^2(E)}\sum_{T\in\phi_E}\|\pi_m^2(R_T(\omega;u_{\mathcal{T}}))\|_{L^2(T)},\end{aligned} \tag{4.54}$$

$$\Big|\sum_{T\in\phi_E}\int_T(\mathrm{id}-\pi_m^2)(R_T(\omega;u_\mathcal{T}))w_E\,dx\Big|$$
$$\lesssim h_E^{\frac{1}{2}}\|\pi_n^2(R_E(\omega;u_\mathcal{T}))\|_{L^2(E)}\sum_{T\in\phi_E}\|(\mathrm{id}-\pi_m^2)(R_T(\omega;u_\mathcal{T}))\|_{L^2(T)}, \quad (4.55)$$

$$\Big|\int_E(\mathrm{id}-\pi_n^2)(R_E(\omega;u_\mathcal{T}))w_E\,dx\Big|$$
$$\leq \|(\mathrm{id}-\pi_n^2)(R_E(\omega;u_\mathcal{T}))\|_{L^2(E)}\|\pi_n^2(R_E(\omega;u_\mathcal{T}))\|_{L^2(E)}, \quad (4.56)$$

where the hidden constants depend only on the shape-regularity Γ and the orders m,n.

Together with (4.50), relations (4.51)-(4.56) provide

$$h_E^{\frac{1}{2}}\|\pi_n^2(R_E(\omega;u_\mathcal{T}))\|_{L^2(E)}\lesssim \alpha_{\max}(\omega)\|u-u_\mathcal{T}\|_{H^1(\phi_E)}$$
$$+h_E\sum_{T\in\phi_E}\|(\mathrm{id}-\pi_m^2)(R_T(\omega;u_\mathcal{T}))\|_{L^2(T)}+h_E^{\frac{1}{2}}\|(\mathrm{id}-\pi_n^2)(R_E(\omega;u_\mathcal{T}))\|_{L^2(E)}. \quad (4.57)$$

Finally, (4.50) and (4.57), the triangle and Young's inequalities, shape-regularity of the mesh $\mathcal{T}$ and the finite overlap property of patches ϕ_E, $E\in\mathcal{E}$ prove the lemma. □

Another property of the oscillations that is used later for deriving the order of convergence of the adaptive method is presented in the following proposition.

Proposition 4.3.2. *For all $T\in\mathcal{T}$, $\mathcal{T}\in\mathbb{T}$ and $v,w\in S_\mathcal{T}$ there almost surely holds*

$$osc_T(\omega;v)\lesssim osc_T(\omega;w)+h_T\|\alpha(\cdot,\omega)\|_{W^{1,\infty}(T)}\|v-w\|_{H^1(\phi_T)}, \quad (4.58)$$

where the hidden constant depends only on Γ and the dimension d.

Proof. We closely follow the derivation of (4.35) (see [26, Proposition 3.3]). The triangle inequality together with linearity of L^2-projections provide

$$osc_T(\omega;v)\leq osc_T(\omega;w)+h_T\|(\mathrm{id}-\pi_m^2)\big(\nabla\cdot(\alpha(\cdot,\omega)(v-w))\big)\|_{L^2(T)}$$
$$+h_T^{\frac{1}{2}}\|(\mathrm{id}-\pi_n^2)R_E(\omega;v-w)\|_{L^2(\partial T\cap D)}. \quad (4.59)$$

We use (A.13) in order to obtain the following estimate

$$\|(\mathrm{id}-\pi_m^2)\big(\nabla\cdot(\alpha(\cdot,\omega)(v-w))\big)\|_{L^2(T)}\leq\|(\mathrm{id}-\pi_m^\infty)\nabla\alpha(\cdot,\omega)\|_{L^\infty(T)}\|\nabla(v-w)\|_{L^2(T)}$$
$$\leq\|\nabla\alpha(\cdot,\omega)\|_{L^\infty(T)}\|\nabla(v-w)\|_{L^2(T)}. \quad (4.60)$$

Using the inverse inequality (A.12), (A.14) and shape-regularity of $\mathcal{T}$ we have

$$\begin{aligned}
&\|(\mathrm{id} - \pi_n^2) R_E(\omega; v - w)\|_{L^2(E)} \\
&\quad\lesssim \|(\mathrm{id} - \pi_n^\infty)\alpha(\cdot,\omega)\|_{L^\infty(T)} \left(h_T^{-\frac{1}{2}} \|\nabla(v-w)\|_{L^2(T)} + h_{T'}^{-\frac{1}{2}} \|\nabla(v-w)\|_{L^2(T')} \right) \\
&\quad\leq h_T^{\frac{1}{2}} \|\nabla\alpha(\cdot,\omega)\|_{L^\infty(T)} \|\nabla(v-w)\|_{L^2(\phi_E)},
\end{aligned} \tag{4.61}$$

where T' is the neighbour of T with a common face E.

Now, (4.60) and (4.61) substituted into (4.59) prove the claim of the proposition. □

Remark 4.3.2. All pathwise results in this subsection hold not only for globally Lipschitz realizations $\alpha(\cdot,\omega)$, $\omega \in \Omega$, but also for realizations that are only piecewise Lipschitz over $\mathcal{T}_0$, as well as for non-convex spatial domains D.

4.3.4 Measurability of solutions

In this section we address the question of measurability of the solution mappings $\Omega \ni \omega \mapsto u_l \in H_0^1(D)$ for $l \in \mathbb{N}$.

We first show measurability of the mappings $\Omega \ni \omega \mapsto u_{(k)}(\cdot,\omega) \in S_{(k)}(\omega) \subset H_0^1(D)$, where $u_{(k)}(\cdot,\omega)$, $k \in \mathbb{N}$ denote the solutions obtained by Algorithm 1.

Lemma 4.3.6. *The mappings $\Omega \ni \omega \mapsto u_{(k)}(\omega) \in S_{(k)}(\omega) \subset H_0^1(D)$ are measurable for all $k \in \mathbb{N}$.*

Proof. According to the result of [48, Theorem 2.3], which is obtained for a more general case of variational inequalities but also holds in the special case we consider, measurability of the set-valued mappings $\Omega \ni \omega \mapsto S_{(k)}(\omega) \subset H_0^1(D)$ implies measurability of the mappings $\Omega \ni \omega \mapsto u_{(k)}(\cdot,\omega) \in S_{(k)}(\omega) \subset H_0^1(D)$ for all $k \in \mathbb{N}$. In this case the mappings $\Omega \ni \omega \mapsto \eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) \in [0,\infty)$ are also measurable, because the $\eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega))$ have an explicit formula defined in (4.26)-(4.28) that includes only deterministic manipulations applied to measurable functions. Note that the element sizes $h_T(\omega)$, $T \in \mathcal{T}_{(k)}(\omega)$ are random, but are determined by the spaces $S_{(k)}(\omega)$.

Taking into account that $S_{(0)}$ is constant for all $\omega \in \Omega$ and that the corresponding set-valued map is trivially measurable, by induction it suffices to show that $\Omega \ni \omega \mapsto S_{(k)}(\omega) \subset H_0^1(D)$ implies measurability of $\Omega \ni \omega \mapsto S_{(k+1)}(\omega) \subset H_0^1(D)$ for any $k \in \mathbb{N} \cup \{0\}$.

Given $\mathbb{T}$, we can define a vector of all elements $(T_i)_{i\in\mathbb{N}}$, such that any partition $\mathcal{T} \in \mathbb{T}$ can be represented as $\mathcal{T} = \bigcup_{i\in\mathbb{N}} T_i e_i^{\mathbb{T}}(\mathcal{T})$, where $e^{\mathbb{T}}(\mathcal{T}) = (e_i^{\mathbb{T}}(\mathcal{T}))_{i\in\mathbb{N}}$ is the indicator vector defined as

$$e_i^{\mathbb{T}}(\mathcal{T}) := \begin{cases} 1, & \text{if} \quad T_i \in \mathcal{T}, \\ 0, & \text{else.} \end{cases}$$

Let $(\lambda_i)_{i\in\mathbb{N}}$ denote the collection of all finite element basis functions corresponding to all partitions $\mathcal{T} \in \mathbb{T}$, i.e. for every $\mathcal{T}$ the finite element space $S(\mathcal{T})$ can be represented as $S(\mathcal{T}) = \operatorname{span}\{\lambda_i e_i(S(\mathcal{T})),\ i \in \mathbb{N}\}$, where $e(S(\mathcal{T})) = (e_i(S(\mathcal{T})))_{i\in\mathbb{N}}$ is the indicator vector defined as

$$e_i(S(\mathcal{T})) := \begin{cases} 1, & \text{if} \quad \lambda_i \in S(\mathcal{T}), \\ 0, & \text{else.} \end{cases}$$

Since any $S_{(k)}(\omega)$, $k \in \mathbb{N} \cup \{0\}$, $\omega \in \Omega$ can be represented as $S_{(k)}(\omega) = \operatorname{span}\{\lambda_i e_i(S_{(k)}(\omega)),\ i \in \mathbb{N}\}$, in order for the mapping $\Omega \ni \omega \mapsto S_{(k)}(\omega) \subset H_0^1(D)$ to be measurable, it suffices to show that the indicator vector $e(S_{(k)})$ is measurable. Notice that there is a deterministic one-to-one correspondence between the indicator vectors $e(S_{(k)})$ and $e^{\mathbb{T}}(\mathcal{T}_{(k)})$. Therefore, it suffices to show measurability of the latter.

Summarizing, having measurability of $e^{\mathbb{T}}(\mathcal{T}_{(k)})$ we would like to show measurability of $e^{\mathbb{T}}(\mathcal{T}_{(k+1)})$. Without loss of generality we assume that for all $\omega \in \Omega$ the elements in $e^{\mathbb{T}}(\mathcal{T}_{(k)}(\omega))$ are ordered according to the decreasing order of corresponding $\eta_{T_i}(\omega; u_{(k)}(\cdot,\omega))$, where $\eta_{T_i}(\omega; u_{(k)}(\cdot,\omega))$ are set to 0 for $T_i \notin \mathcal{T}_{(k)}(\omega)$.

Each refinement step providing a mesh $\mathcal{T}_{(k+1)}(\omega)$ given $\mathcal{T}_{(k)}(\omega)$, $\omega \in \Omega$, consists of two substeps. The first substep is based on the Dörfler marking and the second substep performs closure. Let us denote the uniform refinement of $\mathcal{T}_{(k)}(\omega)$ by $\mathcal{T}^{\mathrm{un}}_{(k)}(\omega)$. The indicator vector $e^{\mathbb{T}}(\mathcal{T}^{\mathrm{un}}_{(k)})$ inherits its measurability from the $e^{\mathbb{T}}(\mathcal{T}_{(k)})$, because it is obtained in a deterministic procedure. Now, we denote the refinement of $\mathcal{T}_{(k)}(\omega)$ obtained by adaptive refinement based on the Dörfler marking by $\mathcal{T}^{\mathrm{D}}_{(k)}(\omega)$. The coefficients of $e^{\mathbb{T}}(\mathcal{T}^{\mathrm{D}}_{(k)}(\omega))$ can be represented as

$$\begin{aligned} e_i^{\mathbb{T}}(\mathcal{T}^{\mathrm{D}}_{(k)}(\omega)) = e_i^{\mathbb{T}}(\mathcal{T}^{\mathrm{un}}_{(k)}(\omega)) \Bigg(& H_1\Big(\theta \eta^2_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)) - \sum_{j=1}^{m} \eta^2_{T_j}(u_{(k)}(\cdot,\omega))\Big) \\ & + H_1\Big(\sum_{j=1}^{m} \eta^2_{T_j}(u_{(k)}(\cdot,\omega)) - \theta \eta^2_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega))\Big) \\ & \cdot H_0\Big(\theta \eta^2_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)) - \sum_{j=1}^{m-1} \eta^2_{T_j}(u_{(k)}(\cdot,\omega))\Big)\Bigg), \end{aligned}$$

where m is such that $T_i \subset T_m$ and $H_0(\cdot)$ and $H_1(\cdot)$ are piecewise constant functions, defined as

$$H_0(z) = \begin{cases} 1, & \text{if} \quad z > 0, \\ 0, & \text{if} \quad z \leq 0, \end{cases} \qquad H_1(z) = \begin{cases} 1, & \text{if} \quad z \geq 0, \\ 0, & \text{if} \quad z < 0. \end{cases} \tag{4.62}$$

Since all components of $e^{\mathbb{T}}(\mathcal{T}^{\mathrm{D}}_{(k)})$ can be represented as compositions, sums and multiplications of measurable functions, they are measurable themselves. Finally, in the closure step we obtain the refinement $\mathcal{T}_{(k+1)}(\omega)$ from the $\mathcal{T}^{\mathrm{D}}_{(k)}(\omega)$. Since closure is a deterministic procedure, the coefficient vector $e^{\mathbb{T}}(\mathcal{T}_{(k+1)})$ inherits its measurability from $e^{\mathbb{T}}(\mathcal{T}^{\mathrm{D}}_{(k)})$, which concludes the proof. □

Now, we are able to show measurability of the mappings $\Omega \ni \omega \mapsto u_l(\cdot, \omega) \in H^1_0(D)$, where $u_l(\cdot, \omega)$ are the solutions chosen according to (4.18).

Theorem 4.3.5. *The mappings $\Omega \ni \omega \mapsto u_l(\cdot, \omega) \in H^1_0$ are measurable for all $l \in \mathbb{N}$, i.e. Assumption 4.3.2 holds.*

Proof. According to Lemma 4.3.6, the mappings $\Omega \ni \omega \mapsto u_{(k)}(\cdot, \omega) \in S_{(k)}(\omega)$, where $u_{(k)}(\cdot, \omega)$, $k \in \mathbb{N}$ denote the solutions obtained by Algorithm 1, are measurable. Moreover, the mappings $\Omega \ni \omega \mapsto \eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot, \omega)) \in [0, \infty)$, $k \in \mathbb{N}$ are also measurable (see the proof of Lemma 4.3.6). The solutions u_l, $l \in \mathbb{N}$ can be represented as a sum of multiplications and compositions of measurable functions

$$u_l = \sum_{k=1}^{\infty} u_{(k)} H_1(\eta_{\mathcal{T}_{(k-1)}}(\cdot; u_{(k-1)}) - Tol_l) H_0(Tol_l - \eta_{\mathcal{T}_{(k)}}(\cdot; u_{(k)})),$$

where $H_0(\cdot)$ and $H_1(\cdot)$ are defined in (4.62), and, therefore, the mappings $\Omega \ni \omega \mapsto u_l(\cdot, \omega) \in H^1_0$ are measurable. □

Hence, we showed that Algorithm 1 with the residual error estimator defined in (4.26)-(4.28) fulfills Assumption 4.3.2. The only missing property required for the adaptive MLMC methods introduced in this chapter is a bound for the computational cost, i.e. Assumption 3.1.2.

4.3.5 Convergence order of the adaptive algorithm

In this section we analyse the order of convergence of Algorithm 1 when the residual-based error estimation introduced in Section 4.3.3 is used. Since mesh refinement in the algorithm is steered by the error estimator for a fixed $\omega \in \Omega$, the convergence in terms of the number of elements in the resulting meshes is closely related to the behaviour of the estimator $\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega))$. The best possible convergence order $0 < s < \infty$ obtained by adaptive mesh refinement can be characterized by

$$|\eta_{\mathcal{T}_{\mathrm{opt}}}(\omega; u_{\mathcal{T}_{\mathrm{opt}}}(\cdot, \omega))|_{\mathbb{B}_s} := \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} N^s \eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)),$$

where $\mathbb{T}(N) := \{\mathcal{T} \in \mathbb{T} : |\mathcal{T}| - |\mathcal{T}_0| \leq N\}$, $N \in \mathbb{N}$.

The condition $|\eta_{\mathcal{T}_{\text{opt}}}(\omega; u_{\mathcal{T}_{\text{opt}}}(\cdot,\omega))|_{\mathbb{B}_s} < \infty$ implies $\eta_{\mathcal{T}_{\text{opt}}(\omega)}(\omega; u_{\mathcal{T}_{\text{opt}}(\omega)}(\cdot,\omega)) = \mathcal{O}(N^{-s})$ for the optimal triangulations $\mathcal{T}_{\text{opt}}(\omega) \in \mathbb{T}(N)$, where the constant in $\mathcal{O}(\cdot)$ may, however, depend on $\omega \in \Omega$.

The following theorem shows quasi-optimal convergence of the adaptive Algorithm 1, i.e. if the decay order s is achieved when the optimal meshes are chosen, this order will be realized by the adaptive Algorithm 1. This result is adopted from [25, Theorem 4.1].

Theorem 4.3.6. *Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28)*, mesh refinement described in Section 4.3.2 and any Dörfler marking parameter $\theta(\omega)$ that fulfills $0 < \theta(\omega) < \theta^*(\omega) := \left(1 + C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)^{-1}$, provides almost surely quasi-optimal convergence of the estimator in the sense*

$$|\eta_{\mathcal{T}_{\text{opt}}}(\omega; u_{\mathcal{T}_{\text{opt}}}(\cdot,\omega))|_{\mathbb{B}_s} \lesssim \frac{\eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega))}{\left(|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}|\right)^{-s}} \lesssim C_{\text{opt}}(\omega) |\eta_{\mathcal{T}_{\text{opt}}}(\omega; u_{\mathcal{T}_{\text{opt}}}(\cdot,\omega))|_{\mathbb{B}_s} \tag{4.63}$$

for all $k \in \mathbb{N}$, $s > 0$, where $C_{\text{opt}}(\omega) := 1 + \left(\frac{1}{\theta(\omega)} + \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)\theta^2(\omega)}\right)^{s+\frac{1}{2}}$, the hidden constant in the left inequality depends only on the bisection depth b and the hidden constant in the right inequality depends only on b, s, d, h_0, C_{mesh}, $C_{\min}$, Γ and D.

Proof. We follow the deterministic results in [25], the random analogues of which are presented in Appendix B. We fix an arbitrary $\omega \in \Omega$ and $k \in \mathbb{N}$. The proof holds for almost all $\omega \in \Omega$.

We suppose that the right hand side of (4.63) is finite, otherwise the upper bound holds trivially.

We choose $\theta_0(\omega) := \left(1 + C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)^{-1}(1 - \delta^{-1}) < 1$ with some $\delta > 1$. Then Proposition B.0.1, together with the discrete reliability (4.32) and stability (4.33) of the error estimator, implies that there exist a $\kappa(\omega) > 0$, such that the implication

$$\begin{aligned} \eta^2_{\tilde{\mathcal{T}}}(\omega; u_{\tilde{\mathcal{T}}}(\cdot,\omega)) \leq \kappa(\omega)\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) \implies \\ \theta\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) \leq \sum_{T \in \mathcal{T}_{(k)}(\omega)\setminus\tilde{\mathcal{T}}} \eta^2_T(\omega; u_{(k)}(\cdot,\omega)) \end{aligned} \tag{4.64}$$

holds for all $0 < \theta \leq \theta_0$ and all refinements $\tilde{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}_{(k)}(\omega)$. Moreover, $\kappa(\omega)$ can be represented as (see the proof of Proposition B.0.1)

$$\kappa(\omega) = \frac{1 - \theta_0\left(1 + (1 + \delta^{-1})C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)}{1 + \delta}, \tag{4.65}$$

and it holds that $0 < \kappa(\omega) < 1$.

Substituting θ_0 into (4.65), we can see that

$$\kappa(\omega) = \frac{(1+\delta)\Big(1 + C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\Big)}{\delta^{-1} + \delta^{-2} C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}} \geq \delta^2(1+\delta).$$

This, together with the monotonicity (4.29) of the error estimator guaranteed by Proposition B.0.2, ensures that there exists a certain refinement $\tilde{\mathcal{T}}^*(\omega) \in \mathbb{T}$ of $\mathcal{T}_{(k)}(\omega)$ for which the set of refined elements $\mathcal{T}_{(k)}(\omega) \setminus \tilde{\mathcal{T}}^*(\omega) \subseteq \mathcal{T}_{(k)}(\omega)$ satisfies

$$|\mathcal{T}_{(k)}(\omega) \setminus \tilde{\mathcal{T}}^*(\omega)| \leq c_0(\omega) |\eta_{\mathcal{T}_{\text{opt}}}(\omega; u_{\mathcal{T}_{\text{opt}}}(\cdot,\omega))|^{\frac{1}{s}}_{\mathbb{B}_s} \eta^{-\frac{1}{s}}_{\mathcal{T}_{(k)}}(\omega; u_{(k)}(\cdot,\omega)), \tag{4.66}$$

where the constant $c_0(\omega)$ depends only on $\kappa(\omega)$ and s. Since $\kappa(\omega)$ is bounded from above and below by constants independent of ω, $c_0(\omega)$ is uniformly bounded from above by a constant that we denote by $c_{0,\max}$. By Proposition B.0.2, the set $\mathcal{T}_{(k)}(\omega) \setminus \tilde{\mathcal{T}}^*(\omega)$ also satisfies the Dörfler marking condition (4.21) for all $0 < \theta \leq \theta_0(\omega)$. Therefore, we have

$$|\mathcal{M}_{(k)}(\omega)| \lesssim |\mathcal{T}_{(k)}(\omega) \setminus \tilde{\mathcal{T}}^*(\omega)|, \tag{4.67}$$

where the hidden constants depend only on $C_{\min}$ from Remark 4.3.1.

Utilizing the closure estimate (4.22), together with (4.66) and (4.67), we then have

$$|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}| \lesssim \sum_{i=0}^{k-1} |\mathcal{M}_{(i)}(\omega)| \lesssim |\eta_{\mathcal{T}_{\text{opt}}}(\omega; u_{\mathcal{T}_{\text{opt}}}(\cdot,\omega))|^{\frac{1}{s}}_{\mathbb{B}_s} \sum_{i=0}^{k-1} \eta^{-\frac{1}{s}}_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)), \tag{4.68}$$

where the hidden constants depend only on C_{mesh}, $C_{\min}$, $c_{0,\max}$ and s.

The stability (4.33) and reduction (4.34) properties of the error estimator imply (see Proposition B.0.3) the estimator reduction

$$\eta^2_{\mathcal{T}_{(k+1)}(\omega)}(\omega; u_{(k+1)}(\cdot,\omega)) \leq \rho_1 \eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) + c_1(\omega) \|u_{(k+1)}(\cdot,\omega) - u_{(k)}(\cdot,\omega)\|^2, \tag{4.69}$$

for all $k \in \mathbb{N} \cup \{0\}$ with constants

$$\rho_1 := (1+\tilde{\delta})(1 - (1-\rho_{\text{red}})\theta) > 0, \quad c_1(\omega) := 2(1+\tilde{\delta}^{-1}) C_{\text{stab}}^2 \|\alpha\|^2_{W^{1,\infty}(D)} > 0,$$

where $\tilde{\delta} > 0$ can be chosen sufficiently small such that $\rho_1 < 1$, and $0 < \theta \leq \theta_0(\omega)$. We fix $\tilde{\delta} = \frac{(1-\rho_{\text{red}})\theta}{2(1-(1-\rho_{\text{red}})\theta)}$, then

$$0 < \rho_1 = 1 - \frac{1}{2}(1-\rho_{\text{red}})\theta < 1, \quad c_1(\omega) = 2\big(2\theta^{-1}(1-\rho_{\text{red}})^{-1} - 1\big) C_{\text{stab}}^2 \|\alpha\|^2_{W^{1,\infty}(D)} > 0.$$

According to Proposition B.0.4, the estimator reduction (4.69) implies

$$\sum_{i=k+1}^{\infty} \eta^2_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)) \leq c_2(\omega) \eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)), \tag{4.70}$$

where

$$
\begin{aligned}
c_2(\omega) &:= \left(\rho_1 + c_1(\omega)\frac{C_{\mathrm{qo}}}{\alpha^2_{\min}(\omega)}\right)(1-\rho_1)^{-1} \\
&= 2(1-\rho_{\mathrm{red}})^{-1}\theta^{-1} - 1 + 8C_{\mathrm{qo}}C^2_{\mathrm{stab}}(1-\rho_{\mathrm{red}})^{-2}\theta^{-2}\frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)} \\
&\qquad - 4C_{\mathrm{qo}}C^2_{\mathrm{stab}}(1-\rho_{\mathrm{red}})^{-1}\theta^{-1}\frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}.
\end{aligned}
\tag{4.71}
$$

In turn, (4.70) implies (see Proposition (B.0.5)) that

$$
\sum_{i=0}^{k-1} \eta^{-1/s}_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)) \le c_3(\omega)\eta^{-1/s}_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)), \tag{4.72}
$$

where $c_3(\omega) := (1+c_2(\omega))^{1/2s}(1-(1+c_2^{-1}(\omega))^{-1/2s})^{-1} = \frac{(1+c_2(\omega))^{\frac{1}{s}}}{(1+c_2(\omega))^{1/2s} - c_2^{1/2s}(\omega)}$.

By using monotonicity properties of the function $x^{1/2s}$, which is convex/concave for different values of s, one can show that

$$
c_3(\omega) \le \begin{cases} 2s(1+c_2(\omega))^{1+\frac{1}{2s}}, & s \ge \frac{1}{2}, \\ 2s(1+c_2(\omega))^{\frac{1}{s}} c_2^{1-\frac{1}{2s}}(\omega), & s < \frac{1}{2}. \end{cases}
$$

Now, combining (4.72) with (4.68), we obtain

$$
|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}| \lesssim c_3(\omega)|\eta_{\mathcal{T}_{\mathrm{opt}}}(\omega; u_{\mathcal{T}_{\mathrm{opt}}}(\cdot,\omega))|^{1/s}_{\mathbb{B}_s}\eta^{-1/s}_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)).
$$

Note that since $c_2(\omega) > 1$, there holds

$$
c_3^s(\omega) \lesssim \begin{cases} (1+c_2(\omega))^{s+\frac{1}{2}}, & s \ge \frac{1}{2}, \\ 1 + c_2^{s+\frac{1}{2}}(\omega), & s < \frac{1}{2}. \end{cases} \tag{4.73}
$$

Substituting (4.71) into (4.73) one can derive that $c_3^s(\omega) \lesssim C_{\mathrm{opt}}(\omega)$ for any $s > 0$. Hence, we obtained the upper bound in the statement of the theorem.

To prove the lower bound in (4.63) we now suppose that the middle part of (4.63) is finite, otherwise the bound holds trivially. Let $N \in \mathbb{N}$ be arbitrary, and let k be the largest integer such that $\mathcal{T}_{(k)}(\omega) \in \mathbb{T}(N)$. Then, by definition of $\mathbb{T}(N)$ and properties of refinement by bisection, we have

$$
N < |\mathcal{T}_{(k+1)}(\omega)| - |\mathcal{T}_{(0)}| < 2^b|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}|,
$$

which leads to

$$
\min_{\mathcal{T}\in\mathbb{T}(N)} N^s \eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)) \lesssim (|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}|)^s \eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)),
$$

which concludes the proof. □

Let us notice that θ^* in Theorem 4.3.6 depends on $\omega \in \Omega$ and in general does not have a uniform lower bound. In what follows we fix

$$\theta(\omega) := \left(1 + C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)^{-1} (1 - \delta^{-1}), \tag{4.74}$$

for some $\delta > 1$.

According to the reliability property (4.30) of the error estimator, the convergence order of the estimator directly implies at least the same convergence order for the error. However, according to the efficiency property (4.45), overestimation may occur and the error may decay faster than the estimator. The following proposition provides a relation between convergence of the error estimator $\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega))$ and convergence of the so-called *total error* (see [26]) that includes the H^1-error and the oscillations.

Proposition 4.3.3. *For a given $\mathcal{T} \in \mathbb{T}$ there a.s. holds*

$$\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)) \lesssim \inf_{v \in S(\mathcal{T})} (C_{\text{eq}}(\omega) \|u(\cdot, \omega) - v\|_{H^1(D)} + osc_{\mathcal{T}}(\omega; v)), \tag{4.75}$$

$$\inf_{v \in S(\mathcal{T})} (\|u(\cdot, \omega) - v\|_{H^1(D)} + osc_{\mathcal{T}}(\omega; v)) \lesssim \left(\frac{1}{\alpha_{\min}(\omega)} + 1\right) \eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)), \tag{4.76}$$

where $C_{\text{eq}}(\omega) := \frac{\alpha^{3/2}_{\max}(\omega)}{\alpha^{1/2}_{\min}(\omega)} + \left(\frac{\alpha^{1/2}_{\max}(\omega)}{\alpha^{1/2}_{\min}(\omega)} + 1\right) \|\alpha(\cdot, \omega)\|_{W^{1,\infty}(D)}$.

Proof. We follow the deterministic result in [25, Theorem 4.4]. The efficiency (4.45), oscillations property (4.58), Céa's lemma [28, Lemma 3.8] and triangle inequality provide (4.75). Indeed,

$$\begin{aligned}
&\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)) \lesssim \alpha_{\max}(\omega) \|u(\cdot, \omega) - u_{\mathcal{T}}(\cdot, \omega)\|_{H^1(D)} + osc_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot, \omega)) \\
&\lesssim \frac{\alpha^{3/2}_{\max}(\omega)}{\alpha^{1/2}_{\min}(\omega)} \|u(\cdot, \omega) - v\|_{H^1(D)} + \|\alpha(\cdot, \omega)\|_{W^{1,\infty}(D)} \|v - u_{\mathcal{T}}(\cdot, \omega)\|_{H^1(D)} + osc_{\mathcal{T}}(\omega; v) \\
&\lesssim \frac{\alpha^{3/2}_{\max}(\omega)}{\alpha^{1/2}_{\min}(\omega)} \|u(\cdot, \omega) - v\|_{H^1(D)} + \|\alpha(\cdot, \omega)\|_{W^{1,\infty}(D)} \left(\frac{\alpha^{1/2}_{\max}(\omega)}{\alpha^{1/2}_{\min}(\omega)} + 1\right) \|v - u(\cdot, \omega)\|_{H^1(D)} \\
&\qquad + osc_{\mathcal{T}}(\omega; v),
\end{aligned}$$

for all $v \in S(\mathcal{T})$.

The reliability (4.30) and (4.44) imply (4.76). □

According to Proposition 4.3.3, the condition $|\eta_{\mathcal{T}_{\text{opt}}}(\omega; u_{\mathcal{T}_{\text{opt}}}(\cdot, \omega))|_{\mathbb{B}_s} < \infty$ is equivalent to $|u(\cdot, \omega), \alpha(\cdot, \omega), f(\cdot, \omega)|_{\mathbb{A}_s} < \infty$, where

$$|u, \alpha, f|_{\mathbb{A}_s} := \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} \inf_{v \in S(\mathcal{T})} N^s (\|u - v\|_{H^1(D)} + osc_{\mathcal{T}}(\omega; v)).$$

Note that the approximation class $\mathbb{A}_s := \{(u,\alpha,f) : |u,\alpha,f|_{\mathbb{A}_s} < \infty\}$ involves non-linear interactions between u and the functions f, α through the oscillations. Therefore, characterization of this class is far from trivial and is still available only partly in the literature, see Remark 4.3.4 below.

Although Theorem 4.3.6 provides quasi-optimal convergence of the error estimators, one does not always know what the optimal order is. The following theorem states that convergence of the adaptive finite element method is at least as good as convergence of the uniform one.

Theorem 4.3.7. *For all $k \in \mathbb{N}$ and almost all $\omega \in \Omega$, Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28) *and mesh refinement described in Section 4.3.2 produces $\mathcal{T}_{(k)}(\omega)$ and corresponding $u_{(k)}(\cdot,\omega)$, $\eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega))$, such that*

$$\eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) \lesssim C_{H^2}(\omega)\big(|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}|\big)^{-\frac{1}{d}},$$

where $C_{H^2}(\omega) := C_{\mathrm{opt}}(\omega)\left(\alpha_{\max}(\omega)C_{\mathrm{un}}(\omega) + \left(1 + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha_{\min}(\omega)}\right)\|f(\cdot,\omega)\|_{L^2(D)}\right).$

Proof. According to Theorem 4.3.6 we have

$$\begin{aligned}\eta_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) &\lesssim C_{\mathrm{opt}}(\omega)\big(|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}|\big)^{-s}|\eta_{\mathcal{T}_{\mathrm{opt}}}(\omega; u_{\mathcal{T}_{\mathrm{opt}}}(\cdot,\omega))|_{\mathbb{B}_s}\\ &= C_{\mathrm{opt}}(\omega)\big(|\mathcal{T}_{(k)}(\omega)| - |\mathcal{T}_{(0)}|\big)^{-s}\sup_{N\in\mathbb{N}}\min_{\mathcal{T}\in\mathbb{T}(N)} N^s\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)).\end{aligned}$$

The efficiency (4.45) of the error estimator provides

$$\begin{aligned}\sup_{N\in\mathbb{N}}\min_{\mathcal{T}\in\mathbb{T}(N)} N^s\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)) \le \sup_{N\in\mathbb{N}}\min_{\mathcal{T}\in\mathbb{T}(N)} N^s\big(&\alpha_{\max}(\omega)\|u(\cdot,\omega) - u_{\mathcal{T}}(\cdot,\omega)\|_{H^1(D)}\\ &+ osc_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega))\big).\end{aligned} \tag{4.77}$$

We choose the $\mathcal{T}_{\mathrm{un}} \in \mathbb{T}$ of minimal cardinality obtained by uniform refinement of $\mathcal{T}_{(0)}$ such that $|\mathcal{T}_{\mathrm{un}}| \ge N$, then Lemma 4.2.1 yields

$$\|u(\cdot,\omega) - u_{\mathcal{T}_{\mathrm{un}}}(\cdot,\omega)\|_{H^1(D)} \lesssim C_{\mathrm{un}}(\omega)N^{-1/d}. \tag{4.78}$$

Using the oscillations property (4.58) with $v = u_{\mathcal{T}_{\mathrm{un}}}(\cdot,\omega)$, $w = 0$ and (4.5) we have

$$\begin{aligned}osc_{\mathcal{T}_{\mathrm{un}}}(\omega; u_{\mathcal{T}_{\mathrm{un}}}(\cdot,\omega)) &\lesssim h_{\mathrm{un}}\big(\|(\mathrm{id} - \pi_m^2)f(\cdot,\omega)\|_{L^2(D)} + \|\alpha\|_{W^{1,\infty}(D)}\|u_{\mathcal{T}_{\mathrm{un}}}(\cdot,\omega)\|_{H^1(D)}\big)\\ &\lesssim N^{-\frac{1}{d}}\left(1 + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha_{\min}(\omega)}\right)\|f(\cdot,\omega)\|_{L^2(D)}.\end{aligned} \tag{4.79}$$

Since $\mathcal{T}_{\mathrm{un}} \in \mathbb{T}(N)$, combining (4.77)-(4.79) we have

$$\sup_{N\in\mathbb{N}}\min_{\mathcal{T}\in\mathbb{T}(N)} N^{\frac{1}{d}}\eta_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)) \lesssim \alpha_{\max}(\omega)C_{\mathrm{un}}(\omega) + \left(1 + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha_{\min}(\omega)}\right)\|f(\cdot,\omega)\|_{L^2(D)},$$

which proves the theorem. ∎

Returning to the multilevel Monte Carlo method, one can show using Theorem 4.3.7 that Assumption 3.1.2 holds with the appropriate constants when the adaptive Algorithm 1 with residual-based error estimation and multigrid methods are used. Note that Algorithm 1 includes the computation of the solutions and error estimators for all $k = 0, \dots, k_l(\omega)$.

Algorithm 1 terminates when the condition (4.18) is fulfilled, which implies

$$\eta_{\mathcal{T}_{(k_l(\omega))}(\omega)}(u_{(k_l(\omega))}(\cdot,\omega)) \leq Tol_l < \eta_{\mathcal{T}_{(k_l(\omega)-1)}(\omega)}(u_{(k_l(\omega)-1)}(\cdot,\omega)). \tag{4.80}$$

Incorporating the result of Theorem 4.3.7, we obtain

$$Tol_l < C_{H^2}(\omega)\big(|\mathcal{T}_{(k_l(\omega)-1)}(\omega)| - |\mathcal{T}_{(0)}|\big)^{-\frac{1}{d}},$$

which yields

$$|\mathcal{T}_{(k_l(\omega))}(\omega)| < 2^b|\mathcal{T}_{(k_l(\omega)-1)}(\omega)| < 2^b|\mathcal{T}_{(0)}| + 2^b C_{H^2}^d(\omega) Tol_l^{-d}.$$

According to (4.15), the uniform partition $\mathcal{T}_{(0)}$ fulfills $|\mathcal{T}_{(0)}| \simeq Tol_0^{-d}$, with the hidden constant that depends only on the size of the domain D, the dimension d, $\|C_{\mathrm{ad}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}$ and $\|C_{\mathrm{un}}\|_{L^{p_{\mathrm{ad}}}(\Omega)}$. Then, taking into account that $Tol_0^{-d} < Tol_l^{-d}$ for all $l \in N$, we have

$$|\mathcal{T}_{(k_l(\omega))}(\omega)| \lesssim (1 + C_{H^2}^d(\omega)) Tol_l^{-d}. \tag{4.81}$$

Theorem 4.3.7 also yields the relation

$$|\mathcal{T}_{(k)}(\omega)| \lesssim |\mathcal{T}_{(0)}| + C_{H^2}^d(\omega)\eta^{-d}_{\mathcal{T}_{(k)}(\omega)}(u_{(k)}(\cdot,\omega)),$$

for all $k \in \mathbb{N} \cup \{0\}$, which, in combination with the convergence stated in Theorem 4.3.3, relation (4.80) and properties of geometric series, provides

$$\begin{aligned}
\sum_{k=0}^{k_l(\omega)-1} |\mathcal{T}_k(\omega)| &\leq k_l(\omega)|\mathcal{T}_{(0)}| + C_{H^2}^d(\omega) C_{\mathrm{conv}}^{d/2}(\omega)\eta^{-d}_{\mathcal{T}_{(k_l(\omega)-1)}}(u_{(k_l(\omega)-1)}(\cdot,\omega)) \\
&\qquad \cdot \sum_{k=0}^{k_l(\omega)-1} \rho_{\mathrm{conv}}^{d(k_l(\omega)-k-1)/2}(\omega) \\
&\leq k_l(\omega)|\mathcal{T}_{(0)}| + C_{H^2}^d(\omega) C_{\mathrm{conv}}^{d/2}(\omega)(1-\rho_{\mathrm{conv}}^{d/2}(\omega))^{-1} Tol_l^{-d}.
\end{aligned} \tag{4.82}$$

The definition of the residual-based estimator (4.26)-(4.28), together with the triangle and Young's inequalities and (4.5), provides

$$\eta_{\mathcal{T}_{(0)}}(u_{(0)}(\cdot,\omega)) \lesssim C_0(\omega) := \Big(1 + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha_{\min}(\omega)}\Big)\|f(\cdot,\omega)\|_{L^2(D)}, \tag{4.83}$$

where the hidden constant depends only on h_0, D and the dimension d. Since

$C_{\text{opt}}(\omega) > 1$ for all $\omega \in \Omega$ (see Theorem 4.3.6), we have $C_0(\omega) < C_{H^2}(\omega)$ by definition of $C_{H^2}(\omega)$.

Now, combining (4.80) and the result of Theorem 4.3.3, we have

$$Tol_l^2 \leq C_{\text{conv}}(\omega)\rho_{\text{conv}}^{k_l(\omega)-1}(\omega)\eta^2_{\mathcal{T}_{(0)}}(u_{(0)}(\cdot,\omega)),$$

which together with (4.83) yields

$$\begin{aligned} k_l(\omega) - 1 &\leq \log_{\rho_{\text{conv}}(\omega)}\left(C_{\text{conv}}^{-1}(\omega)\eta^{-2}_{\mathcal{T}_{(0)}}(u_{(0)}(\cdot,\omega))Tol_l^2\right) \\ &< \log_{\rho_{\text{conv}}(\omega)}\left(C_{\text{conv}}^{-1}(\omega)C_{H^2}^{-2}(\omega)Tol_l^2\right) \\ &< \frac{2}{d}\ln(\rho_{\text{conv}}^{-1}(\omega))^{-1}C_{\text{conv}}^{d/2}(\omega)C_{H^2}^{d}(\omega)Tol_l^{-d}. \end{aligned} \tag{4.84}$$

We again utilize multigrid methods [20, 22] for computing the approximations $\tilde{u}_{(k)}(\omega)$ to the solutions $u_{(k)}(\omega)$, $\omega \in \Omega$, $k \in \mathbb{N} \cup \{0\}$. The multigrid methods have linear complexity in $|\mathcal{T}_{(k)}(\omega)|$ at the mesh level k, i.e.

$$Cost(\tilde{u}_{(k)}(\omega)) \leq C_{\text{MG}}(\omega)|\mathcal{T}_{(k)}(\omega)|, \tag{4.85}$$

for all $k \in \mathbb{N} \cup \{0\}$ and almost all $\omega \in \Omega$, where C_{MG} depends only on Γ, $\alpha_{\min}$ and $\alpha_{\max}$. The error estimators can be computed with the cost

$$Cost(\eta_{\mathcal{T}_{(k)}}(u_{(k)}(\cdot,\omega))) \leq C_\eta|\mathcal{T}_{(k)}(\omega)|, \tag{4.86}$$

where C_η depends only on the dimension d. Then, utilizing (4.81), (4.82), (4.84), (4.85) and (4.86), we have

$$\begin{aligned} Cost(\tilde{u}_l(\omega)) &\leq C_{\text{MG}}(\omega)|\mathcal{T}_{(k_l(\omega))}(\omega)| + \left(C_\eta + C_{\text{MG}}(\omega)\right)\sum_{k=0}^{k_l(\omega)-1}|\mathcal{T}_{(k)}(\omega)| \\ &\lesssim \left(C_{\text{MG}}(\omega)(1 + C_{H^2}^d(\omega)) + (C_\eta + C_{\text{MG}}(\omega))(1 + C_{H^2}^d(\omega)C_\alpha(\omega))\right)Tol_l^{-d}, \end{aligned} \tag{4.87}$$

where $C_\alpha(\omega) := \ln(\rho_{\text{conv}}^{-1}(\omega))^{-1}C_{\text{conv}}^{d/2}(\omega) + C_{\text{conv}}^{d/2}(\omega)(1 - \rho_{\text{conv}}^{d/2}(\omega))^{-1}$.

Note that the random variables C_{MG} and C_α depend only on $\alpha_{\min}$, $\alpha_{\max}$, $\|\alpha\|_{W^{1,\infty}(D)}$ and some deterministic constants. We make the following assumption in order to be able to bound the expected computational cost for computing solutions $u_l(\cdot,\omega)$, $\omega \in \Omega$ for all $l \in \mathbb{N} \cup \{0\}$.

Assumption 4.3.3. $C_{\text{MG}}, C_\alpha \in L^p(\Omega)$ for all $p \in [1, \infty)$.

We are then ready to state the following theorem.

Theorem 4.3.8. *Let Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28), *mesh refinement described in Section 4.3.2 and*

$0 < \theta(\omega) < \theta^*(\omega) := \left(1 + C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)^{-1}$ *be utilized for constructing the pathwise hierarchies of spaces* $S_{(k)}(\omega)$, $k \in \mathbb{N} \cup \{0\}$, *let Assumption 4.3.3 hold and* $p_f \in (\max\{d, 2\}, \infty]$. *Then multigrid methods applied for iterative solution of the pathwise discretized problems of the form* (4.12) *fulfill Assumption 3.1.2 with* $\gamma = d$ *and a constant* C_{cost} *that does not depend on* $l \in \mathbb{N} \cup \{0\}$ *and* $\omega \in \Omega$.

Proof. According to Proposition C.0.7, $\|C^d_{H^2}(\omega))\|_{L^p(\Omega)} < \infty$ for all $p \in [1, \frac{p_f}{d})$. Then relation (4.87), Assumption 4.3.3 and Hölder's inequality imply

$$\begin{aligned}\mathbb{E}[Cost(\tilde{u}_l(\omega))] \lesssim \Big(&C_\eta + \mathbb{E}[C_{\text{MG}}] + \|C_{\text{MG}}\|_{L^{p_1}(\Omega)}\|C^d_{H^2}(\omega))\|_{L^p(\Omega)} \\ &+ C_\eta \|C_\alpha\|_{L^{p_1}(\Omega)}\|C^d_{H^2}(\omega))\|_{L^p(\Omega)} \\ &+ \|C_{\text{MG}}\|_{L^{p_2}(\Omega)}\|C_\alpha\|_{L^{p_3}(\Omega)}\|C^d_{H^2}(\omega))\|_{L^p(\Omega)}\Big) Tol_l^{-d},\end{aligned}$$

where $p \in [1, \frac{p_f}{d})$, $p_1 = \frac{p}{p-1}$, $\frac{1}{p_2} + \frac{1}{p_3} = \frac{1}{p}$ and the hidden constant depends only on b, d, D, h_0, Γ, $\|C_{\text{ad}}\|_{L^{p_{\text{ad}}}(\Omega)}$ and $\|C_{\text{un}}\|_{L^{p_{\text{ad}}}(\Omega)}$. This concludes the proof. □

As it is seen in the proof of Theorem 4.3.8, it is possible to weaken Assumption 4.3.3 and only assume that $C_{\text{MG}} \in L^{p_{\text{MG}}}(\Omega)$ and $C_\alpha \in L^{p_\alpha}(\Omega)$ for some $p_{\text{MG}}, p_\alpha \in (1, \infty)$. For the sake of simplicity however, we have chosen to use Assumption 4.3.3.

Remark 4.3.3. If we let the stronger set of assumptions described in Remark 2.2.2 hold, $C_{\text{MG}}, C_\alpha \in L^\infty(\Omega)$ and it is enough to assume $p_f \in [d, \infty]$ in Theorem 4.3.8.

According to Theorem 4.3.8, in order to have Assumption 3.1.2 fulfilled for adaptive MLMC methods described in this section, we need to strengthen Assumption (2.2.1) (iii). Particularly, in order to have the result of Theorem 4.3.2 we need to require $p_f > \max\{2\frac{p_Q}{p_Q - 2}, d\}$.

Although Assumption 3.1.2 is fulfilled for both classical and adaptive finite element methods, the constant C_{cost} corresponding to one of the methods might be larger than the constant corresponding to the other. This depends on properties of the random solutions. If the random solutions exhibit locally rapidly changing properties, the adaptive finite element method might perform better than the classical uniform method.

Remark 4.3.4. In case when the functions $\alpha(\cdot, \omega)$ are piecewise polynomials over $\mathcal{T}_0$ (not fulfilled for log-normal fields), it is possible to characterize the class $\mathbb{A}_s$ in terms of standard approximation classes [26, Lemma 5.3]

$$\mathcal{A}_s := \{u \in H^1_0(D) : \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} \inf_{v \in S(\mathcal{T})} N^s \|u - v\|_{H^1(D)}\},$$

$$\tilde{\mathcal{A}}_s := \{f \in L^2(D) : \sup_{N \in \mathbb{N}} \min_{\mathcal{T} \in \mathbb{T}(N)} N^s \|(\text{id} - \pi^2_m) f\|_{L^2(D)}\}.$$

The class $\mathcal{A}_s$ can, in turn, be characterized by some particular Besov spaces in the case $d = 2$, see [17, 18]. Therefore, the theory of adaptive methods presented in this thesis can be extended in this case to a larger class of problems, which is beyond the scope of this work.

Remark 4.3.5. In the case when one is interested in an expected output of interest and this output is described by a linear continuous functional, it is possible to improve both the theoretical guarantees and practical performance of the adaptive MLMC-FE method by utilizing goal-oriented error estimation [46, 71, 75]. Based on the results in [71] it is likely possible to extend the results presented in this section to adaptive FE algorithms based on goal-oriented error estimation. Some numerical results for a method combining the MLMC ideas and pathwise goal-oriented error estimation are shown in [33].

Remark 4.3.6. All results in this section are presented for problem (4.1). As for problem (2.8), there exist convergent adaptive FE methods based on residual error estimation [21] and hierarchical error estimation [81] for variational inequalities. The order of convergence is, however, still an open question in both approaches, which makes the analysis of the computational cost of the methods difficult.

To summarise, we have shown that our assumptions required for convergence of the adaptive MLMC-FE methods introduced in Section 4.3 can be fulfilled by Algorithm 1 with the residual-based error estimator defined in (4.26)-(4.28) and mesh refinement described in Section 4.3.2. Theorem 4.3.4 provides the validity of Assumption 4.3.1. Theorem 4.3.5 shows that Assumption 4.3.2 is fulfilled. Finally, Assumption 3.1.2 holds under Assumption 4.3.3, according to Theorem 4.3.8.

4.4 Level dependent selection of sample numbers

The multilevel Monte Carlo methods described in Sections 3.1 and 3.2 cannot be implemented according to Theorems 3.1.1 and 3.2.1, because the numbers of samples on each level introduced in the proofs and the spatial error estimation depend on constants that are usually not available in practice. We will follow the heuristic approach suggested by Giles in [41, 40] and give two versions of the multilevel Monte Carlo algorithms corresponding to Sections 3.1 and 3.2 respectively.

Approximation of the expected solution

The algorithm is based on the identity (3.5) and aims at approximating $\mathbb{E}[u]$ up to a tolerance Tol, i.e

$$\|\mathbb{E}[u] - E^L[\tilde{u}_L]\|_{L^2(\Omega;H^1_0(D))} \leq Tol. \tag{4.88}$$

Utilizing (3.5), one can achieve (4.88), for example, requiring

$$\|\mathbb{E}[u - \tilde{u}_L]\|_{H^1(D)} \leq \frac{Tol}{\sqrt{2}}, \tag{4.89}$$

$$\sum_{l=0}^{L} M_l^{-1} V[\tilde{u}_l - \tilde{u}_{l-1}] \leq \frac{Tol^2}{2}. \tag{4.90}$$

It is not possible to verify (4.89) directly because it includes the unknown u. We assume that for each $l \in \mathbb{N}$ and for almost all $\omega \in \Omega$ there holds

$$\|\tilde{u}_{l+1}(\omega) - \tilde{u}_l(\omega)\|_{H^1(D)} \leq q\|\tilde{u}_l(\omega) - \tilde{u}_{l-1}(\omega)\|_{H^1(D)}, \tag{4.91}$$

where q is as in (3.2). Then, utilizing Jensen's inequality, the triangle inequality and (4.91), we have

$$\begin{aligned}\|\mathbb{E}[u - \tilde{u}_L]\|_{H^1(D)} &\leq \mathbb{E}[\|u - \tilde{u}_L\|_{H^1(D)}] = \mathbb{E}[\|\textstyle\sum_{l=L+1}^{\infty}(\tilde{u}_l - \tilde{u}_{l-1})\|_{H^1(D)}] \\ &\leq \mathbb{E}[\textstyle\sum_{l=L+1}^{\infty}\|\tilde{u}_l - \tilde{u}_{l-1}\|_{H^1(D)}] \leq (q^{-1} - 1)^{-1}\mathbb{E}[\|\tilde{u}_L - \tilde{u}_{L-1}\|_{H^1(D)}].\end{aligned}$$

Therefore, in practical computations, condition (4.89) can be replaced by

$$\mathbb{E}[\|\tilde{u}_L - \tilde{u}_{L-1}\|_{H^1(D)}] \leq (q^{-1} - 1)\frac{1}{\sqrt{2}}Tol. \tag{4.92}$$

In order to ensure (4.90) and define the number of samples on different levels we solve the following optimization problem

$$\begin{aligned}&\min_{M_l} \textstyle\sum_{l=0}^{L} M_l\mathbb{E}[N_l], \\ &\text{s.t.} \textstyle\sum_{l=0}^{L} M_l^{-1}V[\tilde{u}_l - \tilde{u}_{l-1}] = Tol^2/2,\end{aligned} \tag{4.93}$$

treating the M_l as continuous variables, where $N_l(\omega) = \dim S_l(\omega)$.

The solution of this problem is given by

$$M_l^{\text{opt}} = 2Tol^{-2}\sum_{i=0}^{L}\sqrt{V[\tilde{u}_i - \tilde{u}_{i-1}]\mathbb{E}[N_i]}\sqrt{\frac{V[\tilde{u}_l - \tilde{u}_{l-1}]}{\mathbb{E}[N_l]}}. \tag{4.94}$$

We replace $\mathbb{E}$ and V in (4.92) and (4.94) by computable sample approximations E_M and V_M obtained from M samples and obtain

$$E_{M_L}[\|\tilde{u}_L - \tilde{u}_{L-1}\|_{H^1(D)}] \leq (q^{-1} - 1)\frac{1}{\sqrt{2}}Tol, \tag{4.95}$$

$$M_l^{\text{opt}} = 2Tol^{-2}\sum_{i=0}^{L}\sqrt{V_{M_i}[\tilde{u}_i - \tilde{u}_{i-1}]E_{M_i}[N_i]}\sqrt{\frac{V_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}]}{E_{M_l}[N_l]}}. \tag{4.96}$$

Now, we are ready to list the multilevel Monte Carlo algorithm for approximating the expected solution, see Algorithm 2.

Note that in the uniform case there is no need to compute the average partition sizes in (4.96), because they are constant for each level. Furthermore, evaluation of samples in Algorithm 2 in the adaptive case is performed by applying Algorithm 1.

Algorithm 2 Adaptive MLMC algorithm for expected solution

Input: $\mathcal{T}_0, q, Tol, M_{\text{init}}$
Initialize $L := 1$, set $M_l := M_{\text{init}},\ l = 0, 1$
while not converged **do**
  Evaluate required samples
  Compute / update $V_{M_l}[\tilde{u}_l - \tilde{u}_{l-1}],\ l = 0, \ldots, L$
  Compute optimal $M_l^{\text{opt}},\ l = 0, \ldots, L$ according to (4.96)
  Evaluate additional samples
  Check convergence according to (4.95)
  if not converged **then**
    Set $L := L + 1$ and $M_L := M_{\text{init}}$
  end if
end while
Compute $E^L[\tilde{u}_L]$ according to (3.4)
Output: $E^L[\tilde{u}_L]$

Approximation of the expected output of interest

The algorithm for approximating the expected output of interest is based on the identity (3.10) and aims at approximating the mean root square error $e(E^L[Q(\tilde{u}_L)])$ up to the tolerance Tol, which can be done by enforcing the inequalities

$$|\mathbb{E}[Q(u) - Q(\tilde{u}_L)]| \leq \frac{Tol}{\sqrt{2}}, \tag{4.97}$$

$$\sum_{l=0}^{L} M_l^{-1} \mathbb{V}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})] \leq \frac{Tol^2}{2}. \tag{4.98}$$

We make an assumption similar to (4.91), namely that for each $l \in \mathbb{N}$ and for almost all $\omega \in \Omega$ there holds

$$|Q(\tilde{u}_{l+1}) - Q(\tilde{u}_l)| \leq q|Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})|. \tag{4.99}$$

Then, condition (4.97) can be replaced by

$$\mathbb{E}[|Q(\tilde{u}_L) - Q(\tilde{u}_{L-1})|] \leq (q^{-1} - 1)\frac{1}{\sqrt{2}} Tol. \tag{4.100}$$

We ensure condition (4.98) by solving the optimization problem

$$\begin{gathered} \min_{M_l} \textstyle\sum_{l=0}^{L} M_l \mathbb{E}[N_l], \\ \text{s.t.} \textstyle\sum_{l=0}^{L} M_l^{-1} \mathbb{V}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})] = Tol^2/2. \end{gathered} \tag{4.101}$$

The solution is given by

$$M_l^{\mathrm{opt}} = 2Tol^{-2} \sum_{i=0}^{L} \sqrt{\mathbb{V}[Q(\tilde{u}_i) - Q(\tilde{u}_{i-1})]\mathbb{E}[N_i]} \sqrt{\frac{\mathbb{V}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})]}{\mathbb{E}[N_l]}}. \qquad (4.102)$$

We replace $\mathbb{E}$ in (4.100) and (4.102) by a computable sample approximation E_M with M samples. Abusing the notation we use V_M to denote a sample approximation of $\mathbb{V}$ with M samples here and use this approximation in (4.102). Therefore, we have

$$E_{M_L}[|Q(\tilde{u}_L) - Q(\tilde{u}_{L-1})|] \leq (q^{-1} - 1)\frac{1}{\sqrt{2}}Tol, \qquad (4.103)$$

$$M_l^{\mathrm{opt}} = 2Tol^{-2} \sum_{i=0}^{L} \sqrt{V_{M_i}[Q(\tilde{u}_i) - Q(\tilde{u}_{i-1})]E_{M_i}[N_i]} \sqrt{\frac{V_{M_l}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})]}{E_{M_l}[N_l]}}. \qquad (4.104)$$

The multilevel Monte Carlo algorithm for approximating the expected output of interest is listed in Algorithm 3.

Algorithm 3 Adaptive MLMC algorithm for expected output of interest

Input: $\mathcal{T}_0$, q, Tol, M_{init}
Initialize $L := 1$, set $M_l := M_{\mathrm{init}}$, $l = 0, 1$
while not converged **do**
    Evaluate required samples
    Compute / update $V_{M_l}[Q(\tilde{u}_l) - Q(\tilde{u}_{l-1})]$, $l = 0, \dots, L$
    Compute optimal M_l^{opt}, $l = 0, \dots, L$ according to (4.104)
    Evaluate additional samples
    Check convergence according to (4.103)
    if not converged **then**
        Set $L := L + 1$ and $M_L := M_{\mathrm{init}}$
    end if
end while
Compute $E^L[Q(\tilde{u}_L)]$ according to (3.9)
Output: $E^L[Q(\tilde{u}_L)]$

Again, the partition sizes in (4.104) are constant for each level in the uniform case, therefore the averages do not need to be computed. In the adaptive case, Algorithm 1 is used for evaluation of the samples in Algorithm 3.

Chapter 5

Numerical Experiments

In this chapter we investigate the MLMC methods presented in Sections 4.2 and 4.3 from a numerical perspective. We use the algorithms presented in Section 4.4.

5.1 Poisson problem with random right-hand side

We first consider the Poisson problem

$$u(\cdot,\omega) \in \{w \in H^1(D) : \gamma_D(w) = u_D(\cdot,\omega) \text{ a.e. on } \partial D\} :$$
$$a(\omega; u(\cdot,\omega), v) = \ell(\omega; v), \qquad \forall v \in H_0^1(D), \tag{5.1}$$

where $D = (-1,1)^2$ and γ_D denotes the trace map that associates $w \in H^1(D)$ to the restriction $\gamma_D(w)$ in $H^{1/2}(\partial D)$. The bilinear and the linear forms are defined as in (2.5) with the coefficient $\alpha(x,\omega) = 1$ and the uncertain source term

$$f(x,\omega) = e^{-\beta|x-Y(\omega)|^2}(4\beta^2|x-Y(\omega)|^2 - 4\beta). \tag{5.2}$$

The uncertain inhomogeneous boundary condition is given by

$$u_D(x,\omega) = e^{-\beta|x-Y(\omega)|^2}, \quad x \in \partial D,\ \omega \in \Omega.$$

Here β is a positive constant and $Y = (Y_1, Y_2)^\top$ is a random vector which components are uniformly distributed random variables $Y_1, Y_2 \sim \mathcal{U}(-0.25, 0.25)$. For each $\omega \in \Omega$ the pathwise solution to (5.1) is given by

$$u(x,\omega) = e^{-\beta|x-Y(\omega)|^2}, \quad x \in D. \tag{5.3}$$

The stronger version of Assumption 2.2.1 described in Remark 2.2.2 is satisfied for this problem. Therefore, the solution is unique and we have spatial regularity in the sense that $u(\cdot,\omega) \in H^2(D)$ for almost all $\omega \in \Omega$. However, $u(\cdot,\omega)$ exhibits a peak at $(Y_1(\omega), Y_2(\omega)) \in D$ that becomes more pronounced with increasing β, thus leading to a larger constant C_{un} in the uniform error estimate (4.6). We have $p_f = \infty$ and $C_{\text{un}}, C_{\text{ad}} \in L^\infty(\Omega)$ in this toy problem. We are interested in the expected solution $\mathbb{E}[u]$ and compare the performance of MLMC finite element methods based on uniform and adaptive refinements for $\beta = 10, 50, 150$.

Pathwise adaptive refinement is performed as described in Section 4.3, with the exact finite element solution replaced by an approximation provided by an iterative method to be described below. Since selection of $\mathcal{T}_{(0)}$ according to (4.15) is difficult in practice, we choose the initial partition $\mathcal{T}_{(0)}$ that consists of 128 congruent triangles for both uniform and adaptive MLMC and set

$$Tol_0 := \|\eta_{\mathcal{T}_{(0)}}(\cdot; u_{(0)})\|_{L^2(\Omega)},$$

where we approximate the $L^2(\Omega)$-norm by a Monte Carlo method with 1000 samples. We choose Tol_l according to (3.2) with $q = \frac{1}{2}$. This choice guarantees

$$\|u - u_l\|_{L^2(\Omega;H^1_0(D))} \leq \|C_{\text{ad}}\|_{L^\infty(\Omega)} Tol_l$$

for all $l \in \mathbb{N} \cup \{0\}$ in the adaptive case. The accuracy criterion (4.18) then takes the form

$$\eta_{\mathcal{T}_{(k_l(\omega))}(\omega)}(\omega; u_{(k_l(\omega))}(\cdot,\omega)) \leq Tol_l = q^l \|\eta_{\mathcal{T}_{(0)}}(\cdot; u_{(0)})\|_{L^2(\Omega)}, \tag{5.4}$$

and is used as the stopping criterion on each level in the adaptive MLMC method.

Discretized equations of the form (4.4) and (4.12) are solved iteratively by the classical multigrid method with Gauß-Seidel smoothing. The accuracy condition (4.10) is replaced by the stopping criterion

$$\|\tilde{u}_l^j(\cdot,\omega) - \tilde{u}_l^{j-1}(\cdot,\omega)\|_{H^1(D)} \leq \sigma_{\text{alg}} Tol_l, \tag{5.5}$$

where $\tilde{u}_l^j$ denotes the j-th multigrid iterate and $\sigma_{\text{alg}} = 0.001$ is a safety factor that accounts for estimating the algebraic error $\|u_l(\cdot,\omega) - \tilde{u}_l(\cdot,\omega)\|_{H^1(D)}$ by $\|\tilde{u}_l^j(\cdot,\omega) - \tilde{u}_l^{j-1}(\cdot,\omega)\|_{H^1(D)}$.

The implementation was carried out in the finite element software environment DUNE [19]. As a mesh manager we utilized the module `dune-alugrid` [6]. We used the `dune-subgrid` module [45] for the evaluation of the sum of different approximate evaluations of $u_l(\omega)$ on different grids. Due to the incompatibility of `dune-subgrid` and refinement by bisection implemented in `dune-alugrid` we utilized local "red" mesh refinement [12, 15] with hanging nodes [43, Section 3.1]. Our results show, however, that the adaptive Algorithm 1 converges also when this type of refinement is used.

We used the Dörfler marking strategy (4.21) with $\theta = 0.5$.

For all simulations we set the initial number of samples $M_{\text{init}} = 50$.

The cost for the evaluation of $\tilde{u}_l(\cdot,\omega) \in S_l$ in the uniform case is set to the corresponding number of unknowns $N_l = \dim S_l$ multiplied by the number of multigrid iterations $\nu(\tilde{u}_l(\cdot,\omega))$ performed for computing the solution that fulfills (5.5), i.e.

$$Cost_{\text{un}}(\tilde{u}_l(\cdot,\omega)) = N_l \nu(\tilde{u}_l(\cdot,\omega)) + N_{l-1} \nu(\tilde{u}_{l-1}(\cdot,\omega)).$$

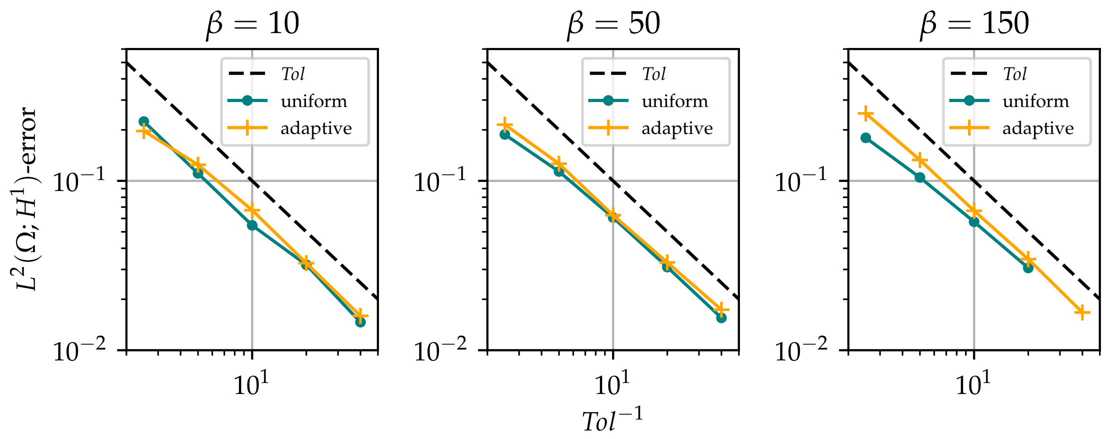

FIGURE 5.1: Errors achieved by uniform and adaptive MLMC against Tol^{-1} for the Poisson problem with random right-hand side.

Let us notice that in the adaptive case seeking a solution that satisfies (5.4) is an iterative process and involves solving the problem (4.12) for all $k = 0, \ldots, k_l(\omega)$. We therefore set the cost for the evaluation of $\tilde{u}_l(\cdot, \omega) \in S_l(\omega)$ to be

$$Cost_{\text{ad}}(\tilde{u}_l(\cdot, \omega)) = \sum_{k=0}^{k_l(\omega)} N_k(\omega)\nu(\tilde{u}_{(k)}(\cdot, \omega)),$$

where $N_k(\omega) = \dim S_{(k)}(\omega)$ and $\nu(\tilde{u}_{(k)}(\cdot, \omega))$ is the number of multigrid iterations performed for computing different $\tilde{u}_{(k)}(\cdot, \omega)$, $k = 0, \ldots, k_l(\omega)$. We note that this definition of the cost does not take into account the cost of computing the error estimators and mesh handling. The computational cost for the adaptive MLMC method with L levels is then given by

$$Cost(E^L[\tilde{u}_L]) = \sum_{l=0}^{L} \sum_{i=1}^{M_l} \sum_{k=0}^{k_l(\omega)} N_{k,i}(\omega)\nu(\tilde{u}_{(k),i}(\cdot, \omega)), \tag{5.6}$$

which reduces to

$$Cost(E^L[\tilde{u}_L]) = N_0 \sum_{i=1}^{M_0} \nu(\tilde{u}_{0,i}(\cdot, \omega)) + \sum_{l=1}^{L} \left(N_l \sum_{i=1}^{M_l} \nu(\tilde{u}_{l,i}(\cdot, \omega)) + N_{l-1} \sum_{i=1}^{M_l} \nu(\tilde{u}_{l-1,i}(\cdot, \omega)) \right)$$

in the case of uniform refinement.

Figure 5.1 illustrates the convergence properties of uniform and adaptive MLMC methods for different values of β by showing the realised error against the inverse of the required tolerance Tol. Here, the error $\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{L^2(\Omega; H_0^1(D))}$ is approximated by a Monte Carlo method utilizing $M = 10$ independent realizations of

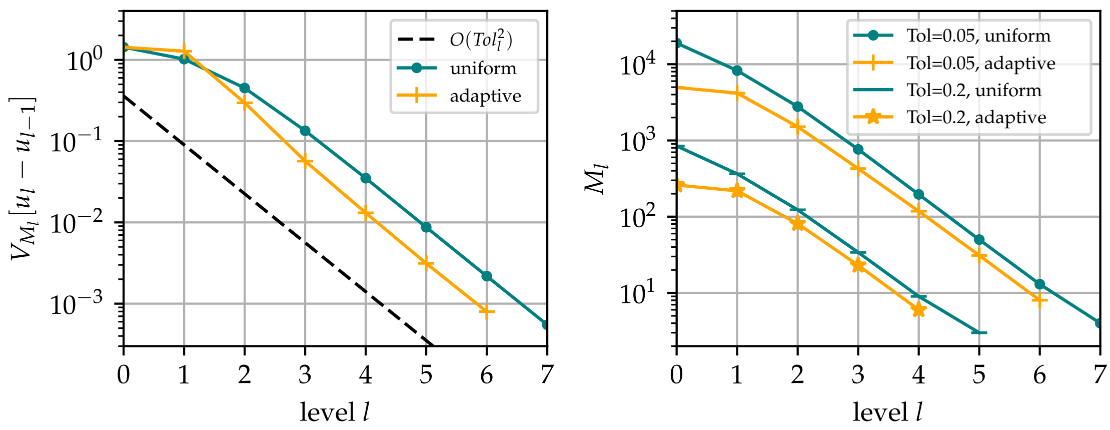

FIGURE 5.2: The values of $V_{M_l}[u_l - u_{l-1}]$ against levels l for $Tol = 0.05$ (left) and the optimal number of samples against levels l for $Tol = 0.05, 0.2$ (right) for uniform and adaptive MLMC for the Poisson problem with random right-hand side with $\beta = 150$.

$\|\mathbb{E}[u] - \mathbb{E}^L[\tilde{u}_L]\|_{H^1(D)}$. For all values of β, both uniform and adaptive MLMC match the required accuracy *Tol* as indicated by the dotted line, thus confirming our theoretical results (cf. Theorems 4.2.1 and 4.3.1) also in this slightly more general case of random boundary conditions. Due to limited memory resources the accessible accuracy of the uniform MLMC is exceeded by the adaptive MLMC for $\beta = 150$.

We now investigate the corresponding computational effort in terms of required numbers of samples and mesh sizes. Figure 5.2 (left) shows an example of the values of $V[u_l - u_{l-1}]$, approximated by sample averages with M_l samples, against levels l for uniform and adaptive MLMC with $Tol = 0.05$ and $\beta = 150$. We illustrate the results only for one value of β, because the qualitative behaviour of the values is the same for all values of β considered in this section. For both methods the values of $V_{M_l}[u_l - u_{l-1}]$ decrease as Tol_l^2 indicated by the dotted line, in agreement with the theoretical results (see Remark 3.1.1). We also see that the values of $V_{M_l}[u_l - u_{l-1}]$ corresponding to the levels of the adaptive MLMC are smaller than the ones corresponding to the levels of the uniform MLMC for almost all values of l. The effect of this difference is seen on the right side of Figure 5.2 showing the optimal numbers of MLMC samples M_l (sometimes smaller than M_{init}), computed according to (4.96), against the corresponding levels $l = 0, \ldots, L$. We note that the optimal numbers of samples corresponding to $Tol = 0.05$ are obtained from the same runs of uniform and adaptive MLMC methods as the results on the left side of the figure. We see that the numbers of samples required for the adaptive MLMC are smaller than for the uniform MLMC. Moreover, in this particular example the adaptive MLMC requires fewer levels than the uniform method. This indicates that the values of $E_{M_L}[\|\tilde{u}_L - \tilde{u}_{L-1}\|_{H^1(D)}]$ (see (4.95)) for the adaptive MLMC method are smaller compared to the values for the uniform MLMC method. We also present the optimal numbers of samples corresponding to $Tol = 0.2$, to illustrate that this effect applies to other values of *Tol* as well.

TABLE 5.1: Average number of unknowns on different levels of uniform and adaptive MLMC with $Tol = 0.05$ for the Poisson problem with random right-hand side, $\beta = 10$.

l	0	1	2	3	4	5
uniform	289	1089	4225	16641	66049	263169
adaptive	289	811	1837	7255	30057	124199

TABLE 5.2: Average number of unknowns on different levels of uniform and adaptive MLMC with $Tol = 0.05$ for the Poisson problem with random right-hand side, $\beta = 50$.

l	0	1	2	3	4	5	6
uniform	289	1089	4225	16641	66049	263169	1050625
adaptive	289	440	1042	2568	8541	34137	141349

Tables 5.1, 5.2 and 5.3 report on the average mesh sizes or, equivalently, the average number of the unknowns $N_{l,i}(\omega)$ among $i = 1, \ldots, M_l$, on the levels $l = 0, \ldots, L$ for one run of the uniform and one run of the adaptive MLMC methods with $Tol = 0.05$ for $\beta = 10$, 50, and 150. The mesh sizes in the adaptive MLMC for all values of β are considerably smaller than for the uniform MLMC and this difference becomes more noticeable for larger values of β. Even though most of the work in MLMC methods is performed on coarser levels, this already indicates a gain of efficiency by adaptive mesh refinement. Note that the adaptive MLMC reached the desired tolerances already on level $L = 6$ in the case $\beta = 150$.

Figure 5.3 presents an example of partitions and corresponding solutions obtained for one sample and different levels in uniform and adaptive MLMC methods. It is clearly seen that adaptive meshes are better suited for problems with locally changing behaviour of solutions.

Finally, we expect from Theorems 4.2.1 and 4.3.1 that the average computational cost of both uniform and adaptive MLMC asymptotically behaves like $\mathcal{O}(L^2 Tol^{-2})$. Figure 5.4 shows the average of $Cost(E^L[\tilde{u}_L])$, as defined in (5.6), against the inverse of the required accuracy Tol together with the asymptotic behaviour (dotted line). The average cost behaves as $\mathcal{O}(Tol^{-2})$ for both uniform and adaptive MLMC methods, which is better than predicted. As in Figure 5.1, the average is taken over $M = 10$ runs. Observe that the adaptive MLMC always outperforms uniform MLMC and that the gain increases with β. These experiments confirm that adaptive MLMC can substantially reduce the computational cost in the presence of random singularities.

TABLE 5.3: Average number of unknowns on different levels of uniform and adaptive MLMC with $Tol = 0.05$ for the Poisson problem with random right-hand side, $\beta = 150$.

l	0	1	2	3	4	5	6	7
uniform	289	1089	4225	16641	66049	263169	1050625	4198401
adaptive	289	364	638	1553	4832	17758	67596	

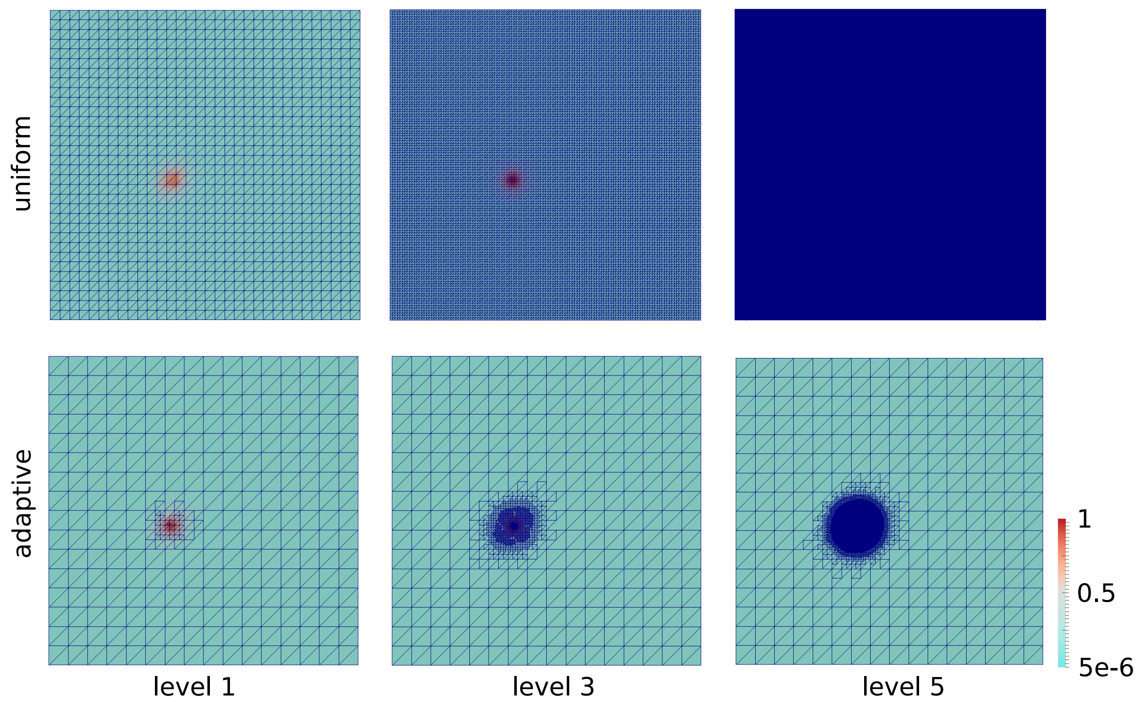

FIGURE 5.3: Examples of partitions $\mathcal{T}_l$ and solutions u_l corresponding to one sample and different levels in uniform and adaptive MLMC methods with $Tol = 0.05$ for the Poisson problem with random right-hand side, $\beta = 150$.

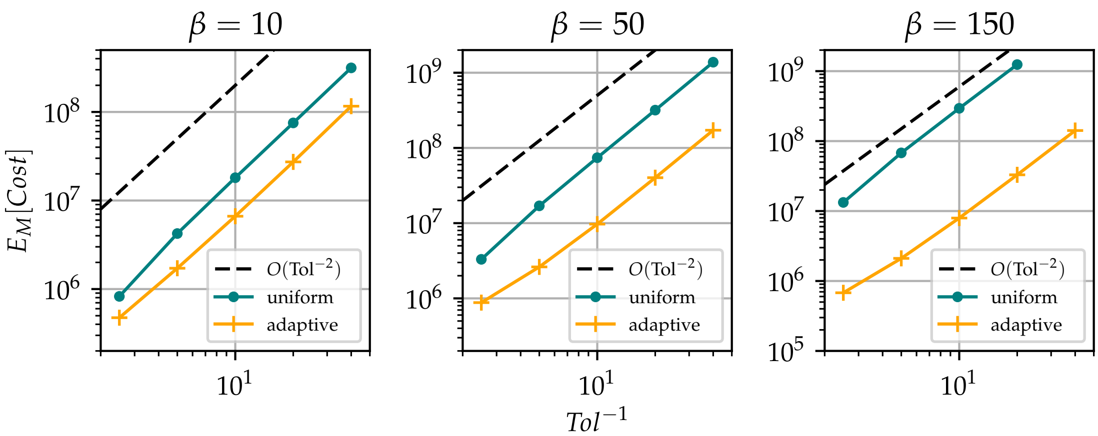

FIGURE 5.4: Average computational cost of uniform and adaptive MLMC against Tol^{-1} for the Poisson problem with random right-hand side.

5.2 Poisson problem with log-normal coefficient and random right-hand side

In this section we consider the problem

$$u(\cdot,\omega) \in H_0^1(D): \quad a(\omega; u(\cdot,\omega), v) = \ell(\omega; v), \quad \forall v \in H_0^1(D) \tag{5.7}$$

with $D = (0,1)^2$ in $d = 2$ space dimensions. The bilinear and the linear forms are again defined as in (2.5). We consider a log-normal coefficient $\alpha(x,\omega) = \exp g(x,\omega)$, where $g(x,\omega)$ is a Gaussian random field with zero mean and exponential covariance function

$$r_g(x,x') = \exp\left(-\frac{\|x-x'\|_1}{\lambda_C}\right), \tag{5.8}$$

where$\|\cdot\|_1$ denotes the l^1-norm in $\mathbb{R}^2$ and λ_C is the correlation length. We fix $\lambda_C = 0.3$. The field α is parametrized with the KL expansion. We follow [29], where analytical eigenvalues and eigenfunctions are found for the covariance operator with the kernel defined by (5.8). We truncate the KL expansion after 40 terms. The uncertain source term f is defined as in (5.2) with $\beta = 150$ and $Y = (Y_1, Y_2)^\top$, where $Y_1, Y_2 \sim \mathcal{U}(0.25, 0.75)$.

Assumption 2.2.1 (i) is satisfied for this problem, we also have Lipschitz continuity of pathwise realizations of α, but we need to assume its Bochner integrability indicated in Assumption 2.2.1 (ii). Assumption 2.2.1 (iii) is fulfilled with $p_f = \infty$. As in the previous section, the solution is unique and we have spatial regularity in the sense that $u(\cdot,\omega) \in H^2(D)$ for almost all $\omega \in \Omega$.

As it is mentioned in Remark 4.2.4, it is possible to obtain a sharper bound for the error provided by the uniform MLMC than presented in this thesis, using duality arguments. The similar error bound shown for the adaptive method is, however, sharp. Improving the error bound for the adaptive MLMC method engaging duality arguments would require goal-oriented error estimation, which stayed out of the scope of this work. However, for some functionals, such as some particular norms of the pathwise solutions, one can expect the adaptive method to perform at least as well as the uniform one. In this section we consider the expected output of interest $\mathbb{E}[Q(u)]$, where $Q(u) = \|u\|_{L^2(D)}$. This output of interest formally fulfills Assumption 2.4.1 with $C_Q = 1$ and $p_Q = \infty$, since

$$|Q(u) - Q(v)| = |\|u\|_{L^2(D)} - \|v\|_{L^2(D)}| \leq \|u-v\|_{L^2(D)} \leq \|u-v\|_{H^1(D)}, \tag{5.9}$$

for any $u, v \in H_0^1(D)$. We are interested in comparing the performance of MLMC finite element methods based on uniform and adaptive refinements.

As in the previous section, pathwise adaptive refinement is performed as described in Section 4.3 with the exact finite element solution replaced by an approximation provided by an iterative method to be described below. The initial partition $\mathcal{T}_0$ consists of 32 congruent triangles for both uniform and adaptive MLMC and

$$Tol_0 := \|\eta_{\mathcal{T}_{(0)}}(\cdot; u_{(0)})\|_{L^2(\Omega)},$$

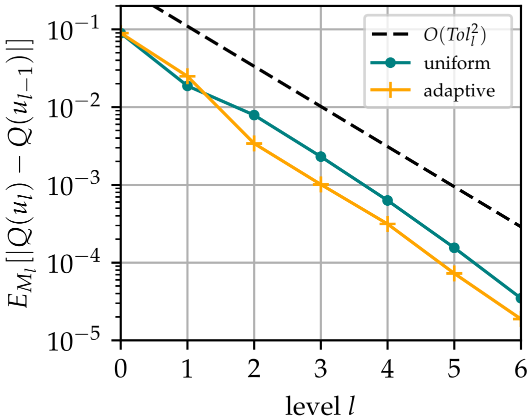

FIGURE 5.5: The values of $E_{M_l}[|Q(u_l) - Q(u_{l-1})|]$ against levels l for $Tol = 0.0001$ for uniform and adaptive MLMC for the problem with log-normal coefficient and random right-hand side.

where we again approximate the $L^2(\Omega)$-norm by a Monte-Carlo method with 1000 samples. The accuracy criterion (4.18) then takes the form of (5.4) with Tol_l defined according to (3.2) with $q = \frac{1}{2}$.

The discretized equations of the form (4.4) and (4.12) are solved iteratively by the classical multigrid method with Gauß-Seidel smoothing and the stopping criterion (5.5).

For this example we used bisection refinement implemented in `dune-alugrid` with 2 bisections applied to each element that is marked for refinement, for both the uniform and adaptive methods.

Recall that, according to Theorem 4.3.8, the value of θ for the Dörfler marking strategy (4.21) that guarantees convergence of the adaptive algorithm (Algorithm 1) is random and depends on $\|\alpha\|_{W^{1,\infty}(D)}$ and $\alpha_{\min}(\omega)$. However, we chose $\theta = 0.5$ to be constant for the results below, for simplicity.

For all simulations we set the initial number of samples to be $M_{\text{init}} = 50$.

We define the computational cost for MLMC methods with L levels by (5.6), as in the previous section.

Figure 5.5 illustrates the behaviour of the expected value of $|Q(u_l) - Q(u_{l-1})|$ approximated by a sample average E_{M_l}, against levels l corresponding to uniform and adaptive MLMC with $Tol = 0.0001$. Although, according to (5.9) we have

$$
\begin{aligned}
|Q(u_l(\omega)) - Q(u_{l-1}(\omega))| &\leq \|u_l(\omega) - u_{l-1}(\omega)\|_{H^1(D)} \\
&\leq \|u_l(\omega) - u(\omega)\|_{H^1(D)} + \|u(\omega) - u_{l-1}(\omega)\|_{H^1(D)} \\
&\leq C_{\text{dis}}(\omega)(1 + q^{-1})Tol_l
\end{aligned}
\tag{5.10}
$$

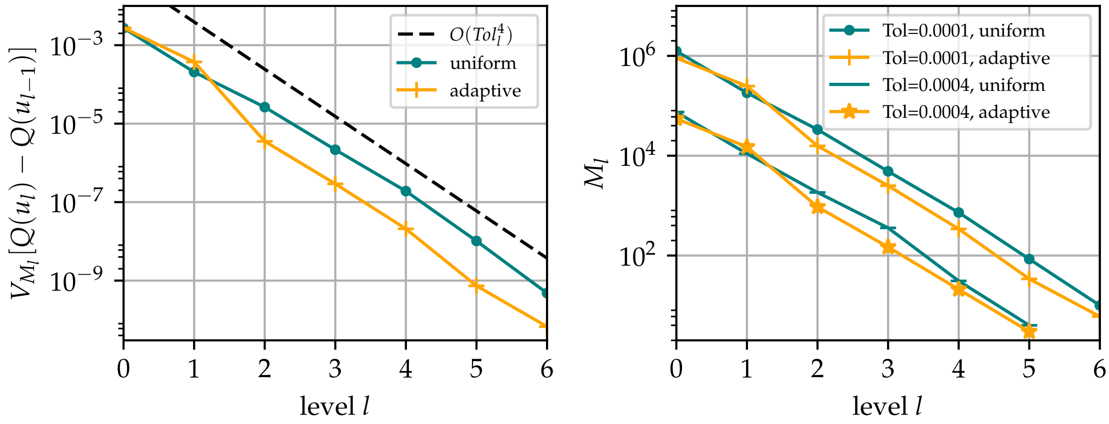

FIGURE 5.6: The values of $V_{M_l}[Q(u_l) - Q(u_{l-1})]$ against levels l for $Tol = 0.0001$ (left) and the optimal number of samples against levels l for $Tol = 0.0001, 0.0004$ (right) for uniform and adaptive MLMC for the problem with log-normal coefficient and random right-hand side.

for almost all $\omega \in \Omega$, which implies $\mathbb{E}[|Q(u_l) - Q(u_{l-1})|] = \mathcal{O}(Tol_l)$, we actually observe that $\mathbb{E}[|Q(u_l) - Q(u_{l-1})|]$ behaves as $\mathcal{O}(Tol_l^2)$. This, however, corresponds to the behaviour, predicted in the uniform case in [86, Proposition 4.4.], where the duality arguments are taken into account

Figure 5.6 (left) shows the values of $\mathbb{V}[Q(u_l) - Q(u_{l-1})]$ approximated by sample averages V_{M_l} (we emphasize that we abuse the notation, denoting both sample approximations to V and $\mathbb{V}$ by V_{M_l} in this chapter) against levels l for uniform and adaptive MLMC with $Tol = 0.0001$. Since

$$\mathbb{V}[Q(u_l) - Q(u_{l-1})] \leq \mathbb{E}[(Q(u_l) - Q(u_{l-1}))^2],$$

the relation (5.10) yields that $\mathbb{V}[Q(u_l) - Q(u_{l-1})] = \mathcal{O}(Tol_l^2)$. However, our numerical results presented in Figure 5.5 show that a stronger result than (5.10) can be obtained and a faster convergence of $\mathbb{V}[Q(u_l) - Q(u_{l-1})]$ can be expected. We observe in Figure 5.6 (left) that for both methods the values of $V_{M_l}[Q(u_l) - Q(u_{l-1})]$ decrease as $\mathcal{O}(Tol_l^4)$, as indicated by the dotted line. This behaviour complies with the numerical results reported in Figure 5.5. We also see that the values of $V_{M_l}[Q(u_l) - Q(u_{l-1})]$ corresponding to the levels of adaptive MLMC are smaller than the ones corresponding to the levels of uniform MLMC for almost all values of l. Figure 5.6 (right) shows the optimal numbers of MLMC samples M_l, computed according to (4.104), against the corresponding levels $l = 0, \dots, L$ for two different values of Tol. We see that the numbers of samples required for adaptive MLMC are in most cases smaller than the numbers required for uniform MLMC.

Table 5.4 presents the average mesh sizes corresponding to different levels $l = 0, \dots, L$ for uniform and adaptive MLMC methods with $Tol = 0.0001$. As in the previous section (see Figure 5.3) we observe a considerable reduction of mesh sizes corresponding to the adaptive MLMC method, compared to the uniform MLMC method.

TABLE 5.4: Average number of unknowns on different levels of uniform and adaptive MLMC with $Tol = 0.0001$ for the problem with log-normal coefficient and random right-hand side.

l	0	1	2	3	4	5	6
uniform	81	289	1089	4225	16641	66049	263169
adaptive	81	146	344	1118	4324	15280	61369

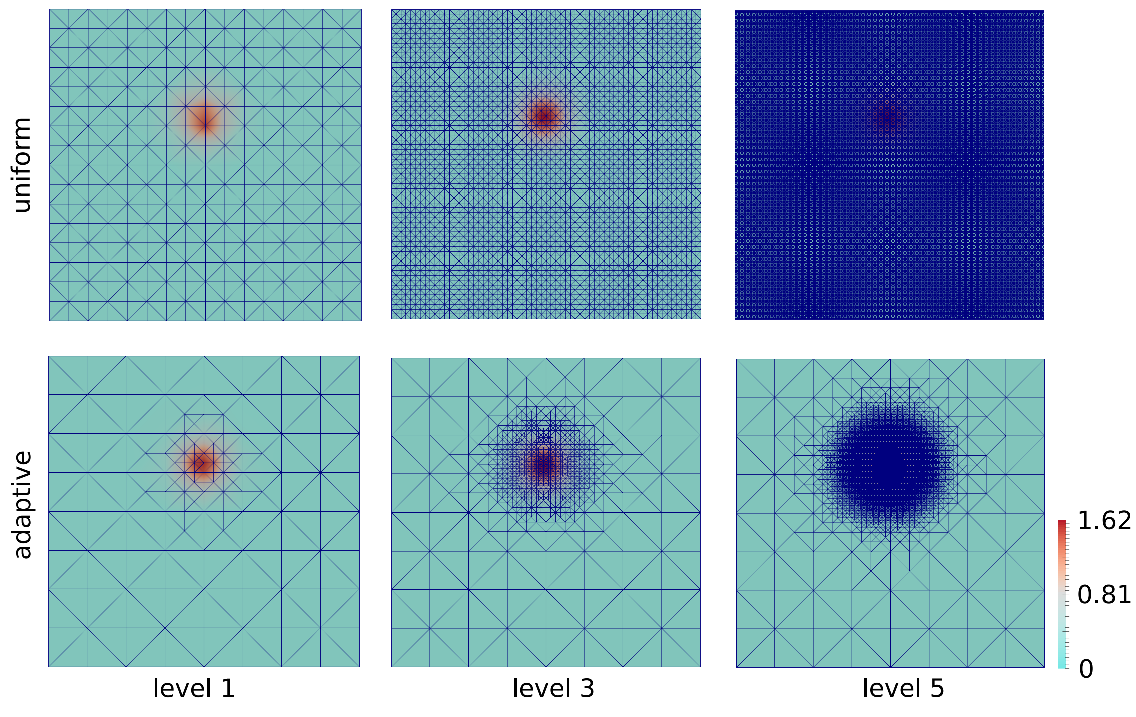

FIGURE 5.7: Examples of partitions $\mathcal{T}_l$ and solutions u_l corresponding to one sample and different levels for the uniform and adaptive MLMC methods applied to the problem (5.7) with Tol=0.0001.

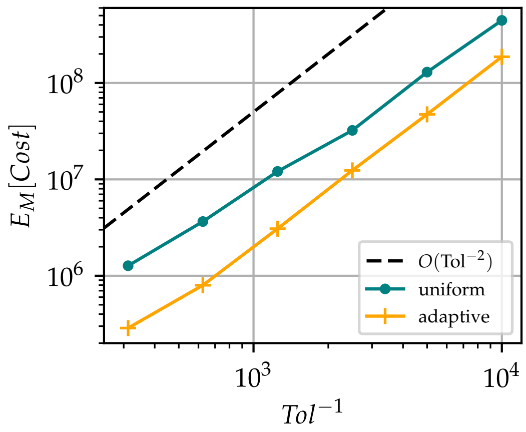

FIGURE 5.8: Average computational cost of uniform and adaptive MLMC against Tol^{-1} for the problem with log-normal coefficient and random right-hand side.

Figure 5.7 illustrates an example of partitions and corresponding solutions obtained for one sample and different levels in uniform and adaptive MLMC methods, which again demonstrates the advantage of adaptive mesh refinement.

Finally, according to Theorems 4.2.2 and 4.3.2 the average computational cost of both uniform and adaptive MLMC asymptotically behaves as $\mathcal{O}(L^2 Tol^{-2})$. However, since we observe a faster convergence of the values $\mathbb{V}[Q(u_l) - Q(u_{l-1})]$, according to e.g. [28, Theorem 4.1] we can expect that the computational cost behaves like $\mathcal{O}(Tol^{-2})$. Figure 5.8 plots the average of $Cost(E^L[\tilde{u}_L])$ as defined in (5.6) over $M = 10$ realizations against the Tol^{-1}, together with the asymptotic behaviour (dotted line), which corresponds to the predicted one. Again, adaptive MLMC outperforms uniform MLMC, which confirms its advantages in the presence of random singularities.

Chapter 6

Uncertainty Quantification in Wear Tests of Knee Implants

In this chapter we present a practically relevant problem that includes uncertain parameters and apply the adaptive MLMC method introduced in Section 4.3 for quantification of uncertainties.

6.1 Motivation

Total knee replacement surgery is one of the common and every day performed operations in Germany. The knee implants used in such surgeries consist of two parts. During the surgery the two implant parts are attached to two bones, namely to the femur and the tibia, replacing the femoro-tibial joint of the human knee. The material properties of the two implant components differ: the femur part is much harder than the tibial plateau, which is usually made of a soft polyethylene. The repeated contact of implant components during knee movements causes the soft material to be worn down, and small polyethylene particles from the tibial component to detach themselves. The detached particles stay in the knee joint which might lead to inflammation and osteolysis [47] which necessitates secondary surgery.

In order to be able to predict the amount of material abrading during daily activities and to keep this amount under the critical values, in vitro testing of the implants is required before they are approved for the market. Such tests are performed in knee wear testing machines, see Figure 6.1. An implant design process might include numerous in vitro tests, each of which is very time consuming (one test might last up to several months) and costly.

Computer simulations of the wear tests can help to accelerate the implant design process. Some of the mechanical tests that must be performed before an implant is approved for use can be replaced by numerical simulations. The run times for these simulations are expected to be considerably less than the time needed for the corresponding mechanical tests. Moreover, the simulations do not require real implants but only their geometry and they can be run unsupervised on a computer, which considerably decreases the expenses required for the implant design.

FIGURE 6.1: A knee wear testing machine, digital photograph, Questmed GmbH, accessed 6 June 2018, <http://www.questmed.de/>.

Several mathematical models for the numerical simulation of the wear tests have been developed in the recent years (e.g. [1, 24, 74]). We concentrate here on the model developed in [24] for a load-controlled testing gait cycle for normal walking, precisely described in the document [56] published by the International Standards Organisation. One of the open problems that appear in the setup of wear tests and, therefore, also in the numerical simulations, is that the initial positioning of the two implant parts with respect to each other is not specified in [56] and is, therefore, a source of uncertainty. It is also not always possible to ensure the precise application of the external forces prescribed in [56]. Additionally, some model parameters might have to be estimated and are, hence, not precise. In this thesis we address these problems by introducing a random formulation of the problem that describes the tests and apply the adaptive MLMC method for approximating the expected wear.

6.2 Mathematical model

In this section we introduce a mathematical model for one type of mechanical wear tests, described in [56]. During the tests the implant parts (see Figure 6.2) are placed into a knee wear testing machine, where external forces are applied to the tibial part in order to imitate the stresses incurred during normal walking. Each such test consists of five million gait cycles. The loads applied to the implants during one gait cycle are given in [56] for 100 time steps. The mass loss of the tibial component is monitored during the tests.

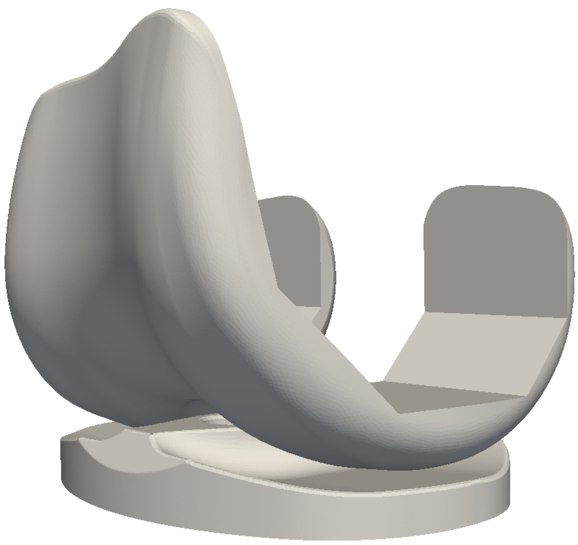

FIGURE 6.2: Femoral (upper) and tibial (lower) components of the knee implants.

We first introduce the deterministic model developed in [24] and then extend it to a random framework that takes into account the uncertainties in the initial position of the implant parts, in the applied forces and in the model parameters.

The femoral and tibial components are modelled as linear elastic bodies. The femoral part is made of a much harder material than the tibial component, and it is reasonable to assume that it behaves as a rigid body. Nevertheless, it is modelled as an elastic body as well, because, as mentioned in [24], the rigid body assumption leads to only little computational savings. Since the deformation of the implant components remains small, the linearised model of elasticity is chosen. Furthermore, a quasi-static approach, assuming that the forces applied to a system vary slowly over time, is utilized. As a result, at each moment of time the contact of two bodies can be described by a system of static equations.

We cite [58, 78] for more detailed introductions to deterministic contact problems. Let us denote the domain occupied by the femoral component by $\overline{D}_1$ and the domain occupied by the tibial part by $\overline{D}_2$. The domains D_1 and D_2 are bounded, open, connected subsets of $\mathbb{R}^3$. The boundaries of the domains ∂D_1 and ∂D_2 are assumed to be piecewise Lipschitz continuous. The boundary ∂D_1 is partitioned into three open disjoint parts

$$\partial D_1 = \overline{\Gamma}_D \cup \overline{\Gamma}_{1,N} \cup \overline{\Gamma}_{1,C},$$

and the boundary ∂D_2 is partitioned into open disjoint parts

$$\partial D_2 = \overline{\Gamma}_R \cup \overline{\Gamma}_{2,N} \cup \overline{\Gamma}_{2,C}.$$

Dirichlet boundary conditions are prescribed on Γ_D, Neumann conditions on $\Gamma_{1,N}$,

$\Gamma_{2,N}$ and Robin type conditions on Γ_R. The boundary parts $\Gamma_{1,C}$ and $\Gamma_{2,C}$ are the contact boundaries, which is where the contact is expected to occur. More precisely, while the actual zone of contact is an unknown of the problem, the model contains the assumption that it is a subset of $\Gamma_{1,C} \cup \Gamma_{2,C}$.

Under applied loads, the two bodies deform and take on new configurations. The deformation of the bodies is specified in linear elasticity by their displacement functions $\mathbf{u}_1 : D_1 \to \mathbb{R}^3$ and $\mathbf{u}_2 : D_2 \to \mathbb{R}^3$.

We consider a homogeneous, isotropic, linearised Saint Venant–Kirchhoff material, i.e the strain-displacement relation corresponding to the small strain assumption is defined as $\boldsymbol{\varepsilon}(\mathbf{u}) := 1/2(\nabla\mathbf{u} + (\nabla\mathbf{u})^\top)$ and the constitutive equation for the stress tensor is given by $\boldsymbol{\sigma}(\mathbf{u}) = \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{u})$, where $\mathbf{C}$ is the forth-order Hooke tensor. The linearised elastic equilibrium for the displacement $\mathbf{u} := (\mathbf{u}_1, \mathbf{u}_2)$ is given by

$$\begin{array}{rcll} -\mathrm{div}\,\boldsymbol{\sigma}(\mathbf{u}) & = & 0 & \text{in} \quad D := D_1 \cup D_2, \\ \mathbf{u} & = & 0 & \text{on} \quad \Gamma_D, \\ \boldsymbol{\sigma}(\mathbf{u})\mathbf{n} + \mathbf{K}\mathbf{u} & = & \mathbf{f}_R & \text{on} \quad \Gamma_R, \\ \boldsymbol{\sigma}(\mathbf{u})\mathbf{n} & = & 0 & \text{on} \quad \Gamma_N := \Gamma_{1,N} \cup \Gamma_{2,N}, \end{array} \tag{6.1}$$

where $\mathbf{n}$ denotes the outer unit normal field on ∂D, $\mathbf{K}$ is the stiffness tensor and $\mathbf{f}_R \in (L^2(\Gamma_N))^3$ is the Neumann load.

In order to model contact between the two bodies, this system is supplemented by a non-penetration constraint on $\Gamma_{k,C}$, $k \in \{1,2\}$. We introduce a homeomorphism $\Phi : \Gamma_{1,C} \to \Gamma_{2,C}$ which we call the *contact mapping*. It forms an a priori identification of points on $\Gamma_{1,C}$ and $\Gamma_{2,C}$ which may come into contact with each other. We take Φ to be the normal projection of $\Gamma_{1,C}$ onto $\Gamma_{2,C}$ and assume that $\Gamma_{1,C}$ and $\Gamma_{2,C}$ are chosen such that construction of Φ is possible. The contact mapping allows to define the initial gap function $g : \Gamma_{1,C} \to \mathbb{R}$ with $g(x) := |\Phi(x) - x|$ and the relative displacement

$$[\mathbf{u}]_\Phi := \mathbf{u}_1|_{\Gamma_{1,C}} - \mathbf{u}_2|_{\Gamma_{2,C}} \circ \Phi,$$

where $\circ$ denotes function composition. The non-penetration condition for two bodies in the framework of small deformations then takes the form

$$[\mathbf{u}]_\Phi \cdot \mathbf{n}_1 \le g \quad \text{on} \quad \Gamma_{1,C}, \tag{6.2}$$

where $\mathbf{n}_1$ is the outer unit normal to ∂D_1.

We set

$$H^1_D(D) := \{v \in H^1(D) \ : \ \gamma_D(v) = 0 \text{ a.e. on } \Gamma_D\},$$

where γ_D denotes the trace map that associates $v \in H^1(D)$ to the restriction $\gamma_D(v)$ in $H^{1/2}(\Gamma_D)$.

The variational formulation of the two-body contact problem (6.1)-(6.2) then takes the form

$$\mathbf{u} \in K : \qquad a(\mathbf{u}, \mathbf{v} - \mathbf{u}) \ge \ell(\mathbf{v} - \mathbf{u}), \qquad \forall \mathbf{v} \in K\,, \tag{6.3}$$

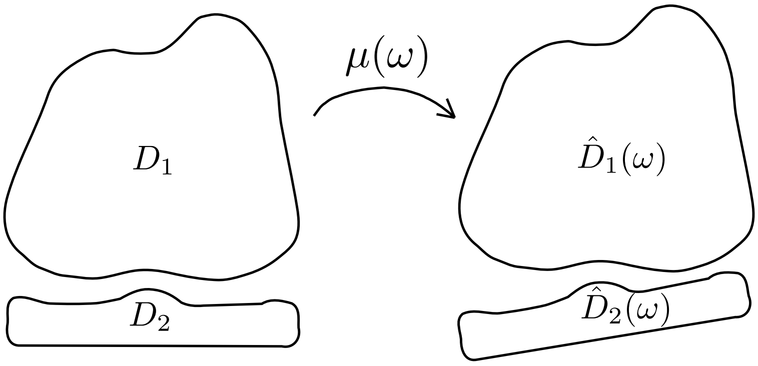

FIGURE 6.3: Transformation of domains.

with the bilinear form

$$a(\mathbf{v},\mathbf{w}) := \int_D \varepsilon(\mathbf{v}) : \mathbf{C} : \varepsilon(\mathbf{w})\, dx + \int_{\Gamma_R} (\mathbf{K}\mathbf{v})\mathbf{w}\, dx, \quad \mathbf{v},\mathbf{w} \in (H^1(D))^3,$$

the linear form

$$\ell(\mathbf{v}) := \int_{\Gamma_R} \mathbf{f}_R \mathbf{v}\, dx, \quad \mathbf{v} \in (H^1(D))^3$$

and the set of admissible displacements

$$K := \{\mathbf{v} \in (H^1_D(D))^3 \ : \ [\mathbf{v}]_\Phi \cdot \mathbf{n}_1 \leq g, \quad \text{a.e. on } \Gamma_{1,C}\}.$$

We note that K is a closed and convex subset of $(H^1_D(D))^3$.

Let us now extend the problem formulation (6.3) to the case when the domains and the external forces are random.

We first assume that we have measurable mappings $\Omega \ni \omega \mapsto \mu_\omega \in \mathrm{SE}(3)$ and $\Omega \ni \omega \mapsto \mathbf{f}_R(\omega) \in (L^2(\mu_\omega \Gamma_R))^3$, where SE(3) is the group of all 3-dimensional rigid body transformations and by $\mu_\omega \Gamma_R$ we denote the boundary transformed by the rigid body motion μ_ω.

We now assume that the position of the tibial component is random and perturbed by a rigid body transformation μ_ω, i.e. it occupies the domain $\overline{\hat{D}}_2(\omega)$, where $\hat{D}_2 = \mu_\omega D_2$. For convenience of notation we define μ_ω as

$$\mu_\omega = \begin{cases} \mathrm{id}\,, & \text{if } x \in D_1, \\ \mu_\omega, & \text{if } x \in D_2, \end{cases}$$

where id is the identity transformation. Then $\hat{D}(\omega) = \mu_\omega D$ and $D_1 = \hat{D}_1 = \mu_\omega D_1$. We note that $\partial D = \mu_\omega^{-1} \partial \hat{D}(\omega)$, $\Gamma_D = \mu_\omega^{-1} \hat{\Gamma}_D(\omega)$ and $\Gamma_R = \mu_\omega^{-1} \hat{\Gamma}_R(\omega)$. In what follows we denote $\Gamma_{k,N}(\omega) := \mu_\omega^{-1} \hat{\Gamma}_{k,N}(\omega)$, $\Gamma_{k,C}(\omega) := \mu_\omega^{-1} \hat{\Gamma}_{k,C}(\omega)$, $k \in \{1,2\}$ and note that the sets $\Gamma_{k,N}$ and $\Gamma_{k,C}$ now also depend on $\omega \in \Omega$, because the contact area is random.

We apply standard techniques in order to formulate the random problem on the reference domain D. The random bilinear form is defined as

$$a(\omega;\mathbf{v},\mathbf{w}) := \int_D \boldsymbol{\varepsilon}_\omega(\mathbf{v}) : \mathbf{C} : \boldsymbol{\varepsilon}_\omega(\mathbf{w})\,dx + \int_{\Gamma_R} (\mathbf{K}\mathbf{v})\mathbf{w}\,dx \quad \mathbf{v},\mathbf{w} \in (H^1(D))^3, \tag{6.4}$$

where $\boldsymbol{\varepsilon}_\omega(\mathbf{v}) := 1/2(J_\mu(\omega)^{-\top}\nabla\mathbf{v} + \nabla\mathbf{v}^\top J_\mu(\omega)^{-1})$ and $J_\mu(\omega)$ is the Jacobian matrix of the transformation μ_ω. We note that $J_\mu(\omega)$ is an orthogonal matrix with determinant 1, because μ_ω and the identity are rigid body transformations.

The linear form is defined as

$$\ell(\omega;\mathbf{v}) := \int_{\Gamma_R} \mathbf{f}_R(\omega) \circ \mu_\omega \mathbf{v}\,dx, \quad \mathbf{v} \in (H^1(D))^3. \tag{6.5}$$

We introduce the contact mapping $\Phi_\omega : \hat{\Gamma}_{1,C}(\omega) \to \hat{\Gamma}_{2,C}(\omega)$ and the corresponding initial gap function $g_\omega : \hat{\Gamma}_{1,C}(\omega) \to \mathbb{R}$ with $g_\omega(x) := |\Phi_\omega(x) - x|$. The relative displacement with respect to Φ_ω is then defined as

$$[\mathbf{u}]_{\Phi_\omega} := \mathbf{u}_1|_{\Gamma_{1,C}(\omega)} - \mathbf{u}_2|_{\Gamma_{2,C}(\omega)} \circ \mu_\omega^{-1} \circ \Phi_\omega \circ \mu_\omega.$$

Then the random variational pathwise formulation for a given $\omega \in \Omega$ reads

$$\mathbf{u}(\cdot,\omega) \in K(\omega) : \qquad a(\omega;\mathbf{u}(\cdot,\omega),\mathbf{v}-\mathbf{u}(\cdot,\omega)) \geq \ell(\omega;\mathbf{v}-\mathbf{u}(\cdot,\omega)), \qquad \forall \mathbf{v} \in K(\omega)\,, \tag{6.6}$$

where

$$K(\omega) := \{\mathbf{v} \in (H_D^1(D))^3 \;:\; [\mathbf{v}]_{\Phi_\omega} \cdot \hat{\mathbf{n}}_1(\omega) \circ \mu_\omega \leq g_\omega \circ \mu_\omega, \text{ a.e. on } \Gamma_{1,C}(\omega)\},$$

and $\hat{\mathbf{n}}_1(\omega)$ is the outer normal to $\partial\hat{D}_1(\omega)$.

Again, for each fixed $\omega \in \Omega$ the set of admissible displacements $K(\omega)$ is a closed and convex subset of $(H^1(D))^3$.

The problem is, therefore, formulated as a random variational inequality, similar to the one introduced in Section 2.3. We do not perform any theoretical analysis for the problem (6.6). We assume this problem is well-posed.

Given partitions $\mathcal{T}_{(k)}(\omega)_{k\in\mathbb{N}\cup\{0\}}$ of the domain D for a fixed $\omega \in \Omega$, we denote the first order finite element subspace of $(H_D^1(D))^3$ defined on $\mathcal{T}_{(k)}(\omega)$ by $S_{(k)}(\omega)$ for any $k \in \mathbb{N}\cup\{0\}$. Discretization of the problem (6.6) in a space $S_{(k)} \subset H_D^1(D)$, together with the mortar approach introduced in [93] for the variationally consistent discretization of the non-penetration condition, leads to the problem

$$\mathbf{u}_{(k)}(\cdot,\omega) \in K_k(\omega) : \;\; a(\omega;\mathbf{u}_{(k)}(\cdot,\omega),\mathbf{v}-\mathbf{u}_{(k)}(\cdot,\omega)) \geq \ell(\mathbf{v}-\mathbf{u}_{(k)}(\cdot,\omega)), \;\; \forall \mathbf{v} \in K_k(\omega), \tag{6.7}$$

where

$$K_k(\omega) := \{\mathbf{v} \in S_{(k)}(\omega) \;:\; \lambda_d([\mathbf{v}]_{\Phi_\omega} \cdot \hat{\mathbf{n}}_1(\omega) \circ \mu_\omega) \leq \lambda_d(g_\omega \circ \mu_\omega),\; \forall \lambda_d \in M_k(\Gamma_{1,C}(\omega))\},$$

the bilinear and the linear forms are defined as in (6.4)-(6.5) and $M_k(\Gamma_{1,C}(\omega))$ is the dual mortar basis corresponding to the space $S_{(k)}(\omega)$, see [93] for the definition and details.

Quantity of interest

The quantity of interest in the simulations of the mechanical tests is the expected wear or, in other words, the expected mass loss of the tibial component. Given a solution $\mathbf{u}(\cdot,\omega)$ to (6.6) for fixed $\omega \in \Omega$, we utilize Archard's wear law [7] as a model for the wear. In the quasi-static context the wear depth on the contact surface is given by [24]

$$w(\mathbf{u}(\cdot,\omega)) = kp(\mathbf{u}(\cdot,\omega))s(\mathbf{u}(\cdot,\omega)), \tag{6.8}$$

where k is a material constant, $p(\mathbf{u}) = -\frac{1}{3}\mathrm{tr}(\sigma(\mathbf{u}))$ is the pressure and s is the sliding distance, i.e. the relative movement of the two implant components. The model parameter k can be estimated from experimental data. Due to uncertainties that appear in the estimation, we model k as a random variable.

The total mass loss is then given by

$$Q_{\text{wear}}(\mathbf{u}(\cdot,\omega)) = \int_{\Gamma_{2,C}(\omega)} w(\mathbf{u}(\cdot,\omega))\, dx,$$

and we are interested in $\mathbb{E}[Q_{\text{wear}}(\mathbf{u})]$. We assume that Q_{wear} fulfills Assumption 2.4.1.

In order to approximate $\mathbb{E}[Q_{\text{wear}}(\mathbf{u})]$ we apply a slightly modified version of the adaptive MLMC-FE method introduced in Section 4.3. It is sometimes difficult to achieve a balance between the statistical and the discretization errors in (3.10) when trying to bound the mean square error of an MLMC-FE method. One of the possible reasons is that high resolution meshes required for balancing the statistical error are not computationally feasible. This is also the case in the simulations of the wear tests. Therefore, instead of the bound for the MSE of the form

$$e(E^L[Q_{\text{wear}}(\tilde{\mathbf{u}}_L)]) \leq Tol$$

we aim at the bound

$$\alpha(\mathbb{E}[Q_{\text{wear}}(\mathbf{u}) - Q_{\text{wear}}(\tilde{\mathbf{u}}_L)])^2 + (1-\alpha)\sum_{l=0}^{L} M_l^{-1}\mathbb{V}[Q_{\text{wear}}(\tilde{\mathbf{u}}_l) - Q_{\text{wear}}(\tilde{\mathbf{u}}_{l-1})] \leq Tol^2 \tag{6.9}$$

with some $\alpha \in [0,1]$. Note that $\alpha = 0$ corresponds to the case when we are only interested in the discretization error and in the case $\alpha = 1$ we are only interested in the statistical error of the applied MLMC method.

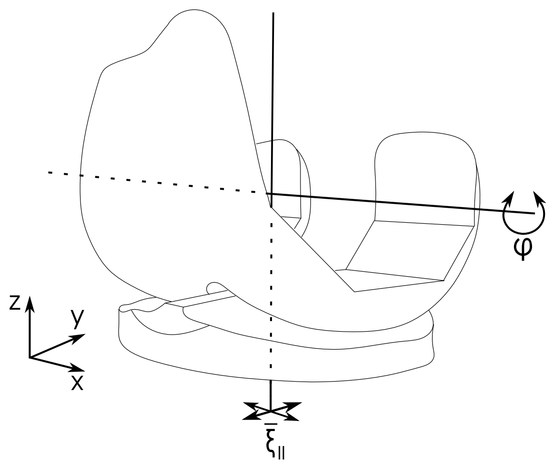

FIGURE 6.4: Illustration of the components characterizing the transformation μ_ω.

6.3 Numerical results

Finally, we report on some implementation details and numerical results for approximating the expected wear.

The values of the material parameters that define the tensor $\mathbf{C}$ (i.e. Young's modulus and Poisson ratio) are taken from [3]. The tensor $\mathbf{K}$ and the boundary patches Γ_D and Γ_R for the femur and tibial components are defined in [56] and discussed in more detail in [24]. We define the random load $\mathbf{f}_R$ by perturbing the values given in [56], which we denote by $\mathbf{f}_R^{\text{iso}}$, for each time step. We set $\mathbf{f}_R = \mathbf{f}_R(\mathbf{f}_R^{\text{iso}}, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6)$, where $\xi_i \sim \mathcal{U}(a_i, b_i)$, $i = 1, \dots, 6$ and the values a_i, b_i, $i = 1, \dots, 6$ are chosen such that the direction and the absolute value of $\mathbf{f}_R^{\text{iso}}$ are perturbed in a suitable range. The maximum absolute deviation of $|\mathbf{f}_R|$ around $\mathbb{E}[|\mathbf{f}_R|] = |\mathbf{f}_R^{\text{iso}}|$ is set to $1 \cdot 10^{-2}|\mathbf{f}_R^{\text{iso}}|$. The first component that defines the transformation μ_ω is a random translation of the tibial (lower) component of the implant in the XY-plane. The values of the vector that defines the translation are uniformly distributed within a circle with radius 3. The second component that defines μ_ω is a rotation with respect to the axis illustrated in Figure 6.4. The values of the angle that defines the rotation are uniformly distributed on the interval $(-3\pi \cdot 10^{-2}, 3\pi \cdot 10^{-2})$. Thus, we set $\mu_\omega = \mu_\omega(\xi_7, \xi_8, \xi_9)$, where $\xi_i \sim \mathcal{U}(a_i, b_i)$ with suitable a_i, b_i, $i = 7, \dots, 9$. A value for the coefficient k from (6.8) can be found in [74]. However, it is not clear, whether this value was obtained from laboratory experiments, from numerical simulations, or from some other source. We, therefore, set $k = \xi_{10}[mm^3/Nm]$, where $\xi_{10} \sim \mathcal{U}(1.75 \cdot 10^{-7}, 2.25 \cdot 10^{-7})$ and $\mathbb{E}[k]$ equals the value presented in [74].

Since solving the discrete problem (6.7) for a fixed $\omega \in \Omega$ for every of the 100 time steps for five million gait cycles is a computationally demanding task, we divide the wear test into 10 blocks. For each of the blocks the finite element solution and the corresponding wear are computed for the first cycle and then the wear is extrapolated for the remaining cycles as described in [24]. For each of the cycle batches the

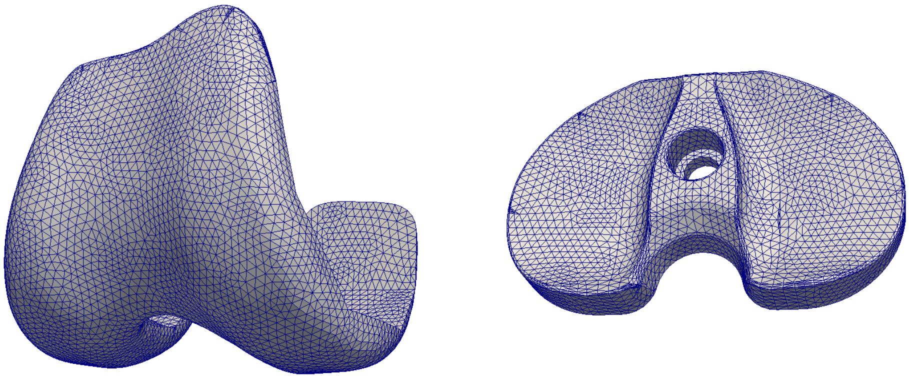

FIGURE 6.5: Initial meshes for the femoral (left) and the tibial (right) components of the knee implants.

geometry of the tibial component is updated according to the computed depth of the wear.

The geometry of the bodies in our simulations corresponds to the commercially sold implant "Genius Pro (Fixed Bearing, PCL retaining)". The tetrahedral domain partitions consisted of 8745 elements for the femoral component and 5038 elements for the tibial component, see Figure 6.5.

The implementation of the deterministic solver producing the values of the wear for a fixed $\omega \in \Omega$ given partitions of the domains D_1 and D_2 was done in DUNE [19]. This implementation was provided by Ansgar Burchardt and Oliver Sander (co-authors of the article [24]), together with the initial partitions that resolve the geometry of the implants. The code was parallelized using OpenMP, allowing for the computation of 100 independent time steps of each cycle in parallel.

The implementation was further extended by adaptive mesh refinement, producing the hierarchies $\mathcal{T}_{(k)}(\omega)$ based on hierarchical error estimation as presented in [78]. We used a heuristic approach for the refinement that was performed once for one whole gait cycle of the wear tests, based on the error indicators averaged over all time steps in the cycle. The refinement for the next cycle was then started from the last refined mesh obtained from the current cycle and not from the initial mesh. This approach allowed for some unnecessary computations to be eliminated. Since the wear only appears on the softer tibial component, the mesh corresponding to the femoral part of the implant was not refined. We used "red-green" mesh refinement as implemented in `dune-uggrid` [79] and set the Dörfler marking parameter $\theta = 0.2$. We assume that the adaptive algorithm based on hierarchical error estimation and the described heuristics fulfills Assumption 4.3.1.

The discretized equations of the form (6.7) were solved using the interior point solver implemented in IPOPT [91]. This solver is not optimal in the sense of computational cost, which behaves only like $\mathcal{O}(N_k^2(\omega))$, where $N_k(\omega) = \dim(S_{(k)}(\omega))$, $k \in \mathbb{N} \cup \{0\}$. We take this into account when computing the numbers of samples corresponding to different levels.

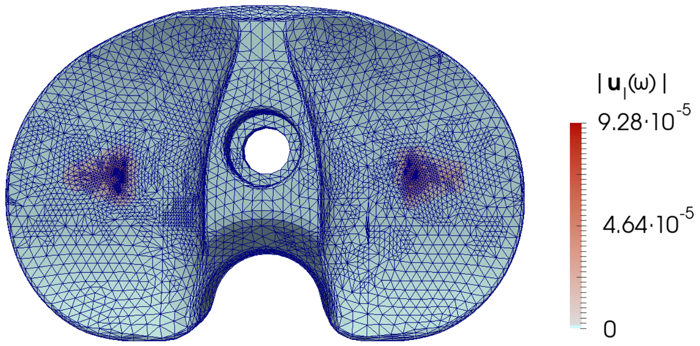

FIGURE 6.6: An adaptively refined mesh corresponding to the tibial component of the knee implant with the values of $|\mathbf{u}_l(\omega)|[m]$, $\omega \in \Omega$, $l = 2$.

We utilized Algorithm 3 for the adaptive MLMC-FE method, where condition (4.103) is replaced by

$$E_{M_L}[|Q_{\text{wear}}(\tilde{\mathbf{u}}_L) - Q_{\text{wear}}(\tilde{\mathbf{u}}_{L-1})|] \leq (q^{-1} - 1)\frac{1}{\sqrt{2\alpha}} Tol$$

and the formula for the number of samples (4.104) is replaced by

$$M_l^{\text{opt}} = 2(1-\alpha)Tol^{-2} \sum_{i=0}^{L} \sqrt{V_{M_i}[Q_{\text{wear}}(\tilde{\mathbf{u}}_i) - Q_{\text{wear}}(\tilde{\mathbf{u}}_{i-1})] E_{M_i}[N_i^2]} \cdot \sqrt{\frac{V_{M_l}[Q_{\text{wear}}(\tilde{\mathbf{u}}_l) - Q_{\text{wear}}(\tilde{\mathbf{u}}_{l-1})]}{E_{M_l}[N_l^2]}},$$

which ensures (6.9).

We set

$$Tol_0 := \|\eta_{\mathcal{T}_0}(\cdot; \mathbf{u}_0)\|_{L^2(\Omega)},$$

where the $L^2(\Omega)$-norm was approximated by a Monte Carlo method with 100 samples. The values of Tol_l were chosen according to (3.2) with $q = \frac{1}{2}$. We assume that the solutions $\mathbf{u}_l(\omega)$, $\omega \in \Omega$, $l \in \mathbb{N} \cup \{0\}$ fulfill Assumption 4.3.2.

The initial number of samples was set to $M_{\text{init}} = 50$, the tolerance $Tol = 0.8$ and the parameter $\alpha = 0.01$. The resulting adaptive MLMC method consisted of $L = 3$ levels with $M_0 = 339$, $M_1 = 286$, $M_2 = 206$, $M_3 = 113$.

The computations corresponding to different samples on different levels were run in parallel on the hybrid x86 and GPU cluster Allegro at Freie Universität Berlin with MLMC post-processing.

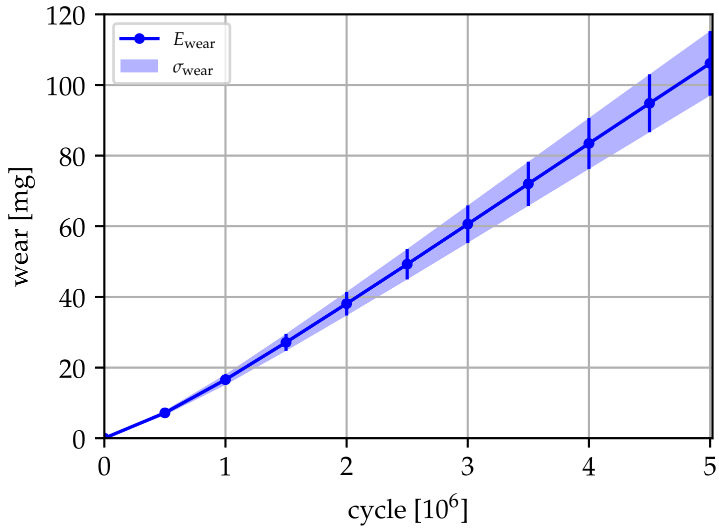

FIGURE 6.7: Wear mean and standard deviation over cycle numbers obtained by the adaptive MLMC method.

Figure 6.6 presents an adaptively constructed mesh and the corresponding absolute value of the solution obtained for one sample by the adaptive algorithm based on hierarchical error estimation described in this section. We see that the areas with higher absolute values of displacement are adaptively refined.

Finally, Figure 6.7 illustrates the expected total wear measured in milligrams over 5 million cycles obtained by the adaptive MLMC method. We present the MLMC approximation of the wear mean $E_{\text{wear}} := E^L[Q_{\text{wear}}(\tilde{\mathbf{u}}_L]$ together with the standard deviation $\sigma_{\text{wear}} := (V_M[Q_{\text{wear}}(\tilde{\mathbf{u}}_L)])^{\frac{1}{2}}$, where V_M is a sample based approximation of the variance. The mean wear lies in the range of results obtained in [24] with different meshes. Hence, providing estimates of the uncertainty, the results presented in Figure 6.7 complement the results presented in [24], where single deterministic problems were solved.

In this chapter, we successfully applied the adaptive MLMC method for uncertainty quantification to a practically relevant problem. Since application of the uniform MLMC method is computationally infeasible for the considered problem due to the dimensionality and the complicated geometry of the spatial domains, this demonstrates a wider applicability of the adaptive MLMC method.

Appendix A

Appendix

In this appendix we list some results from the finite element theory that are used in some of the proofs in Section 4.3.3.

Quasi-interpolation operator

For a partition $\mathcal{T} \in \mathbb{T}$ we denote the set of its interior vertices by $\mathcal{N}_{\text{int}}$. For any $p \in \mathcal{N}_{\text{int}}$ let $\lambda_p \in S(\mathcal{T})$ denote the first order nodal basis function associated to this vertex.

The quasi-interpolation operator $I_{\mathcal{T}} : L^1(D) \to S(\mathcal{T})$ is then defined as

$$I_{\mathcal{T}} v := \sum_{p \in \mathcal{N}_{\text{int}}} \lambda_p \frac{1}{|\phi_p|} \int_{\phi_p} v \, dx, \tag{A.1}$$

where ϕ_p denotes the union of elements of $\mathcal{T}$ sharing the vertex $p \in \mathcal{N}_{\text{int}}$ and $|\cdot|$ denotes the d-dimensional Lebesgue measure.

For all $v \in H_0^1(D)$ and all $T \in \mathcal{T}$ the quasi-interpolation operator satisfies the following local estimates

$$\|v - I_{\mathcal{T}} v\|_{L^2(T)} \leq c_1 h_T \|v\|_{H^1(\phi_T)}, \tag{A.2}$$

$$\|v - I_{\mathcal{T}} v\|_{L^2(\partial T)} \leq c_2 h_T^{\frac{1}{2}} \|v\|_{H^1(\phi_T)}, \tag{A.3}$$

where the constants $c_1, c_2 > 0$ only depend on the shape-regularity parameter $C_{\mathcal{T}}$ of $\mathcal{T}$, see [90].

Bubble functions

For any element $T \in \mathcal{T}$ we define the element bubble function as

$$\psi_T := \alpha_T \prod_{p \in T} \lambda_p, \quad \alpha_T = \begin{cases} 27, & \text{if } T \text{ is a triangle,} \\ 256, & \text{if } T \text{ is a tetrahedron.} \end{cases} \tag{A.4}$$

The face bubble function is defined for any face $E \in \mathcal{E}$ as

$$\psi_E := \beta_E \prod_{p \in E} \lambda_p, \quad \beta_E = \begin{cases} 4, & \text{if } E \text{ is a line segment,} \\ 27, & \text{if } E \text{ is a triangle.} \end{cases} \tag{A.5}$$

For every polynomial degree $k \in \mathbb{N} \cup \{0\}$ the following inverse estimates hold for all polynomials φ of degree k

$$c_{1,k} \|\varphi\|_{L^2(T)} \leq \|\psi_T^{\frac{1}{2}} \varphi\|_{L^2(T)}, \tag{A.6}$$

$$\|\nabla(\psi_T \varphi)\|_{L^2(T)} \leq c_{2,k} h_T^{-1} \|\varphi\|_{L^2(T)}, \tag{A.7}$$

$$c_{3,k} \|\varphi\|_{L^2(E)} \leq \|\psi_E^{\frac{1}{2}} \varphi\|_{L^2(E)}, \tag{A.8}$$

$$\|\nabla(\psi_E \varphi)\|_{L^2(\phi_E)} \leq c_{4,k} h_E^{-\frac{1}{2}} \|\varphi\|_{L^2(E)}, \tag{A.9}$$

$$\|\psi_E \varphi\|_{L^2(\phi_E)} \leq c_{5,k} h_E^{\frac{1}{2}} \|\varphi\|_{L^2(E)}, \tag{A.10}$$

where the constants $c_{1,k}, c_{2,k}, c_{3,k}, c_{4,k}, c_{5,k} > 0$ depend only on the degree k and the shape-regularity parameter $C_{\mathcal{T}}$ of $\mathcal{T}$, see [5].

Inverse inequalities

Let $T \in \mathcal{T}$ and $E \in T$, then for all $v \in S(\mathcal{T})$ there holds,

$$\|\nabla v\|_{L^2(T)} \leq c_3 h_T^{-1} \|v\|_{L^2(T)}, \tag{A.11}$$

$$\|\nabla v\|_{L^2(E)} \leq c_4 h_T^{-\frac{1}{2}} \|\nabla v\|_{L^2(T)}, \tag{A.12}$$

where the constants $c_3, c_4 > 0$ depend only on the shape-regularity of T, see [20].

Properties of L^p-best approximation operator

Let $T \in \mathcal{T}$, $v \in L^\infty(T)$ and φ be a polynomial of order $n \in \mathbb{N} \cup \{0\}$. Let π_m^p denote the operator of L^p-best approximation onto the set of polynomials of order $\leq m$, where $m \geq n$, over T. Then there holds

$$\|(\mathrm{id} - \pi_m^2)(v\varphi)\|_{L^2(T)} \leq \|(\mathrm{id} - \pi_{m-n}^\infty)v\|_{L^\infty(T)} \|\varphi\|_{L^2(T)}. \tag{A.13}$$

For all $v \in L^\infty(T)$ and $m \in \mathbb{N} \cup \{0\}$ there holds

$$\|(\mathrm{id} - \pi_m^2)v\|_{L^\infty(T)} \leq h_T \|\nabla v\|_{L^\infty(T)}, \tag{A.14}$$

see [26] for both results.

Appendix B

Appendix

In this appendix we list some of the results from [25] adapted to the residual error estimators defined in (4.26)-(4.28) for the solution of the random problem (2.4). These results are used in the proof of Theorem 4.3.6. Our goal is to trace how the appearing constants depend on $\omega \in \Omega$. In the case when the dependency of the constants is explicitly stated in the original results in [25], we list the random versions of these results without proofs. For some of the results we only sketch some parts of the original proofs concerning the constants. We adopt the notations of the Section 4.3 here.

The following proposition follows from [25, Proposition 4.12], the discrete reliability (4.32) and the stability property (4.33) of the residual error estimator.

Proposition B.0.1. *Let $\mathcal{T} \in \mathbb{T}$, then for almost all $\omega \in \Omega$ and for all $0 < \theta_0 < \theta^*(\omega) := \left(1 + C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)^{-1}$ there exists some $0 < \kappa(\omega) < 1$ that depends only on C_{stab}, C_{rel}, $\alpha_{\min}(\omega)$, $\|\alpha\|_{W^{1,\infty}(D)}$ and θ_0, such that the implication*

$$\eta^2_{\tilde{\mathcal{T}}}(\omega; u_{\tilde{\mathcal{T}}}(\cdot,\omega)) \le \kappa(\omega)\eta^2_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)) \implies \theta\eta^2_{\mathcal{T}}(\omega; u_{\mathcal{T}}(\cdot,\omega)) \le \sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}} \eta^2_T(\omega; u_{\mathcal{T}}(\cdot,\omega)) \tag{B.1}$$

holds for all $0 < \theta \le \theta_0$ and all refinements $\tilde{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}$.

Proof. We omit the arguments of $u_{\mathcal{T}}$ and $u_{\tilde{\mathcal{T}}}$ for brevity.

Let us fix a $0 < \theta_0 < \theta^*(\omega)$ and define $\kappa(\omega)$ as

$$\kappa(\omega) = \frac{1 - \theta_0\left(1 + (1+\delta^{-1})C_{\text{stab}}^2 C_{\text{rel}}^2 \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)}{1+\delta}. \tag{B.2}$$

Relation (4.40) provides

$$\begin{aligned}\eta^2_{\mathcal{T}}(\omega;u_{\mathcal{T}}) &= \sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}}\eta^2_T(\omega;u_{\mathcal{T}}) + \sum_{T\in\mathcal{T}\cap\tilde{\mathcal{T}}}\eta^2_{T\in\mathcal{T}}(\omega;u_{\mathcal{T}})\\ &\leq \sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}}\eta^2_T(\omega;u_{\mathcal{T}}) + (1+\delta)\sum_{T\in\mathcal{T}\cap\tilde{\mathcal{T}}}\eta^2_{T\in\tilde{\mathcal{T}}}(\omega;u_{\tilde{\mathcal{T}}})\\ &\qquad + (1+\delta^{-1})C^2_{\text{stab}}\|\alpha(\cdot,\omega)\|^2_{W^{1,\infty}(D)}\|u_{\tilde{\mathcal{T}}}-u_{\mathcal{T}}\|^2_{H^1(D)}\end{aligned}$$

with any $\delta > 0$.

Incorporating (4.32), together with the assumption of the implication (B.1), we obtain

$$\begin{aligned}\eta^2_{\mathcal{T}}(\omega;u_{\mathcal{T}}) &\leq (1+\delta)\kappa(\omega)\eta^2_{\mathcal{T}}(\omega;u_{\mathcal{T}}) +\\ &\left(1+(1+\delta^{-1})C^2_{\text{stab}}C^2_{\text{rel}}\frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right)\sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}}\eta^2_T(\omega;u_{\mathcal{T}}).\end{aligned} \tag{B.3}$$

Definition of $\kappa(\omega)$, (B.2) and (B.3) provide the result of the implication (B.1) for all $\theta \leq \theta_0$.

Note that for each θ_0 the parameter δ can be chosen sufficiently large, such that $0 < \kappa(\omega) < 1$. □

The following proposition follows from [25, Lemma 4.14] and the monotonicity of the residual error estimator (4.29).

Proposition B.0.2. *Let $\|\eta_{\mathcal{T}_{\text{opt}}}(\omega;u_{\mathcal{T}_{\text{opt}}}(\cdot,\omega))\|_{\mathbb{B}_s} < \infty$, $\mathcal{T} \in \mathbb{T}$ and assume that for an $\omega \in \Omega$ the implication*

$$\begin{aligned}&\eta^2_{\tilde{\mathcal{T}}}(\omega;u_{\tilde{\mathcal{T}}}(\cdot,\omega)) \leq \kappa(\omega)\eta^2_{\mathcal{T}}(\omega;u_{\mathcal{T}}(\cdot,\omega)) \implies\\ &\qquad\qquad \theta_0(\omega)\eta^2_{\mathcal{T}}(\omega;u_{\mathcal{T}}(\cdot,\omega)) \leq \sum_{T\in\mathcal{T}\setminus\tilde{\mathcal{T}}}\eta^2_T(\omega;u_{\mathcal{T}}(\cdot,\omega))\end{aligned}$$

holds for a particular choice of $0 < \kappa(\omega)$, $\theta_0(\omega) < 1$ and all refinements $\tilde{\mathcal{T}} \in \mathbb{T}$ of $\mathcal{T}$. Then there exists a certain refinement $\tilde{\mathcal{T}}^*(\omega) \in \mathbb{T}$ of $\mathcal{T}$, such that

$$|\mathcal{T}\setminus\tilde{\mathcal{T}}^*(\omega)| \leq c_0(\omega)\|\eta_{\mathcal{T}\text{opt}}(\omega;u_{\mathcal{T}\text{opt}}(\cdot,\omega))\|^{\frac{1}{s}}_{\mathbb{B}_s}\eta_{\mathcal{T}}^{-\frac{1}{s}}(\omega;u_{\mathcal{T}}(\cdot,\omega)), \tag{B.4}$$

where $c_0(\omega)$ depends only on $\kappa(\omega)$ and s. The set $\mathcal{T}\setminus\tilde{\mathcal{T}}^*(\omega)$ also satisfies the Dörfler marking (4.21) *for all $0 < \theta \leq \theta_0(\omega)$.*

The following proposition follows from [25, Lemma 4.7] and the stability (4.33) and the reduction (4.34) properties of the residual error estimator.

Proposition B.0.3. *Let $u_{(k)}(\cdot,\omega)$, $k \in \mathbb{N} \cup \{0\}$, $\omega \in \Omega$ be the solutions obtained by Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28)*, then there holds*

$$\eta^2_{\mathcal{T}_{(k+1)}(\omega)}(\omega; u_{(k+1)}(\cdot,\omega)) \leq \rho_1 \eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)) + c_1(\omega)\|u_{(k+1)}(\cdot,\omega) - u_{(k)}(\cdot,\omega)\|^2_{H^1(D)}, \tag{B.5}$$

where

$$\rho_1 := (1+\delta)(1-(1-\rho_{\mathrm{red}})\theta) > 0, \quad c_1(\omega) := 2(1+\delta^{-1})C^2_{\mathrm{stab}}\|\alpha\|^2_{W^{1,\infty}(D)} > 0,$$

with sufficiently small $\delta > 0$ such that $\rho_1 < 1$.

The following proposition follows from [25, Proposition 4.10], Proposition B.0.3, (4.43) and the reliability property (4.32) of the residual error estimator.

Proposition B.0.4. *Let $u_{(k)}(\cdot,\omega)$, $k \in \mathbb{N} \cup \{0\}$, $\omega \in \Omega$ be the solutions obtained by Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28)*, then there holds*

$$\sum_{i=k+1}^{\infty} \eta^2_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)) \leq c_2(\omega)\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)), \tag{B.6}$$

for all $k \in \mathbb{N}$, where $c_2(\omega) := \left(\rho_1 + c_1(\omega)\frac{C_{\mathrm{qo}}}{\alpha^2_{\min}(\omega)}\right)(1-\rho_1)^{-1}$.

Proof. The estimator reduction (B.5) together with the property (4.43) implies

$$\begin{aligned}
&\sum_{i=k+1}^{N} \eta^2_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)) \leq \\
&\quad \sum_{i=k+1}^{N} \left(\rho_1 \eta^2_{\mathcal{T}_{(i-1)}(\omega)}(\omega; u_{(i-1)}(\cdot,\omega)) + c_1(\omega)\|u_{(i)}(\cdot,\omega) - u_{(i-1)}(\cdot,\omega)\|^2_{H^1(D)}\right) \leq \\
&\quad \sum_{i=k+1}^{N} \rho_1 \eta^2_{\mathcal{T}_{(i-1)}(\omega)}(\omega; u_{(i-1)}(\cdot,\omega)) + c_1(\omega)\frac{C_{\mathrm{qo}}}{\alpha^2_{\min}(\omega)}\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)).
\end{aligned}$$

Taking the limit $N \to \infty$ and rearranging some terms lead to

$$(1-\rho_1)\sum_{i=k+1}^{\infty} \eta^2_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)) \leq \left(\rho_1 + c_1(\omega)\frac{C_{\mathrm{qo}}}{\alpha^2_{\min}(\omega)}\right)\eta^2_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)),$$

which concludes the proof. □

Finally, the last proposition follows from [25, Lemma 4.9] and Proposition B.0.4.

Proposition B.0.5. *Let $u_{(k)}(\cdot,\omega)$, $k \in \mathbb{N} \cup \{0\}$, $\omega \in \Omega$ be the solutions obtained by Algorithm 1 with the residual error estimator defined in* (4.26)-(4.28)*, then for all $s > 0$ there holds*

$$\sum_{i=0}^{k-1} \eta^{-1/s}_{\mathcal{T}_{(i)}(\omega)}(\omega; u_{(i)}(\cdot,\omega)) \leq c_3(\omega)\eta^{-1/s}_{\mathcal{T}_{(k)}(\omega)}(\omega; u_{(k)}(\cdot,\omega)), \tag{B.7}$$

for all $k \in \mathbb{N}$, *where* $c_3(\omega) := (1 + c_2(\omega))^{1/2s}(1 - (1 + c_2^{-1}(\omega))^{-1/2s})^{-1}$.

Proof. According to the proof of [25, Lemma 4.9], the result of Proposition B.0.4 implies that

$$\eta^2_{\mathcal{T}_{(k+i)}}(\omega; u_{(k+i)}(\cdot, \omega)) \leq c_4(\omega)\rho_4^i(\omega)\eta^2_{\mathcal{T}_{(k)}}(\omega; u_{(k)}(\cdot, \omega)) \tag{B.8}$$

holds for all $i, k \in \mathbb{N} \cup \{0\}$, where $0 < \rho_4(\omega) := (1 + c_2^{-1}(\omega))^{-1} < 1$ and $c_4(\omega) := 1 + c_2(\omega) > 0$. In turn, (B.8) implies (B.7) with $c_3(\omega) = c_4(\omega)^{1/2s}(1 - \rho_4^{1/2s}(\omega))^{-1}$, which concludes the proof. □

Appendix C

Appendix

In this appendix we show that some constants obtained in Chapter 4 are finite and are, therefore, well defined.

Proposition C.0.6. *Let Assumption 2.2.1 hold, then the C_{un} defined in Lemma 4.2.1 fulfills $\|C_{\text{un}}\|_{L^p(\Omega)} < \infty$ for all $p \in [1, p_f)$, where p_f is as in Assumption 2.2.1 (iii).*

Proof. By the definition of C_{un}, it can be represented as

$$C_{\text{un}}(\omega) = C_{\alpha,1}(\omega)\|f(\cdot,\omega)\|_{L^2(D)}, \tag{C.1}$$

where

$$C_{\alpha,1}(\omega) \lesssim \left(\frac{\alpha_{\max}(\omega)}{\alpha_{\min}(\omega)}\right)^{\frac{1}{2}} \left(\frac{1}{\alpha_{\min}(\omega)} + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)}\right).$$

Assumptions 2.2.1 (i-ii), Hölder's and the triangle inequalities imply that $C_{\alpha,1} \in L^q(\Omega)$ for all $q \in [1, \infty)$. Indeed, we have

$$\left\| \left(\frac{\alpha_{\max}}{\alpha_{\min}}\right)^{\frac{1}{2}} \left(\frac{1}{\alpha_{\min}} + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha^2_{\min}}\right) \right\|_{L^q(\Omega)} \leq \left\| \frac{\alpha_{\max}}{\alpha_{\min}} \right\|^{\frac{1}{2}}_{L^q(\Omega)} \left\| \frac{1}{\alpha_{\min}} + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha^2_{\min}} \right\|_{L^{2q}(\Omega)} \leq$$
$$\|\alpha_{\max}\|^{\frac{1}{2}}_{L^{2q}(\Omega)} \left\| \frac{1}{\alpha_{\min}} \right\|^{\frac{1}{2}}_{L^{2q}(\Omega)} \left(\left\| \frac{1}{\alpha_{\min}} \right\|_{L^{2q}(\Omega)} + \|\alpha\|_{L^{4q}(\Omega;W^{1,\infty}(D))} \left\| \frac{1}{\alpha_{\min}} \right\|^2_{L^{8q}(\Omega)} \right) < \infty$$

for all finite $q \geq 1$. Then (C.1) and Hölder's inequality lead to

$$\|C_{\text{un}}\|_{L^p(\Omega)} \leq \|C_{\alpha,1}\|_{L^q(\Omega)} \|f\|_{L^{p_f}(\Omega;L^2(D))},$$

where $\frac{1}{p} = \frac{1}{q} + \frac{1}{p_f}$, which provides the claim of the proof. □

Proposition C.0.7. *Let Assumption 2.2.1 hold and let $p_f > d$, then the C_{H^2} from Theorem 4.3.7 fulfills $\|C^d_{H^2}\|_{L^p(\Omega)} < \infty$ for all $p \in [1, \frac{p_f}{d})$.*

Proof. By the definition of C_{H^2}, it can be represented as

$$C_{H^2}(\omega) = C_{\alpha,2}(\omega)\|f(\cdot,\omega)\|_{L^2(D)}, \tag{C.2}$$

where, using the formula for $\theta(\omega)$, (4.74) and the relation $s = \frac{1}{d}$, we have

$$C_{\alpha,2}(\omega) \lesssim \left(1 + \left(1 + \frac{\|\alpha\|^2_{W^{1,\infty}(D)}}{\alpha^2_{\min}(\omega)} + \frac{\|\alpha\|^4_{W^{1,\infty}(D)}}{\alpha^4_{\min}(\omega)} + \frac{\|\alpha\|^6_{W^{1,\infty}(D)}}{\alpha^6_{\min}(\omega)}\right)^{\frac{1}{d}+\frac{1}{2}}\right)\cdot \\ \left(1 + \frac{\|\alpha\|_{W^{1,\infty}(D)}}{\alpha_{\min}(\omega)} + \frac{\|\alpha\|^{3/2}_{W^{1,\infty}(D)}}{\alpha^{3/2}_{\min}(\omega)} + \frac{\|\alpha\|^{5/2}_{W^{1,\infty}(D)}}{\alpha^{5/2}_{\min}(\omega)}\right).$$

Assumptions 2.2.1 (i-ii), Hölder's and the triangle inequalities imply that $C^d_{\alpha,2} \in L^q(\Omega)$ for all $q \in [1,\infty)$ and $d = 1,2,3$ (see the proof of Proposition C.0.6 for a similar implication). Then (C.2) and Hölder's inequality lead to

$$\|C^d_{H^2}\|_{L^p(\Omega)} \leq \left\|C^d_{\alpha,2}\right\|_{L^q(\Omega)} \|f\|^d_{L^{sd}(\Omega;L^2(D))},$$

where $\frac{1}{p} = \frac{1}{q} + \frac{1}{s}$ and $p < s$, since q can take any finite values larger than 1. Since $f \in L^p(\Omega)$ for $p \leq p_f$, we require $sd \leq p_f$, which implies $p < s \leq \frac{p_f}{d}$. $\square$

Zusammenfassung

Die Quantifizierung von Unsicherheiten ist heutzutage ein wichtiger Teil in vielen Anwendungen aus der Mathematik, da mathematische Modelle oft unbekannte Parameter erhalten, die im Allgemeinen nicht genau durch Experimente bestimmt werden können. In dieser Arbeit beschäftigen wir uns mit Problemen, denen elliptische partielle Differentialgleichungen mit zufälligen hoch-dimensionalen Koeffizienten zu Grunde liegen. Wir interessieren uns für Annäherungen von dem Erwartungswert von der Problemslösung oder einem Output des Interesses, welcher durch ein Lipschitz-kontinuierliches Funktional von der Lösung gegeben ist.

Monte Carlo Methoden sind Sampling Verfahren, deren Rechenkosten dimensionsunabhängig sind. Deswegen sind diese in Kombination mit den Methoden der Finiten Elemente für die in dieser Arbeit betrachteten Probleme gut geeignet. Allerdings, ist die Konvergenz von den Monte Carlo Methoden langsam und deswegen werden für eine hohe Genauigkeit viele Stichproben benötigt. Da jede Stichprobe einer partiellen Differentialgleichungen entspricht, deren Lösung oft kostspielig ist, sind die Rechenkosten von den Monte Carlo Finite Elemente Methoden die Hauptschwäche von diesen Verfahren. Multilevel Monte Carlo Finite Elemente Methoden, in denen die Stichproben auf Ebenen einer Gitterhierarchie verteilt werden, erben die Vorteile von den Monte Carlo Methoden, sind jedoch viel effizienter, da ihre theoretischen Rechenkosten assymptotisch mit den Rechenkosten für ein deterministisches Problem, das einer Stichprobe entspricht, vergleichbar sind.

In dieser Arbeit stellen wir eine Verallgemeinerung von den multilevel Monte Carlo Verfahren vor, in welcher die Ebenen durch eine Abfolge von Toleranzen charakterisiert sind. Wir präsentieren ein abstraktes Framework für das neue so-genannte adaptive multilevel Monte Carlo Verfahren, zusammen mit einer Reihe von Annahmen, die zur Konvergenz und zu optimalen Rechenkosten des Verfahrens führt. Wir zeigen weiter, dass das klassische multilevel Monte Carlo Finite Elemente Verfahren als Sonderfall von dem neuen adaptiven Verfahren betrachtet werden kann. Der Ansatz des neuen Verfahrens ermöglicht außerdem die Anwendung von a posteriori Fehlerschätztechniken, welche die Lösung von den deterministischen Problemen der Stichproben beschleunigen können. Wir präsentieren eine adaptive multilevel Monte Carlo Finite Elemente Methode für lineare elliptische Probleme, die auf a posteriori Residuum-Fehlerschätzung basiert. In einer theoretischen Analyse zeigen wir dann, dass dieses Verfahren alle Annahmen des abstrakten Frameworks erfüllt. Vorgeführte numerische Beispiele demonstrieren die Vorteile von der neuen adaptiven multilevel Monte Carlo Finite Elemente Methode im Vergleich mit dem klassischen Verfahren.

Schließlich, präsentieren wir die Ergebnisse einer Anwendung des neuartigen adaptiven Verfahrens für die Quantifizierung von Unsicherheiten in Abriebtests von Knieimplantaten und demonstrieren damit dessen breitere Anwendbarkeit.

List of Abbreviations

a.e.	**a**lmost **e**verywhere
a.s.	**a**lmost **s**urely
i.i.d.	**i**ndependent **i**dentically **d**istributed
FE	**F**inite **E**lement
KL	**K**arhunen-**L**oève
MC	**M**onte **C**arlo
MLMC	**M**ulti**l**evel **M**onte **C**arlo
MLMC-FE	**M**ulti**l**evel **M**onte **C**arlo – **F**inite **E**lement
MSE	**M**ean **S**quare **E**rror
PDE	**P**artial **D**ifferential **E**quation

List of Symbols

$\|\cdot\|_a$	energy norm	36
$\mathbb{A}_s$	approximation class	53
$\mathcal{A}$	σ-algebra on Ω	6
$a(\omega;\cdot,\cdot)$	bilinear form	8
$\mathbb{B}_s$	approximation class	49
b	number of bisections	27
C_{ad}	ω-dependent constant in (4.14)	31
C_{conv}	ω-dependent constant in Theorem 4.3.3	43
C_{cost}	constant in Assumptions 3.1.2	17
C_{dis}	ω-dependent constant in Assumptions 3.1.1 and 3.2.1	17
C_{eq}	ω-dependent constant in (4.75)	53
C_{H^2}	ω-dependent constant in Theorem 4.3.7	54
$C^k(\overline{D})$	space of k times continuously differentiable functions	6
$C^{k,r}(\overline{D})$	Hölder space	6
C_{mesh}	constant in (4.22)	35
C_{MG}	ω-dependent multigrid constant	29
C_{min}	constant in Remark 4.3.1	34
C_{opt}	ω-dependent constant in (4.63)	50
C_{qo}	constant in (4.43)	42
C_{rel}	constant in (4.32)	40
C_{stab}	constant in (4.33) and (4.34)	40
$C_{\mathcal{T}}$	shape-regularity parameter of $\mathcal{T}$	26
C_Q	ω-dependent Lipschitz constant in Assumption 2.4.1	11
C_{un}	ω-dependent constant in (4.6)	27
C_η	constant in (4.86)	56
$Cost(\cdot)$	computational cost	17
c_a	ω-dependent coercivity constant	8
D	spatial open, bounded, convex, polyhedral domain	5
d	space dimension	5
$\mathbb{E}[\cdot]$	expected value	6
$\mathcal{E}$	set of interior faces of $\mathcal{T}$	25

This is an incomplete list of symbols that appear in Chapters 2-4.

E	face in $\mathcal{E}$	25
$E^L[\cdot]$	MLMC estimator	17
$E_M[\cdot]$	MC estimator	15
f	random source function	7
$H^k(D)$	Sobolev space	5
$H_0^1(D)$	Sobolev space with zero trace on ∂D	5
h_E	size of $E \in \mathcal{E}$	25
h_k	maximum element size of $\mathcal{T}_{(k)}$	27
h_T	size of $T \in \mathcal{T}$	26
k_l	refinement level corresponding to uniform MLMC level l	28
$k_l(\omega)$	refinement level corresponding to adaptive MLMC level l	32
$L^p(D)$	Lebesgue space	5
$L^p(\Omega; B)$	Bochner space of B-valued random variables	7
$L^p(\Omega)$	Bochner space of $\mathbb{R}$-valued random variables	7
$\ell(\omega; \cdot)$	linear form	8
$\mathcal{M}_{(k)}(\omega)$	set of marked elements defined in Algorithm 1	34
$\mathbf{n}_E$	arbitrary unit normal to E	37
$\mathbf{n}_T$	outer unit normal to ∂T	37
$osc_{\mathcal{T}}(\omega; \cdot)$	oscillations associated with mesh $\mathcal{T}$	44
$osc_T(\omega; \cdot)$	oscillations associated with element T	44
$\mathbb{P}$	probability measure	6
p_{ad}	Lebesgue space order for C_{ad}	31
p_{dis}	Lebesgue space order for C_{dis}	21
p_f	Lebesgue space order for f	8
p_Q	Lebesgue space order for C_Q	11
$Q(\omega; \cdot)$	functional that defines an output of interest	10
q	MLMC tolerance reduction factor	16
$R_T(\omega; \cdot)$	element residual	38
$R_E(\omega; \cdot)$	jump residual	38
$S_{(k)}$	first order finite element space defined on $\mathcal{T}_{(k)}$	27
$S_{(k)}(\omega)$	first order finite element space defined on $\mathcal{T}_{(k)}(\omega)$	31
S_l	finite element space corresponding to uniform MLMC level l	28
$S_l(\omega)$	finite element space corresponding to adaptive MLMC level l	32
$S_{\mathcal{T}}$	first order finite element space defined on $\mathcal{T}$	36

$\mathbb{T}$	set of admissible refinements of $\mathcal{T}_{(0)}$	35
$\mathcal{T}$	partition of D	25
$\|\mathcal{T}\|$	number of elements in $\mathcal{T}$	26
$\mathcal{T}_{(k)}$	k times uniformly refined partition of $\mathcal{T}_{(0)}$	26
$\mathcal{T}_{(k)}(\omega)$	k times adaptively refined partition of $\mathcal{T}_{(0)}$	31
$\mathcal{T}_{\text{opt}}$	optimal partition	49
T	element of $\mathcal{T}$	25
Tol_l	MLMC tolerance	16
$u(\omega;\cdot)$	$\in H_0^1(D)$, solution to a random PDE	8, 10
$u_{(k)}(\omega;\cdot)$	$\in S_{(k)}$ or $\in S_{(k)}(\omega)$, finite element solution	27, 31
$u_l(\omega;\cdot)$	$\in S_l$ or $\in S_l(\omega)$, finite element solution	28, 32
$\tilde{u}_l(\omega;\cdot)$	approximation to $u(\omega;\cdot)$ in abstract MLMC	16
$\tilde{u}_l(\omega;\cdot)$	algebraic approximation to $u_l(\omega;\cdot)$	28
$u_{\mathcal{T}}(\omega;\cdot)$	$\in S_{\mathcal{T}}$, finite element solution	36
$\mathbb{V}[\cdot]$	variance	6
$V[\cdot]$	variance-like operator	16
$W^{k,\infty}(D)$	Sobolev space	6
α	random coefficient	7
α_{max}	ω-dependent upper bound of α	7
α_{min}	ω-dependent lower bound of α	7
Γ	shape-regularity parameter for sequence $\mathcal{T}_0 \subset \mathcal{T}_1 \subset \dots$	26
γ	order in Assumption 3.1.2	17
$\eta_{\mathcal{T}}(\cdot)$	error estimator corresponding to partition $\mathcal{T}$	31
$\eta_{\mathcal{T}}(\omega;\cdot)$	residual-based error estimator corresponding to partition $\mathcal{T}$	38
$\eta_T(\cdot)$	error indicator corresponding to element T	34
$\eta_T(\omega;\cdot)$	residual-based error indicator corresponding to element T	38
θ	Dörfler marking parameter	34
ξ	real-valued random variable	6
ϕ_E	union of elements sharing face $E \in \mathcal{E}$	26
ϕ_T	union of elements sharing at least one vertex with $T \in \mathcal{T}$	26
ρ_{conv}	ω-dependent constant in Theorem 4.3.3	43
ρ_{red}	ω-dependent constant in (4.34)	40
Ω	set of events	6
ω	event from Ω	6

Bibliography

[1] A. Abdelgaied et al. "Computational wear prediction of artificial knee joints based on a new wear law and formulation". In: *Journal of Biomechanics* 44(6) (2011), pp. 1108–1116. DOI: 10.1016/j.jbiomech.2011.01.027.

[2] A. Abdulle, A. Barth, and C. Schwab. "Multilevel Monte Carlo Methods for Stochastic Elliptic Multiscale PDEs". In: *Multiscale Modeling and Simulation* 11(4) (2013), pp. 1033–1070. DOI: 10.1137/120894725.

[3] C. Abicht. "Künstliche Kniegelenke nach dem Viergelenkprinzip: Konzeption, Design und tribologische Eigenschaften einer neuen Knieendoprothese mit naturnaher Gelenkgeometrie". PhD thesis. Ernst-Moritz-Arndt-Universität Greifswald, 2005. URL: https://epub.ub.uni-greifswald.de/frontdoor/index/index/year/2006/docId/62.

[4] R. J. Adler. *The Geometry of Random Fields*. New York: Springer-Verlag, 1981. DOI: 10.1137/1.9780898718980.

[5] M. Ainsworth and J. T. Oden. "A posteriori error estimation in finite element analysis". In: *Computer Methods in Applied Mechanics and Engineering* 142(1-2) (1997), pp. 1–88. DOI: 10.1016/S0045-7825(96)01107-3.

[6] M. Alkämper et al. "The DUNE-ALUGrid Module". In: *Archive of Numerical Software* 4(1) (2016), pp. 1–28. DOI: 10.11588/ans.2016.1.23252.

[7] J. F. Archard and W. Hirst. "The wear of metals under unlubricated conditions". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 236(1206) (1956), pp. 397–410. DOI: 10.1098/rspa.1956.0144.

[8] M. Aurada et al. "Efficiency and Optimality of Some Weighted-Residual Error Estimator for Adaptive 2D Boundary Element Methods". In: *Computational Methods in Applied Mathematics* 13(3) (2013), pp. 305–332. DOI: 10.1515/cmam-2013-0010.

[9] I. Babuška, F. Nobile, and R. Tempone. "A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data". In: *SIAM Journal on Numerical Analysis* 45(3) (2007), pp. 1005–1034. DOI: 10.1137/050645142.

[10] I. Babuška, R. Tempone, and G. E. Zouraris. "Galerkin Finite Element Approximations of Stochastic Elliptic Partial Differential Equations". In: *SIAM Journal on Numerical Analysis* 42(2) (2006), pp. 800–825. DOI: 10.1137/S0036142902418680.

[11] L. Badea. "Global convergence rate of a standard multigrid method for variational inequalities". In: *IMA Journal of Numerical Analysis* 34(1) (2014), pp. 197–216. DOI: 10.1093/imanum/drs054.

[12] R. E. Bank, A. H. Sherman, and A. Weiser. "Refinement Algorithms and Data Structures for Regular Local Mesh Refinement". In: *Scientific Computing*. Ed. by R. Stepleman. North Holland: IMACS, 1983, pp. 3–17.

[13] A. Barth, A. Lang, and C. Schwab. "Multilevel Monte Carlo method for parabolic stochastic partial differential equations". In: *BIT Numerical Mathematics* 53(1) (2013), pp. 3–27. DOI: 10.1007/s10543-012-0401-5.

[14] A. Barth, C. Schwab, and N. Zollinger. "Multi-level Monte Carlo Finite Element method for elliptic PDEs with stochastic coefficients". In: *Numerische Mathematik* 119(1) (2011), pp. 123–161. DOI: 10.1007/s00211-011-0377-0.

[15] J. Bey. "Simplicial grid refinement: on Freudenthal's algorithm and the optimal number of congruence classes". In: *Numerische Mathematik* 85(1) (2000), pp. 1–29. DOI: 10.1007/s002110050475.

[16] C. Bierig and A. Chernov. "Convergence analysis of multilevel Monte Carlo variance estimators and application for random obstacle problems". In: *Numerische Mathematik* 130(4) (2015), pp. 579–613. DOI: 10.1007/s00211-014-0676-3.

[17] P. Binev, W. Dahmen, and R. DeVore. "Adaptive Finite Element Methods with convergence rates". In: *Numerische Mathematik* 97(2) (2004), pp. 219–268. DOI: 10.1007/s00211-003-0492-7.

[18] P. Binev et al. "Approximations Classes for Adaptive Methods". In: *Serdica Mathematical Journal* 28 (2002), pp. 391–416. URL: http://www.math.bas.bg/serdica/2002/2002-391-416.pdf.

[19] M. Blatt et al. "The Distributed and Unified Numerics Environment, Version 2.4". In: *Archive of Numerical Software* 4(100) (2016), pp. 13–29. DOI: 10.11588/ans.2016.100.26526.

[20] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. 3rd ed. Cambridge University Press, 2007. DOI: 10.1017/CBO9780511618635.

[21] D. Braess, C. Carstensen, and R. H. W. Hoppe. "Convergence analysis of a conforming adaptive finite element method for an obstacle problem". In: *Numerische Mathematik* 107(3) (2007), pp. 455–471. DOI: 10.1007/s00211-007-0098-6.

[22] S. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. 3rd ed. Springer, 2008. DOI: 10.1007/978-0-387-75934-0.

[23] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. New York: Springer, 2010. DOI: 10.1007/978-0-387-70914-7.

[24] A. Burchardt, C. Abicht, and O. Sander. *Simulating Wear On Total Knee Replacements*. arXiv:1704.08307 [q-bio.TO]. 2017. URL: https://arxiv.org/abs/1704.08307.

[25] C. Carstensen et al. "Axioms of adaptivity". In: *Computers and Mathematics with Applications* 67(6) (2014), pp. 1195–1253. DOI: 10.1016/j.camwa.2013.12.003.

[26] J. M. Cascon et al. "Quasi-Optimal Convergence Rate for an Adaptive Finite Element Method". In: *SIAM Journal on Numerical Analysis* 46(5) (2008), pp. 2524–2550. DOI: 10.1137/07069047X.

[27] J. Charrier. "Strong and Weak Error Estimates for Elliptic Partial Differential Equations with Random Coefficients". In: *SIAM Journal on Numerical Analysis* 50(1) (2011), pp. 216–246. DOI: 10.1137/100800531.

[28] J. Charrier, R. Scheichl, and A. L. Teckentrup. "Finite Element Error Analysis of Elliptic PDEs with Random Coefficients and Its Application to Multilevel

Monte Carlo Methods". In: *SIAM Journal on Numerical Analysis* 51(1) (2012), pp. 322–352. DOI: 10.1137/110853054.

[29] K. A. Cliffe et al. "Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients". In: *Computing and Visualization in Science* 14(3) (2011), pp. 3–15. DOI: 10.1007/s00791-011-0160-x.

[30] N. Collier et al. "A continuation multilevel Monte Carlo algorithm". In: *BIT Numerical Mathematics* 55(2) (2015), pp. 399–432. DOI: 10.1007/s10543-014-0511-3.

[31] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Classics in Applied Mathematics. SIAM, 1992. DOI: 10.1137/1.9780898719000.

[32] R. Courant and D. Hilbert. *Methods of Mathematical Physics, vol. I*. New York: John Wiley and Sons, 1989. DOI: 10.1002/9783527617234.

[33] G. Detommaso, T. Dodwell, and R. Scheichl. *Continuous Level Monte Carlo and Sample-Adaptive Model Hierarchies*. arXiv:1802.07539 [math.NA]. 2018. URL: https://arxiv.org/abs/1802.07539.

[34] W. Dörfler. "A Convergent Adaptive Algorithm for Poisson's Equation". In: *SIAM Journal on Numerical Analysis* 33(3) (1996), pp. 1106–1124. DOI: 10.1137/0733054.

[35] M. Eigel, C. Merdon, and J. Neumann. "An Adaptive Multilevel Monte Carlo Method with Stochastic Bounds for Quantities of Interest with Uncertain Data". In: *SIAM/ASA Journal on Uncertainty Quantification* 4(1) (2016), pp. 1219–1245. DOI: 10.1137/15M1016448.

[36] D. Elfverson, F. Hellman, and A. Målqvist. "A Multilevel Monte Carlo Method for Computing Failure Probabilities". In: *SIAM/ASA Journal on Uncertainty Quantification* 4(1) (2016), pp. 312–330. DOI: 10.1137/140984294.

[37] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. Chapman and Hall/CRC, 2015.

[38] T. Gerstner and S. Heinz. "Dimension- and Time-adaptive Multilevel Monte Carlo Methods". In: *Sparse Grids and Applications, Lecture Notes in Computational Science and Engineering, Volume 88*. Ed. by J. Garcke and M. Griebel. Springer, 2013, pp. 107–120.

[39] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. New York: Springer-Verlag, 1991. DOI: 10.1007/978-1-4612-3094-6.

[40] M. B. Giles. "Multilevel Monte Carlo methods". In: *Acta Numerica* 24 (2015), pp. 259–328. DOI: 10.1017/S096249291500001X.

[41] M. B. Giles. "Multilevel Monte Carlo Path Simulation". In: *Operations Research* 56(3) (2008), pp. 607–617. DOI: 10.1287/opre.1070.0496.

[42] I. G. Graham et al. "Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients". In: *Numerische Mathematik* 131(2) (2015), pp. 329–368. DOI: 10.1007/s00211-014-0689-y.

[43] C. Gräser. "Convex Minimization and Phase Field Models". PhD thesis. Freie Universität Berlin, 2011. URL: http://page.mi.fu-berlin.de/graeser/public_files/c_graeser_dissertation.pdf.

[44] C. Gräser and R. Kornhuber. "Multigrid Methods for Obstacle Problems". In: *Journal of Computational Mathematics* 27(1) (2009), pp. 1–44. URL: http://www.global-sci.org/jcm/openaccess/v27_1.pdf.

[45] C. Gräser and O. Sander. "The dune-subgrid module and some applications". In: *Computing* 86(269) (2009), pp. 269–290. DOI: 10.1007/s00607-009-0067-2.

[46] T. Grätsch and J. B. Klaus. "A posteriori error estimation techniques in practical finite element analysis". In: *Computers and Structures* 83(4-5) (2005), pp. 235–265. DOI: 10.1016/j.compstruc.2004.08.011.

[47] S. K. Gupta et al. "Review Article: Osteolysis After Total Knee Arthroplasty". In: *The Journal of Arthroplasty* 22(6) (2007), pp. 787–799. DOI: 10.1016/j.arth.2007.05.041.

[48] J. Gwinner and F. Raciti. "On a Class of Random Variational Inequalities on Random Sets". In: *Numerical Functional Analysis and Optimization* 27(5–6) (2006), pp. 619–636. DOI: 10.1080/01630560600790819.

[49] W. Hackbusch. *Elliptic Differential Equation: Theory and Numerical Treatment, volume 18 of Springer Series in Computational Mathematics*. Springer-Verlag, 1992.

[50] A.-L. Haji-Ali, F. Nobile, and R. Tempone. "Multi-index Monte Carlo: when sparsity meets sampling". In: *Numerische Mathematik* 132(4) (2016), pp. 767–806. DOI: 10.1007/s00211-015-0734-5.

[51] A.-L. Haji-Ali et al. "Optimization of mesh hierarchies in multilevel Monte Carlo samplers". In: *Stochastics and Partial Differential Equations: Analysis and Computations* 4 (2016), pp. 76–112. DOI: 10.1007/s40072-015-0049-7.

[52] J. Heinonen. *Lectures on Lipschitz Analysis*. URL: http://www.math.jyu.fi/research/reports/rep100.pdf.

[53] S. Heinrich. "Multilevel Monte Carlo Methods". In: *Lecture Notes in Computer Science, vol 2179*. Ed. by S. Margenov, J. Waśniewski, and P. Yalamov. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 3624–3651. DOI: 10.1007/3-540-45346-6_5.

[54] H. Hoel et al. "Adaptive Multilevel Monte Carlo Simulation". In: *Numerical Analysis of Multiscale Computations*. Ed. by B. Engquist, O. Runborg, and Y. H. Tsai. Springer, 2012, pp. 217–234. DOI: 10.1007/978-3-642-21943-6_10.

[55] H. Hoel et al. "Implementation and analysis of an adaptive multilevel Monte Carlo algorithm". In: *Monte Carlo Methods and Applications* 20(1) (2014), pp. 1–41. DOI: 10.1515/mcma-2013-0014.

[56] *ISO 14243-1: Implants for surgery – Wear of total knee-joint prostheses – Part 1: Loading and displacement parameters for wear-testing machines with load control and corresponding environmental conditions for test. International standard*. URL: https://www.iso.org/standard/44262.html.

[57] M. Karkulik, D. Pavlicek, and D. Praetorius. "On 2D Newest Vertex Bisection: Optimality of Mesh-Closure and H^1-Stability of L_2-Projection". In: *Constructive Approximation* 38(2) (2013), pp. 213–234. DOI: 10.1007/s00365-013-9192-4.

[58] C. Klapproth. "Adaptive Numerical Integration of Dynamical Contact Problems". PhD thesis. Freie Universität Berlin, 2011. URL: http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_000000009561/Diss_Klapproth_online.pdf.

[59] R. Kornhuber. "Monotone multigrid methods for elliptic variational inequalities I". In: *Numerische Mathematik* 69(2) (1994), pp. 167–184. DOI: 10.1007/BF03325426.

[60] R. Kornhuber, C. Schwab, and M.-W. Wolf. “Multilevel Monte Carlo Finite Element Methods for Stochastic Elliptic Variational Inequalities”. In: *SIAM Journal on Numerical Analysis* 52(3) (2014), pp. 1243–1268. DOI: 10.1137/130916126.

[61] R. Kornhuber and E. Youett. *Adaptive Multilevel Monte Carlo Methods for Stochastic Variational Inequalities*. arXiv:1611.06012 [math.NA]. To appear in SIAM Journal on Numerical Analysis. 2016. URL: https://arxiv.org/abs/1611.06012.

[62] C. Kreuzer and K. G. Siebert. “Decay rates of adaptive finite elements with Dörfler marking”. In: *Numerische Mathematik* 117(4) (2011), pp. 679–716. DOI: 10.1007/s00211-010-0324-5.

[63] F. Y. Kuo, C. Schwab, and I. H. Sloan. “Quasi-Monte Carlo Finite Element Methods for a Class of Elliptic Partial Differential Equations with Random Coefficient”. In: *SIAM Journal on Numerical Analysis* 50(6) (2012), pp. 3351–3374. DOI: 10.1137/110845537.

[64] F. Y. Kuo et al. “Multi-level Quasi-Monte Carlo Finite Element Methods for a Class of Elliptic Partial Differential Equations with Random Coefficient”. In: *Foundations of Computational Mathematics* 15(2) (2015), pp. 411–449. DOI: 10.1007/s10208-014-9237-5.

[65] F. Y. Kuo et al. “Multilevel Quasi-Monte Carlo methods for lognormal diffusion problems”. In: *Mathematics of Computation* 86 (2017), pp. 2827–2860. DOI: 10.1090/mcom/3207.

[66] J. Lang and R. Scheichl. *Adaptive Multilevel Stochastic Collocation Method for Randomized Elliptic PDEs*. Preprint Nr. 2718, Fachbereich Mathematik, TU Darmstadt. URL: http://www3.mathematik.tu-darmstadt.de/fb/mathe/preprints.html?eID=user_tudpreprints_pi3&user_tud_preprints_list%5Bview%5D=5278.

[67] O. P. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*. Scientific Computation. Dordrecht: Springer, 2010. DOI: 10.1007/978-90-481-3520-2.

[68] A. Liu and B. Joe. “Quality Local Refinement of Tetrahedral Meshes Based on Bisection”. In: *SIAM Journal on Scientific Computing* 16(6) (1994), pp. 1269–1291. DOI: 10.1137/0916074.

[69] J. Mandel. “A multilevel iterative method for symmetric, positive definite linear complementarity problems”. In: *Applied Mathematics and Optimization* 11(1) (1984), pp. 77–95. DOI: 10.1007/BF01442171.

[70] H. Matthies and A. Keese. “Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations”. In: *Computer Methods in Applied Mechanics and Engineering* 194(12-16) (2005), pp. 1295–1331. DOI: 10.1016/j.cma.2004.05.027.

[71] M. S. Mommer and R. Stevenson. “A Goal-Oriented Adaptive Finite Element Method with Convergence Rates”. In: *SIAM Journal on Numerical Analysis* 47(2) (2009), pp. 861–886. DOI: 10.1137/060675666.

[72] R. H. Nochetto, K. G. Siebert, and A. Veeser. “Theory of adaptive finite element methods: An introduction”. In: *Multiscale, Nonlinear and Adaptive Approximation: Dedicated to Wolfgang Dahmen on the Occasion of his 60th Birthday*. Ed. by R. DeVore and A. Kunoth. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 409–542. DOI: 10.1007/978-3-642-03413-8_12.

[73] A. Nowak. "Random Solutions of Equations". In: *Transactions of the 8th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes* B (1978), pp. 77–82.

[74] S. T. O'Brien et al. "An energy dissipation and cross shear time dependent computational wear model for the analysis of polyethylene wear in total knee replacements". In: *Journal of Biomechanics* 47(5) (2014), pp. 1127–1133. DOI: 10.1016/j.jbiomech.2013.12.017.

[75] J. T. Oden and S. Prudhomme. "Goal-oriented error estimation and adaptivity for the finite element method". In: *Computers and Mathematics with Applications* 41(5-6) (2001), pp. 735–756. DOI: 10.1016/S0898-1221(00)00317-5.

[76] M.C. Rivara. "Algorithms for refining triangular grids suitable for adaptive and multigrid techniques". In: *Numerical Methods in Engineering* 20(4) (1984), pp. 745–756. DOI: 10.1002/nme.1620200412.

[77] P. Robbe, D. Nuyens, and S. Vandewalle. "A Multi-Index Quasi-Monte Carlo Algorithm for Lognormal Diffusion Problems". In: *SIAM Journal on Scientific Computing* 39(5) (2017). DOI: 10.1137/16M1082561.

[78] O. Sander. "Multidimensional Coupling in a Human Knee Model". PhD thesis. Freie Universität Berlin, 2008. URL: http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_000000004627/dissertation_sander.pdf.

[79] O. Sander and A. Burchardt. *Dune-uggrid*. https://gitlab.dune-project.org/staging/dune-uggrid. 2018.

[80] C. Schwab and C. J. Gittelson. "Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs". In: *Acta Numerica* 20 (2011), pp. 291–467. DOI: 10.1017/S0962492911000055.

[81] K. G. Siebert and A. Veeser. "A Unilaterally Constrained Quadratic Minimization with Adaptive Finite Elements". In: *SIAM Journal on Optimization* 18(1) (2006), pp. 260–289. DOI: 10.1137/05064597X.

[82] R. Stevenson. "Optimality of a Standard Adaptive Finite Element Method". In: *Foundations of Computational Mathematics* 7(2) (2007), pp. 245–269. DOI: 10.1007/s10208-005-0183-0.

[83] R. Stevenson. "The Completion of Locally Refined Simplicial Partitions Created by Bisection". In: *Mathematics of Computation* 77(261) (2008), pp. 227–241. URL: http://www.jstor.org/stable/40234505.

[84] M. Straatman. "Random variables on non-separable Banach spaces". MA thesis. Mathematical Institute, University of Leiden, 2017. URL: https://www.universiteitleiden.nl/binaries/content/assets/science/mi/scripties/master/masterscriptie-mayke-straatman-7-juli-2017.pdf.

[85] T. J. Sullivan. *Introduction to Uncertainty Quantification*. New York: Springer, 2015. DOI: 10.1007/978-3-319-23395-6.

[86] A. L. Teckentrup. "Multilevel Monte Carlo methods and uncertainty quantification". PhD thesis. University of Bath, 2013. URL: http://www.maths.bath.ac.uk/~masrs/Teckentrup_PhD.pdf.

[87] A. L. Teckentrup et al. "Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients". In: *Numerische Mathematik* 125(3) (2013), pp. 569–600. DOI: 10.1007/s00211-013-0546-4.

[88] A.L. Teckentrup et al. “A Multilevel Stochastic Collocation Method for Partial Differential Equations with Random Input Data”. In: *SIAM/ASA Journal on Uncertainty Quantification* 3(1) (2015), pp. 1046–1074. DOI: 10.1137/140969002.

[89] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*. SIAM, 1992. DOI: 10.1137/1.9781611970692.

[90] R. Verfürth. *A Posteriori Error Estimation Techniques for Finite Element Methods, Numerical Mathematics and Scientific Computation*. Oxford Scholarship Online, 2013. DOI: 10.1093/acprof:oso/9780199679423.001.0001.

[91] A. Wächter and T. L. Biegler. “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming”. In: *Mathematical Programming* 106(1) (2006), pp. 25–57. DOI: 10.1007/s10107-004-0559-y.

[92] L. Wang. “Karhunen-Loeve Expansions and their applications”. PhD thesis. The London School of Economics and Political Science, 2008. URL: http://etheses.lse.ac.uk/2950/.

[93] B. Wohlmuth. “Variationally consistent discretization schemes and numerical algorithms for contact problems”. In: *Acta Numerica* 20 (2011), pp. 569–734. DOI: 10.1017/S0962492911000079.

[94] D. Xiu and G. E. Karniadakis. “Modeling uncertainty in flow simulations via generalized polynomial chaos”. In: *Journal of Computational Physics* 187(1) (2003), pp. 137–167. DOI: 10.1016/S0021-9991(03)00092-5.

[95] D. Xiu and G. E. Karniadakis. “The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations”. In: *SIAM Journal on Scientific Computing* 24(2) (2002), pp. 619–644. DOI: 10.1137/S1064827501387826.